GRAVITY AND DEFECTS
Gravity and Defects

Hagen Kleinert
Professor of Physics
Freie Universität Berlin

The sky seems so crystal clear
as the heat rays have now past,
I stare into the Milky Way
it just looks so big and vast.

Bernhard Howe, Our Night Skies
Preface

In 1922, Cartan extended Riemannian spacetime by torsion [1]. His work instigated Einstein to develop a theory of teleparallelism [2]. Twenty years later, Schrödinger attempted to relate torsion to electromagnetism [3]. He noticed that if torsion fields were present in the universe, this would make photons massive and limit the range of magnetic fields emerging from planets and stars. From the observed ranges he deduced upper bounds on the photon mass [4] which were, even then, extremely small. Further twenty years passed before Utiyama, Sciama, and Kibble [5, 6, 7] clarified the intimate relation between torsion and the spin density of the gravitational field. A detailed review of the theory was given in 1976 by Hehl et al. [3], which led to further investigations of the subject, summarized recently by Hammond [9].

I ran into the subject in the eighties when developing a disorder field theory of crystalline defects. Since the work of Kondo [10] and Kröner [11] it had become clear that dislocation and disclination lines in crystals may be viewed as singular lines of torsion and curvature in a flat space. This led to a geometric formulation of the theory of plastic deformations. Since Volterra, the theory of defects had been based on multivalued functions. The multivaluedness was made physically irrelevant by a local invariance called defect gauge invariance in the textbook I published in 1989 on this subject [12, 13]. The new geometric description had the advantage that only single-valued functions appeared. The defect gauge invariance of Volterra’s formulation reappeared as a local invariance of the geometric formulation under general coordinate transformations and local rotations. This invariance can again be reformulated as a gauge invariance. The resulting nonlinear gauge theory was simplified in a linearized approximation, called tangential approximation in the textbooks [12, 13], in which only simple abelian gauge degrees of freedom survive.

In the same approximation, the relation between the two formulations is completely analogous to the well-known relation between Maxwell’s theory of magnetism formulated in terms of a gauge field, the vector potential, and an alternative formulation in which the magnetic field is the gradient of a multivalued scalar field.

The latter analogy has a perfect defect correspondence in superfluids where the line-like defects are vortex lines. I have used this analogy to set up a field-theoretic formulation of the statistical mechanics of vortex lines in superfluids and superconductors in the textbook [12]. The book explains the properties of the phase transitions from the low-temperature to the normal phase in these systems by the
increase of the configurational entropy of these lines with temperature. For the quantitative formulation I set up a disorder field theory of the vortex lines. The interactions between them were described by a gauge field coupled to the disorder field.

It was a fortunate coincidence that I was also searching, as many other people did at that time, for a simple field-theoretic formulation of the phenomenon of quark confinement. In that context, I took advantage of the analogy of the above structures with the situation in Dirac’s theory of magnetic monopoles. Similar to Volterra’s construction, Dirac used a multivalued vector potential to introduce an infinitely thin magnetic flux tube to simulate a magnetic point source at its end. In this way, the world line of a monopole in spacetime could be viewed as a defect line of Maxwell’s theory. It was known that a condensate of monopoles would create an environment in which the electric flux lines between charges would be pressed into thin tubes. This would naturally create a potential proportional to the distance, thus confining electric charges. The analogy with the theory of defect and vortex lines led to a disorder field theory of monopoles and a simple theory of quark confinement [14].

When extending the statistical mechanics of vortex lines to defect lines in the second volume of the textbook [13], I naturally ran into the geometry of defects. This made me realize that reversing the development in the theory of defects from Volterra to Kondo and Kröner, the geometrical basis for a theory of gravity with torsion could be formulated in a very intuitive way with the help of multivalued translation and rotation fields. In the Volterra theory, these transformations are used to carry an ideal crystal into crystals with translational and rotational defects. Their geometric analogs carry a flat spacetime into a spacetime with curvature and torsion. The mathematical basis expressing the new geometry are multivalued tetrad fields $e^a_{\mu}(x)$. 

In the traditional literature on gravity with spinning particles, a special role is played by single-valued vierbein fields $h^a_{\mu}(x)$ which define local nonholonomic coordinate differentials $dx^a$ which can be reached from the physical coordinate differentials $dx^\mu$ by transformations $dx^a = h^a_{\mu}(x)dx^\mu$. Only infinitesimal vectors $dx^a$ are defined, and the transformation cannot be extended over finite domains, since it is nonholonomic. Such an extension is unnecessary, however, since the infinitesimal nonholonomic coordinates $dx^a$ are completely sufficient to specify the behavior of spinning particles in a Riemannian spacetime.

The theory presented here goes an important step further, leading to a drastic simplification of the description of the geometry. This becomes possible by an efficient use of a set of nonholonomic coordinates $dx^a$ which are more nonholonomic than the traditional $dx^a$ used to describe particles with spin. To emphasize this one might call them hyper-nonholonomic coordinates. They are related to $dx^a$ by a multivalued Lorentz transformation $dx^a = \Lambda^a_{\alpha}(x)dx^\alpha$, and to the physical $dx^\mu$ by a the above-mentioned multivalued tetrad fields as $dx^a = e^a_{\mu}(x)dx^\mu \equiv \Lambda^a_{\alpha}(x)h^a_{\mu}(x)dx^\mu$. The gradients $\partial_{\mu}e^a_{\nu}(x)$ determine directly the full affine connection, their antisymmetric combination $\partial_{\mu}e^a_{\nu}(x) - \partial_{\nu}e^a_{\mu}(x)$ yields the torsion. This is in contrast to the
curl of the usual vierbein fields $h^\alpha_\mu(x)$ which determines the object of anholonomy, a quantity existing also in a purely Riemannian spacetime, i.e., in the absence of torsion.

The purpose of this book is to make students and colleagues working in gravitational physics appreciate the many advantages brought about by the use of the multivalued tetrad fields $e^a_\mu(x)$. Apart from a simple intuitive reformulation of Riemann-Cartan geometry, it suggests a new principle in physics [15], which we have called the nonholonomic mapping principle, to be explained in detail in this book. Nonholonomic coordinate transformations enable us to transform physical laws from flat spacetime to spacetimes with curvature and torsion. It is therefore natural to postulate that the image laws describe correctly the physics in such general affine spacetimes. As a result we are able to make predictions which cannot be made with Einstein’s construction method based merely on covariance under ordinary coordinate transformations which cannot connect different geometries.

It should be emphasized that it is not the purpose of this book to propose creating all geometries studied in gravitational theories with the help of nonholonomic coordinate transformations. In fact, we shall restrict much of the discussion to almost flat auxiliary spacetimes. This will be enough to derive the general form of the physical laws the presence of curvature and torsion. At the end we may always return to the usual geometric description. The intermediate auxiliary spacetime with defects will be referred to as world crystal.

The reader will be pleased to see in Subsection 4.5 that the standard minimal coupling of electromagnetism is a simple consequence of our nonholonomic mapping principle. The similar minimal coupling to gravity will be derived from this principle in Sections 14 and 17.

At the end we shall argue that torsion fields in gravity, if they exists, would lead quite a hidden life, unless they are of a special form. They would not be observable for many generations to come since they could exist only in an extremely small neighborhood of material point particles, limited to distances of the order of the Planck length $10^{-33}$ cm, which no presently conceivable experiment can probe. As such, a textbook on gravity containing torsion fields such as this is somewhat academic from the practical point of view. Its main merit should lie in the mathematical virtues of the nonholonomic approach to Riemann-Cartan geometry.

Special thanks go to my wife Dr. Annemarie Kleinert for her sacrifices, inexhaustible patience, and constant encouragement.

H. Kleinert  
Berlin, September 2004
Notes and References


# Contents

Notes and References ........................................... x

1 Basics ..................................................................... 1
  1.1 Galilean Invariance of Newtonian Mechanics ................. 1
    1.1.1 Translations .................................................. 1
    1.1.2 Rotations ...................................................... 2
    1.1.3 Galilei Boosts ................................................ 3
    1.1.4 Galilei Group ............................................... 3
  1.2 Lorentz Invariance of Maxwell Equations ....................... 3
    1.2.1 Lorentz Boosts .............................................. 4
    1.2.2 Lorentz Group ............................................. 6
  1.3 Infinitesimal Lorentz Transformations ............................. 6
    1.3.1 Generators of Group Transformations ..................... 7
    1.3.2 Group Multiplication and Lee Algebra .................. 9
  1.4 Vectors, Tensors, Scalars ........................................ 11
    1.4.1 Discrete Lorentz Transformations ......................... 13
    1.4.2 Poincaré group ............................................. 14
  1.5 Differential Operators for Lorentz Transformations .......... 14
  1.6 Vector and Tensor Operators ................................... 16
  1.7 Finite Operator Lorentz Transformations ....................... 16
    1.7.1 Rotations .................................................... 16
    1.7.2 Lorentz Boosts ............................................ 17
    1.7.3 Lorentz Group ............................................. 18
  1.8 Relativistic Point Mechanics .................................... 19
  1.9 Quantum Mechanics ............................................. 22
  1.10 Relativistic Particles with Electromagnetic Interactions .... 24
  1.11 Dirac Field ....................................................... 29
  1.12 Spacetime-Dependent Lorentz Transformations ............... 31
    1.12.1 Angular Velocities ...................................... 31
    1.12.2 Angular Gradients ..................................... 33
  1.13 Energy-Momentum Tensors ..................................... 33
    1.13.1 Point Particles ........................................... 33
    1.13.2 Electromagnetic Field ................................... 35
  1.14 Angular Momentum and Spin ................................... 37
  1.15 Energy-Momentum Tensor of Perfect Fluid .................... 42
Appendix 1A Tensor Identities ................................... 43
Notes and References

1A.1 Product Formulas ................................................. 44
1A.2 Determinants ......................................................... 45
1A.3 Expansion of Determinants ..................................... 46
Notes and References .................................................. 46

2 Action Approach ....................................................... 48
2.1 General Particle Dynamics ......................................... 49
2.2 Single Relativistic Particle ........................................ 50
2.3 Scalar Fields .......................................................... 52
  2.3.1 Locality ............................................................ 52
  2.3.2 Lorentz Invariance ............................................... 53
  2.3.3 Field Equations .................................................. 54
  2.3.4 Plane Waves ...................................................... 55
  2.3.5 Schrödinger Quantum Mechanics as Nonrelativistic Limit .... 55
  2.3.6 Natural Units ...................................................... 56
  2.3.7 Hamiltonian Formalism .......................................... 57
  2.3.8 Conserved Current ............................................... 57
2.4 Maxwell’s Equation from Extremum of Field Action ............... 59
  2.4.1 Electromagnetic Field Action ................................... 59
  2.4.2 Alternative Action for Electromagnetic Field ................. 61
  2.4.3 Hamiltonian of Electromagnetic Fields ......................... 61
  2.4.4 Gauge Invariance of Maxwell’s Theory ......................... 63
2.5 Maxwell-Lorentz Action for Charged Point Particles ............... 65
2.6 Scalar Field with Electromagnetic Interaction .................... 66
2.7 Dirac Fields .......................................................... 67
Notes and References .................................................. 69

3 Continuous Symmetries and Conservation Laws. Noether’s Theorem 70
3.1 Continuous Symmetries and Conservation Laws .................... 70
  3.1.1 Group Structure of Symmetry Transformations .................. 70
  3.1.2 Substantial Variations .......................................... 71
  3.1.3 Conservation Laws ............................................... 71
  3.1.4 Alternative Derivation of Conservation Laws .................. 72
3.2 Time Translation Invariance and Energy Conservation ............. 74
3.3 Momentum and Angular Momentum .................................. 75
  3.3.1 Translational Invariance in Space ............................. 76
  3.3.2 Rotational Invariance .......................................... 76
  3.3.3 Center-of-Mass Theorem ........................................ 78
  3.3.4 Conservation Laws from Lorentz Invariance .................... 79
3.4 Generating the Symmetries ......................................... 81
3.5 Field Theory .......................................................... 83
  3.5.1 Continuous Symmetry and Conserved Currents ................. 83
  3.5.2 Alternative Derivation .......................................... 84
### 3.5.3 Local Symmetries ................................. 85
### 3.6 Canonical Energy-Momentum Tensor .................. 87
#### 3.6.1 Electromagnetism ............................... 89
#### 3.6.2 Dirac Field .................................. 90
### 3.7 Angular Momentum ................................ 91
### 3.8 Four-Dimensional Angular Momentum ................. 93
### 3.9 Spin Current ...................................... 94
#### 3.9.1 Electromagnetic Fields ........................ 95
#### 3.9.2 Dirac Field .................................. 98
### 3.10 Symmetric Energy-Momentum Tensor ................ 99
### 3.11 Internal Symmetries ................................ 101
#### 3.11.1 U(1)-Symmetry and Charge Conservation ....... 102
#### 3.11.2 Broken Internal Symmetries ..................... 103
### 3.12 Generating the Symmetry Transformations on Quantum Fields .......... 103
### 3.13 Energy-Momentum Tensor of Relativistic Massive Point Particle .... 104
### 3.14 Energy-Momentum Tensor of Massive Charged Particle in Electromagnetic Field .................................. 106
### Notes and References .................................. 109

### 4 Multivalued Gauge Transformations in Magnetostatics 110
#### 4.1 Vector Potential of Current Distribution .............. 110
#### 4.2 Multivalued Gradient Representation of Magnetic Field ........ 111
#### 4.3 Generating Magnetic Fields by Multivalued Gauge Transformations 118
#### 4.4 Magnetic Monopoles ................................ 119
#### 4.5 Minimal Magnetic Coupling of Particles from Multivalued Gauge Transformations .................................. 122
#### 4.6 Equivalence of Multivalued Scalar and Single-Valued Vector Potential Representation .................. 124
#### 4.7 Multivalued Field Theory of Magnetic Monopoles and Electric Currents .................................. 126
### Notes and References .................................. 128

### 5 Multivalued Fields in Superfluids and Superconductors 130
#### 5.1 Superfluid Transition ............................... 130
##### 5.1.1 Configuration Entropy .......................... 132
##### 5.1.2 Origin of Massless Excitations .................. 134
##### 5.1.3 Vortex Density ................................ 136
##### 5.1.4 Partition Function ............................... 137
##### 5.1.5 Continuum Derivation of Interaction Energy .......... 142
##### 5.1.6 Physical Jumping Surfaces ....................... 143
##### 5.1.7 Canonical Representation of Superfluid ............ 144
##### 5.1.8 Yukawa Loop Gas ................................ 148
##### 5.1.9 Gauge Field of Superflow ........................ 149
##### 5.1.10 Disorder Field Theory ........................... 151

H. Kleinert, GRAVITY WITH TORSION
8 Relativistic Magnetic Monopoles and Electric Charge Confinement 249
8.1 Monopole Gauge Invariance 249
8.2 Charge Quantization 253
8.3 Electric and Magnetic Current-Current Interactions 254
8.4 Dual Gauge Field Representation 256
8.5 Monopole Gauge Fixing 258
8.6 Quantum Field Theory of Spinless Electric Charges 259
8.7 Theory of Magnetic Charge Confinement 260
8.8 Second Quantization of the Monopole Field 262
8.9 Quantum Field Theory of Electric Charge Confinement 264
Notes and References 268

9 Multivalued Mapping from Ideal Crystals to Crystals with Defects 272
9.1 Defects 272
9.2 Dislocation Lines and Burgers Vector 276
9.3 Disclination Lines and Frank Vector 279
9.4 Interdependence of Dislocation and Disclinations 281
9.5 Defect Lines with Infinitesimal Discontinuities in Continuous Media 283
9.6 Multivaluedness of Displacement Field 284
9.7 Smoothness Properties of Displacement Field and Weingarten’s Theorem 285
9.8 Integrability Properties of Displacement Field 288
9.9 Dislocation and Disclination Densities 290
9.10 Mnemonic Procedure for Constructing Defect Densities 293
9.11 Defect Gauge Invariance 296
9.12 Branching Defect Lines 297
9.13 Defect Density and Incompatibility 298
Notes and References 303

10 Defect Melting 305
10.1 Specific Heat 305
10.2 Elastic and Plastic Energies 306
Notes and References 310

11 Relativistic Mechanics in Curvilinear Coordinates 312
11.1 Equivalence Principle 312
11.2 Free Particle in General Coordinates Frame 313
11.3 Minkowski Geometry formulated in General Coordinates 316
11.3.1 Local Basis tetrads 317
11.3.2 Vector- and Tensor Fields, and Lorentz Invariance 319
11.3.3 Affine Connections and Covariant Derivatives 323
11.4 Covariant Time Derivative and Acceleration 326
11.5 Torsion tensor 326
11.6 Curvature Tensor as Covariant Curl of Affine Connection 328
11.7 Riemann Curvature Tensor 332
Appendix 11A Curvilinear Versions of Levi-Civita Tensor 334
Notes and References 337

12 Torsion and Curvature from Defects 338
12.1 Multivalued Infinitesimal Coordinate Transformations 339
12.2 Examples for Nonholonomic Coordinate Transformations 344
12.2.1 Dislocation 344
12.2.2 Disclination 346
12.3 Differential-Geometric Properties of Affine Spaces 347
12.3.1 Integrability of Metric and Affine Connection 347
12.3.2 Local Parallelism 348
12.4 Circuit Integrals in Affine Spaces with Curvature and Torsion 351
12.4.1 Closed Contour Integral over Parallel Vector Field 351
12.4.2 Closed Contour Integral over Coordinates 352
12.4.3 Closure Failure and Burgers Vector 353
12.4.4 Alternative Circuit Integral for Curvature 354
12.4.5 Parallelism in World Crystal 355
12.5 Bianchi Identities for Curvature and Torsion Tensors 355
12.6 Special Coordinates in Riemann Spacetime 358
12.6.1 Geodesic Coordinates 358
12.6.2 Canonical Geodesic Coordinates 359
12.6.3 Harmonic Coordinates 362
12.6.4 Coordinates with det($g_{\mu\nu}$) = 1 363
12.6.5 Orthogonal Coordinates 363
12.7 Number of Independent Components of $R_{\mu\nu\lambda}$ and $S_{\mu\nu}^{\lambda}$ 365
12.7.1 Two Dimensions 365
12.7.2 Three Dimensions 367
12.7.3 Four or More Dimensions 367
12.7.4 Constant Curvature 369
Notes and References 370

13 Curvature and Torsion from Embedding 372
13.1 Curvature 372
13.2 Torsion 374
13.2.1 Strategy 374
13.2.2 Nonholonomic Embedding 375
13.2.3 Torsion 377
Notes and References 377
### 14 Nonholonomic Mapping Principle

14.1 Motion of Point Particle ............................................. 379
14.1.1 Classical Action Principle for Spaces with Curvature .......... 379
14.1.2 Autoparallel Trajectories in Spaces with Torsion .......... 380
14.1.3 Special Properties of Gradient Torsion ...................... 385
14.2 Autoparallel Trajectories from Embedding ........................ 387
14.2.1 Special Role of Autoparallels ................................. 387
14.2.2 Gauss Principle of Least Constraint ......................... 387
14.3 Maxwell-Lorentz Orbits as Autoparallel Trajectories .......... 389
14.4 Bargmann-Michel-Telegdi Equation from Torsion .......... 389

Notes and References ................................................. 390

### 15 Field Equations of Gravitation

15.1 Invariant Action ...................................................... 391
15.2 Energy-Momentum Tensor and Spin Density ........................ 393
15.3 Total Energy-Momentum Tensor and Defect Density ............ 399

Notes and References ................................................. 400

### 16 Minimally Coupled Fields of Integer Spin

16.1 Scalar Fields in Riemann-Cartan Space .......................... 401
16.2 Electromagnetism in Riemann-Cartan Space ..................... 404

Notes and References ................................................. 405

### 17 Particles with Half-Integer Spin

17.1 Local Lorentz Invariance and Anholonomic Coordinates .......... 406
17.1.1 Nonholonomic Image of Dirac Action ........................ 407
17.1.2 Vierbein Fields .................................................. 409
17.1.3 Local Inertial Frames .......................................... 410
17.1.4 Difference between Vierbein and Multivalued Tetrad Fields 412
17.2 Dirac Action in Riemann-Cartan Space ........................... 416
17.3 Ricci Identity ....................................................... 417
17.4 Alternative Form of Coupling ..................................... 418
17.5 Invariant Action for Vector Fields ................................ 419
17.6 Verifying Local Lorentz Invariance ................................ 421
17.7 Field Equations with Gravitational Spinning Matter .......... 422

Notes and References ................................................. 426

### 18 Covariant Conservation Law

18.1 Spin Density .......................................................... 427
18.2 Energy-Momentum Density ......................................... 429
18.3 Covariant Derivation of Conservation Laws .................... 432
18.4 Matter with Integer Spin ........................................... 433
18.5 Relations between Conservation Laws and Bianchi Identities .... 435
18.6 Particle Trajectories from Energy-Momentum Conservation .... 436

H. Kleinert, GRAVITY WITH TORSION
Notes and References ................................................. 438

19 Gravitation of Spinning Matter as a Gauge Theory 439
  19.1 Local Lorentz Transformations ............................... 439
  19.2 Local Translations ........................................... 441
Notes and References ................................................. 442

20 Linearized Einstein Gravity of Point Particles 443
  20.1 Action and Equation of Motion ................................. 443
  20.2 Gauge Invariance .............................................. 447
  20.3 Newton's Field around a Mass Point at Rest ................ 449
  20.4 Newtonian Limit of Particle Motion ......................... 450
  20.5 The Cosmological Term ....................................... 451
  20.6 Energy-Momentum Tensors in Linearized Gravity .......... 453
  20.7 Relation between Canonical Energy-Momentum Tensor and Einstein Tensor .............................................. 454
  20.8 First-Order Linear Correction in Symmetric Space ......... 456
  20.9 Weak Gravitational Field of a Spinning Star; Lense-Thirring Effect .............................................. 457
  20.10 Free Fall of a Spinning Point Particle ..................... 460
  20.11 Precession of Spinning Planet due to Tidal Forces ....... 461
  20.12 Precession of Gyroscope in a Satellite Orbit ............. 462
    20.12.1 Geodetic Precession ...................................... 462
    20.12.2 Lense-Thirring of Frame-Dragging Precession .......... 465
  20.13 Torsion in Linearized Theory ................................ 466
Notes and References ................................................. 467

21 Experimental Tests in Weak Field 468
  21.1 Gravitational Red Shift ....................................... 468
  21.2 Deflection of Light Grazing the Sun ......................... 470
  21.3 Time Delay of Light Grazing the Sun ......................... 474
  21.4 Neighbouring Trajectories and Tidal Forces ................. 475
    21.4.1 Simplification .............................................. 476
    21.4.2 Weak-Field Limit .......................................... 476
Notes and References ................................................. 477

22 Simple Exact Solutions of Einstein Equations 479
  22.1 Static Spherically Symmetric Coordinates .................. 479
  22.2 The Schwarzschild Solution ................................... 483
    22.2.1 Standard Coordinate System .............................. 483
    22.2.2 Spatially Isotropic Coordinate System ................ 486
  22.3 Planetary Motion in Schwarzschild Metric .................. 487
  22.4 Orbit of Light Ray ............................................ 493
  22.5 Schwarzschild Singularity .................................... 494
  22.6 Black Holes .................................................... 494
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Infinitesimally thin closed current loop $L$ and magnetic field</td>
<td>112</td>
</tr>
<tr>
<td>4.2</td>
<td>Single- and multi-valued definitions of arctan $\varphi$</td>
<td>116</td>
</tr>
<tr>
<td>5.1</td>
<td>Specific heat of superfluid $^4$He</td>
<td>130</td>
</tr>
<tr>
<td>5.2</td>
<td>Energies of the elementary excitations in superfluid $^4$He</td>
<td>131</td>
</tr>
<tr>
<td>5.3</td>
<td>Rotons join side by side to form surfaces whose boundary appears as a large vortex loop</td>
<td>131</td>
</tr>
<tr>
<td>5.4</td>
<td>Vortex loops in XY-model for different $\beta = 1/k_B T$</td>
<td>133</td>
</tr>
<tr>
<td>5.5</td>
<td>Lattice Yukawa potential at the origin and the associated trace-log</td>
<td>140</td>
</tr>
<tr>
<td>5.6</td>
<td>Specific heat of Villain model in three dimensions</td>
<td>142</td>
</tr>
<tr>
<td>5.7</td>
<td>Critical temperature of a loop gas with Yukawa interactions</td>
<td>148</td>
</tr>
<tr>
<td>5.8</td>
<td>Specific heat of superconducting aluminum</td>
<td>155</td>
</tr>
<tr>
<td>5.9</td>
<td>Potential for the order parameter $\rho$ with cubic term</td>
<td>176</td>
</tr>
<tr>
<td>5.10</td>
<td>Phase diagram of a two-dimensional layer of superfluid $^4$He</td>
<td>181</td>
</tr>
<tr>
<td>5.11</td>
<td>Experimental phase diagram of a two-dimensional layer of superfluid $^4$He by $^3$He</td>
<td>181</td>
</tr>
<tr>
<td>5.12</td>
<td>Order parameter $\bar{\rho} =</td>
<td>\phi</td>
</tr>
<tr>
<td>7.1</td>
<td>Energy gap of superconductor as a function of temperature</td>
<td>221</td>
</tr>
<tr>
<td>7.2</td>
<td>Energies of the low-energy excitations in superconductor</td>
<td>224</td>
</tr>
<tr>
<td>7.3</td>
<td>Contour plot of zeros for energy eigenvalues in superconductor</td>
<td>225</td>
</tr>
<tr>
<td>7.4</td>
<td>Temperature behavior of superfluid density $\rho_s/\rho = \phi(\Delta/T)$ (Yoshida function) and the gap function $\tilde{\rho}_s/\rho = \tilde{\pi}(\Delta/T)$</td>
<td>230</td>
</tr>
<tr>
<td>7.5</td>
<td>Temperature behavior of the functions governing the kinetic term of the pair field in the BCS superconductor</td>
<td>232</td>
</tr>
<tr>
<td>7.6</td>
<td>Spatial variation of order parameter $\rho$ and magnetic field $H$ in the neighborhood of a planar domain wall between normal and superconducting phases $N$ and $S$</td>
<td>241</td>
</tr>
<tr>
<td>7.7</td>
<td>Order parameter $\rho$ and magnetic field $H$ for a vortex line</td>
<td>244</td>
</tr>
<tr>
<td>7.8</td>
<td>Critical field $H_{c1}$ as a function of the parameter $\kappa$</td>
<td>245</td>
</tr>
<tr>
<td>7.9</td>
<td>Lines of equal size of order parameter $\rho(\mathbf{x})$ in a typical mixed state in which the vortex lines form a hexagonal lattice</td>
<td>246</td>
</tr>
<tr>
<td>7.10</td>
<td>Temperature behavior of the critical magnetic fields of a type-II superconductor</td>
<td>247</td>
</tr>
<tr>
<td>7.11</td>
<td>Magnetization curve as a function of the external magnetic field</td>
<td>247</td>
</tr>
</tbody>
</table>
22.1 Orbit of Mercury in Newton’s and Einstein’s theory ................. 493
22.2 Maximal Schwarzschild geometry in Kruskal coordinates .......... 496
22.3 Spatial surface $\theta = \pi/2$ at $t = 0$ for the Schwarzschild solution:
    Schwarzschild throat ........................................... 498
22.4 Evolution of the Schwarzschild geometry .................................. 498

23.1 Energy levels of Dirac particle in gravitational field of an incom-
pressible star .......................................................... 507
23.2 of raising and lowering operators $\hat{L}_+$ and $\hat{L}_-$ upon the states $|s, m\rangle$ 512

24.1 Distortions of a circular array of mass points by the passage of a
    gravitational quadrupole wave .................................... 533
24.2 Field lines of tidal forces of a gravitational wave ....................... 534
24.3 Two equal masses $M$ oscillating at the ends of a spring as a source
    of gravitational radiation ......................................... 545
24.4 Two spherical masses in circular orbits about their center of mass . 547
24.5 Gravitational Amplitudes arriving at the earth from possible sources 549
24.6 Two pulsars orbiting around each other .................................. 550
24.7 Shift of time of periastron passage of PSR 1913+16 ..................... 551
24.8 Two masses in a Keplerian orbit around the common center-of-mass . 551
24.9 Spectrum of the gravitational radiation emitted by a particle of mass
    $m$ falling radially into a black hole of mass $M$ .................. 554
24.10 Particle falling radially toward a mass .................................... 554

28.1 Potential of closed Friedman universe as a function of the radius
    $a/a_{\text{max}}$ .......................................................... 598
28.2 Radius of universe as a function of time in Friedman universe ...... 599
1

Basics

In his fundamental work on theoretical mechanics entitled *Principia*, Newton (1642-1727) assumes the existence of an absolute spacetime. Space is parametrized by vectors $\mathbf{x} = (x^1, x^2, x^3)$, and the movement of point particles is described by trajectories $\mathbf{x}(t)$ whose components $q^i(t) (i = 1, 2, 3)$ specify the coordinates $x^i = q^i(t)$ along which the particles move as a function of the time $t$. In Newton’s absolute spacetime, a single free particle moves without acceleration. Mathematically, this is expressed by the differential equation

$$\ddot{\mathbf{x}}(t) \equiv \frac{d^2}{dt^2} \mathbf{x}(t) = 0. \quad (1.1)$$

The dots denote derivatives with respect to the argument.

A set of $N$ point particles $\mathbf{x}_n(t)$ ($n = 1, \ldots, N$) with masses $m_n$ is subject to gravitational forces which change the free equations of motion to

$$m_n \ddot{\mathbf{x}}_n(t) = G_N \sum_{m \neq n} m_n m_m \frac{\mathbf{x}_m(t) - \mathbf{x}_n(t)}{|\mathbf{x}_m(t) - \mathbf{x}_n(t)|^3}, \quad (1.2)$$

where $G_N$ is Newton’s gravitational constant

$$G_N \approx 6.67259(85) \times 10^{-8} \text{cm}^3/\text{g sec}^2. \quad (1.3)$$

### 1.1 Galilean Invariance of Newtonian Mechanics

The parametrization of absolute spacetime in which the above equations of motion hold is not unique.

#### 1.1.1 Translations

The coordinates $\mathbf{x}$ may always be changed by translated coordinates

$$\mathbf{x}' = \mathbf{x} - \mathbf{x}_0. \quad (1.4)$$
It is obvious that the translated trajectories $\mathbf{x}'_n(t) = \mathbf{x}_n(t) - \mathbf{x}_0$ will again satisfy the equations of motion (1.2). The equations remain also true for a translated time

$$t' = t - t_0,$$

i.e., the trajectories

$$\mathbf{x}'(t) \equiv \mathbf{x}(t + t_0)$$

satisfy (1.2). This property of Newton’s equations (1.2) is referred to as translational symmetry in spacetime.

An alternative way of formulating this invariance is my keeping the coordinate frame fixed and displacing the physical system in spacetime, moving all particles to new coordinates $\mathbf{x}' = \mathbf{x} + \mathbf{x}_0$ at a new time $t' = t + t_0$. The equations of motion are again invariant. The first procedure of reparametrizing the same physical system is called passive symmetry transformation, the second active symmetry transformation. One may use either procedure to discuss symmetries. In this book we shall use active or passive transformations, depending on the circumstance.

### 1.1.2 Rotations

The equations of motion are invariant under more transformations which mix different coordinates linearly with each other, for instance the rotations:

$$x'^i = R^i_j x^j$$

where $R^i_j$ is the rotation matrix

$$R^i_j = \cos \theta \delta_{ij} + (1 - \cos \theta) \hat{\theta}_i \hat{\theta}_j + \sin \theta \epsilon_{ijk} \hat{\theta}_k,$$

in which $\hat{\theta}_i$ denotes the directional unit vector of the rotation axis. The matrices satisfy the orthogonality relation

$$R^T R = 1.$$

In Eq. (1.7) a sum from 1 to 3 is implied over the repeated spatial index $j$. This is called the Einstein summation convention, which will be followed throughout this text. As for the translations, the rotations can be applied in the passive or active sense.

The active rotations are obtained from the above passive ones by changing the sign of $\theta$. For example, the active rotations around the $z$-axis with a rotation vector $\hat{\varphi} = (0, 0, 1)$ are given by the orthogonal matrices

$$R_3(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
1.1.3 Galilei Boosts

A further set of transformations mixes space and time coordinates:

\[ x'^i = x^i - v^i t, \]  
\[ t' = t. \]  

These are called pure Galilei transformations of Galilei boosts. The coordinates \( x'^i, t' \) are positions and time of a particle observed in a frame of reference that moves uniformly through absolute spacetime with velocity \( \mathbf{v} \equiv (v^1, v^2, v^3) \). In the active description, the transformation \( x'^i = x^i + v^i t \) specifies the coordinates of a physical system moving past the observer with uniform velocity \( \mathbf{v} \).

1.1.4 Galilei Group

The combined set of all transformations

\[ x'^i = R^i_{\ j} x^j - v^i t - x'^i, \]  
\[ t' = t - t_0, \]  

forms a group. Group multiplication is defined by performing the transformations successively. This multiplication law is obviously associative, and each element has an inverse. The set of transformations (1.13) and (1.14) is referred to as the Galilei group.

Newton called all coordinate frames in which the equations of motion have the simple form (1.2) inertial frames.

1.2 Lorentz Invariance of Maxwell Equations

Problems with Newton’s theory arose when J. C. Maxwell (1831 - 1879) formulated in 1864 his theory of electromagnetism. His equations for the electric field \( \mathbf{E}(x) \) and the magnetic flux density or magnetic induction \( \mathbf{B}(x) \) in empty space

\[ \nabla \cdot \mathbf{E} = 0 \] (Coulomb’s law),  
\[ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = 0 \] (Ampre’s law),  
\[ \nabla \cdot \mathbf{B} = 0 \] (absence of magnetic monopoles),  
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \] (Faraday’s law),

can be combined to obtain the second-order differential equations

\[ \left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{E}(x,t) = 0, \]  
\[ \left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{B}(x,t) = 0. \]
The equations contain explicitly the light velocity

$$c \equiv 299 792 458 \frac{m}{\text{sec}}, \quad (1.21)$$

and are not invariant under the Galilei group (1.14). Indeed, they contradict Newton’s postulate of the existence of an absolute spacetime. If light were to propagate with the velocity $c$ in absolute spacetime, it could not do so in other inertial frames which have a nonzero velocity with respect to the absolute frame. A precise measurement of the light velocity could therefore single out the absolute spacetime. However, experimental attempts to do this did not succeed. The experiment of Michelson (1852-1931) and Morley (1838-1923) in 1887 showed [1] that light travels parallel and orthogonal to the earth’s orbital motion with the same velocity up to $\pm 5$ km/sec [2]. This led Fitzgerald (1851-1901) [3], Lorentz (1855-1928) [4], Poincaré (1854-1912) [5], and finally Einstein (1879-1955) [6] to suggest that Newton’s postulate of the existence of an absolute spacetime was unphysical [7].

### 1.2.1 Lorentz Boosts

The conflict was resolved by modifying the Galilei transformations (1.11) and (1.12) in such a way that Maxwell’s equations remain invariant. This is achieved by the coordinate transformations

$$x'^i = x^i + (\gamma - 1) \frac{v^i v^j}{v^2} x^j - \gamma v^i t, \quad (1.22)$$

$$t' = \gamma t - \frac{1}{c^2} \gamma v^i x^i, \quad (1.23)$$

where $\gamma$ is the velocity-dependent parameter

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}. \quad (1.24)$$

The transformations (1.23) are referred to as pure Lorentz transformations or Lorentz boosts. The parameter $\gamma$ has the effect that in different moving frames of reference, time elapses differently. This is necessary to make the light velocity the same in all frames.

The pure Lorentz transformations are conveniently written in a four-dimensional vector notation. Introducing the four-vectors $x^a$ labeled by letters $a,b,c,\ldots$ running through the values $0,1,2,3$,

$$x^a = \begin{pmatrix} ct \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \quad (1.25)$$

we rewrite (1.22) and (1.23) as

$$x'^a = \Lambda^a_b x^b \quad (1.26)$$

H. Kleinert, GRAVITY WITH TORSION
where $\Lambda^a_b$ are the $4 \times 4$-matrices
\[
\Lambda^a_b \equiv \begin{pmatrix} \gamma & -\gamma v^i/c \\ -\gamma v^i/c & \delta_{ij} + (\gamma - 1) v_i v_j/v^2 \end{pmatrix}.
\] (1.27)

Note that we adopt Einstein’s summation convention also for repeated labels $a, b, c, \ldots = 0, \ldots, 3$. The matrices $\Lambda^a_b$ satisfy the pseudo-orthogonality relation [compare (1.9)]:
\[
\Lambda^T_a c g_{cd} \Lambda^d_b = g_{ab},
\] (1.28)
where $g_{ab}$ is the Minkowski metric with the matrix elements
\[
g_{ab} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.
\] (1.29)

Equation (1.28) has the consequence that for any two four-vectors $x^a$ and $y^a$, the scalar product formed with the help of the Minkowski metric
\[
x y \equiv x^a g_{ab} y^b
\] (1.30)
is an invariant under Lorentz transformation.

In order to verify the relation (1.28) it is convenient to introduce a dimensionless vector $\zeta$, called rapidity, which points in the direction of the velocity $v$ and has a length $\zeta \equiv |\zeta|$ given by
\[
cosh \zeta = \gamma, \quad \sinh \zeta = \gamma v/c.
\] (1.31)

We also define the unit vectors in three-space
\[
\hat{\zeta} \equiv \zeta/\zeta = \hat{v} \equiv v/v,
\] (1.32)
so that
\[
\zeta = \zeta \hat{\zeta} = \text{atanh} \frac{v}{c} \hat{v}.
\] (1.33)

Then the matrices $\Lambda^a_b$ of the pure Lorentz transformations (1.27) take the form
\[
\Lambda^a_b = B^a_b(\zeta) \equiv \begin{pmatrix} \cosh \zeta & -\sinh \zeta \hat{\zeta}_1 & -\sinh \zeta \hat{\zeta}_2 & -\sinh \zeta \hat{\zeta}_3 \\ -\sinh \zeta \hat{\zeta}_1 & \delta_{ij} + (\cosh \zeta - 1) \hat{\zeta}_i \hat{\zeta}_j \end{pmatrix}.
\] (1.34)

The notation $B^a_b(\zeta)$ emphasizes that the transformations are boosts. Note that the property (1.28) follows directly from the identities $\hat{\zeta}^2 = 1$, $\cosh^2 \zeta - \sinh^2 \zeta = 1$. 

For active transformations of a physical system, the above transformations have to be inverted. For instance, the active boosts with a rapidity pointing in the $z$-direction, $\hat{\zeta} = (0, 0, 1)$, have the pseudo-orthogonal matrix form

$$\Lambda^a_b = B_3(\zeta) = \begin{pmatrix}
\cosh \zeta & 0 & \sinh \zeta \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\sinh \zeta & 0 & \cosh \zeta
\end{pmatrix}. \quad (1.35)$$

### 1.2.2 Lorentz Group

The set of Lorentz boosts (1.34) can be extended by rotations to form the Lorentz group. In four-by-four matrix notation, the rotation matrices (1.8) have the block form

$$\Lambda^a_b(R) = R^a_b \equiv \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & R_{ij} \\
0 & 0 & 1
\end{pmatrix}. \quad (1.36)$$

It is easy to verify that these satisfy the relation (1.28), which reduces to the orthogonality relation (1.9).

The four-dimensional versions of the active rotations (1.10) around the $z$-axis with a rotation vector $\hat{\varphi} = (0, 0, 1)$ are given by the orthogonal matrices

$$\Lambda^b_a = R_3(\varphi) = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \varphi & -\sin \varphi & 0 \\
0 & \sin \varphi & \cos \varphi & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}. \quad (1.37)$$

The rotation matrix (1.37) differs from the boost matrix (1.35) mainly in the presence of trigonometric functions instead of hyperbolic functions. In addition, there is a sign change under transposition accounting for the opposite sign in the time- and space-like parts of the metric (1.29).

When performing all possible Lorentz boosts and rotations in succession, the resulting set of transformations forms a group called the Lorentz group.

### 1.3 Infinitesimal Lorentz Transformations

The transformation laws of continuous groups such as rotation and Lorentz group are conveniently expressed in an infinitesimal form. By performing many infinitesimal transformations after each other it is always possible to reconstruct from these the finite transformation laws. This is a consequence of the fact that the exponential function $e^x$ can always be obtained by a product of many small-$x$ approximations $e^{x_{\epsilon}} \approx 1 + \epsilon x$:

$$e^x = \lim_{\epsilon \to 0} (1 + \epsilon x)^{1/\epsilon}. \quad (1.38)$$
1.3 Infinitesimal Lorentz Transformations

1.3.1 Generators of Group Transformations

Let us illustrate this procedure for the active rotations (1.37). These can be written in the exponential form

\[
R_3(\varphi) = \exp \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \varphi \right\} \equiv e^{-iL_3 \varphi}. \tag{1.39}
\]

The matrix

\[
L_3 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.40}
\]

is called the generator of this rotation in the Lorentz group. There are similar generators for rotations around x- and y-directions

\[
L_1 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{1.41}
\]

\[
L_2 = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{1.42}
\]

The three generators may compactly be written as

\[
L_i \equiv -i \begin{pmatrix} 0 & 0 \\ 0 & \epsilon_{ijk} \end{pmatrix}, \tag{1.43}
\]

where \(\epsilon_{ijk}\) is the completely antisymmetric Levi-Civita tensor with \(\epsilon_{123} = 1\).

Introducing a vector notation for the three generators, \(\mathbf{L} \equiv (L_1, L_2, L_2)\), the general pure rotation matrix (1.36) is given by the exponential

\[\Lambda(R(\varphi)) = e^{-i\varphi \cdot \mathbf{L}}. \tag{1.44}\]

This follows from the fact that all orthogonal \(3 \times 3\)-matrices in the spatial block of (1.36) can be written as an exponential of \(i\) times all antisymmetric \(3 \times 3\)-matrices, and that these can all be reached by the linear combinations \(\varphi \cdot \mathbf{L}\).

Let us now find the generators of the active boosts, first in the z-direction where we see from (1.35) that the boost matrix can be written as an exponential

\[
B_3(\zeta) = \exp \left\{ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \zeta \right\} = e^{-iM_3 \zeta}. \tag{1.45}
\]
with the generator

\[ M_3 = i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \]  

Similarly we find the generators for the \( x \)- and \( y \)-directions:

\[ M_1 = i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \]

\[ M_2 = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \]

Introducing a vector notation for the three boost generators, \( \mathbf{M} \equiv (M_1, M_2, M_3) \), the general Lorentz transformation matrix (1.34) is given by the exponential

\[ \Lambda(B(\xi)) = e^{-i\xi \mathbf{M}}. \]  

The proof is analogous to the proof of the exponential form (1.44).

The Lorentz group is therefore generated by the six matrices \( L_i, M_i \), to be collectively denoted by \( G_a (a = 1, \ldots, 6) \). Every element of the group can be written as

\[ \Lambda = e^{-i(\varphi L + \xi \mathbf{M})} \equiv e^{-i\alpha_a G_a}. \]  

There exists a Lorentz-covariant way of specifying the generators of the Lorentz group. We introduce the \( 4 \times 4 \)-matrices

\[ (L_{ab})^{cd} = i(g^{ac}g^{bd} - g^{ad}g^{bc}), \]  

labeled by the antisymmetric pair of indices \( ab \), i.e.,

\[ L^{ab} = -L^{ba}. \]  

There are 6 independent matrices which coincide with the generators of rotations and boosts as follows:

\[ L_i = \frac{1}{2} \varepsilon_{ijk} L_{jk}, \quad M_i = L^{0i}. \]  

With the generators (1.51), we can write every element (1.50) of the Lorentz group as follows

\[ \Lambda = e^{-i\frac{1}{2} \omega_{ab} L^{ab}}, \]
where the antisymmetric angular matrix \( \omega_{ab} = -\omega_{ba} \) collects both, rotation angles and rapidities:

\[
\omega_{ij} = \epsilon_{ijk}\phi^k, \quad \omega_{0i} = \zeta^i.
\]

(1.55) \hspace{1cm} (1.56)

Summarizing the notation we have set up an exponential representation of all Lorentz transformations

\[
\Lambda = e^{-i(\psi \cdot L + \zeta \cdot M)} = e^{-i(\frac{1}{2} \omega_{ij} L_{ij} + \omega_{0i} L^{0i})} = e^{-i\frac{1}{2} \omega_{ab} L^{ab}}.
\]

(1.57)

Note that if the metric were Euclidean

\[
g = \begin{pmatrix}
1 & & & \\
 & 1 & & \\
 & & 1 & \\
 & & & 1
\end{pmatrix},
\]

(1.58)

the above representation would be familiar from basic matrix theorems: Then \( \Lambda \) would, by Eq. (1.28), comprise all real orthogonal matrices in four dimensions, and these could be written as an exponential of all real antisymmetric \( 4 \times 4 \)-matrices. For the pseudo-orthogonal matrices satisfying (1.28) with the Minkowski metric (1.29), only \( iL \)'s are antisymmetric while \( iM \) are symmetric.

### 1.3.2 Group Multiplication and Lee Algebra

The reason for expressing the group elements as exponentials of the six generators is that, in this way, the multiplication rules of infinitely many group elements can be completely reduced to the knowledge of the finite number of commutation rules among the six generators \( L_i, M_i \). This is a consequence of the Baker-Campbell-Hausdorff formula [8]

\[
e^{A}e^{B} = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}[A-B,[A,B]]-\frac{1}{12}[A,[B,[A,B]]]+...}.
\]

(1.59)

According to this formula, the product of exponentials can be written as an exponential of commutators. Adapting the general notation \( G_r = (L_i, M_i) \) for the six generators in Eqs. (1.53) and (1.57), the product of two group elements is

\[
\Lambda_1\Lambda_2 = e^{-i\alpha^2 G_r}e^{-i\alpha^2 G_s} = \exp\left\{ -i\alpha^1 G_r - i\alpha^2 G_s + \frac{1}{2} i\alpha^1_{G_r} G_r, -i\alpha^2_{G_s} G_s \right\}
\]

(1.60)

The exponent involves only commutators among \( G_r \)'s. For the Lorentz group these can be calculated from the explicit \( 4 \times 4 \)-matrices (1.40)-(1.42) and (1.46)-(1.48). The result is
\[ [L_i, L_j] = i\epsilon_{ijk}L_k, \]
\[ [L_i, M_j] = i\epsilon_{ijk}M_k, \]
\[ [M_i, M_j] = -i\epsilon_{ijk}L_k. \]

This algebra of generators is called the \textit{Lie algebra} of the group. In the general notation with the generators \( G_r \), the algebra reads
\[ [G_r, G_s] = if_{rst}G_t. \]

The number of linearly independent matrices \( G_r \) (here 6) is called the \textit{rank} \( r \) of the Lie algebra.

In any Lie algebra, the commutator of two generators is a linear combination of generators. The coefficients \( f_{abc} \) are called \textit{structure constants}. They are completely antisymmetric in \( a, b, c \), and satisfy the relation
\[ f_{rsu}f_{utv} + f_{stu}f_{urv} + f_{tru}f_{usv} = 0. \]

This guarantees that the generators obey the \textit{Jacobi identity}
\[ [[G_r, G_s], G_t] + [[G_s, G_t], G_r] + [[G_t, G_r], G_s] = 0, \]
which ensures that multiplication of three exponentials \( \Lambda_j = e^{-i\alpha_j G_r} \) (\( i = 1, 2, 3 \)) obeys the law of associativity \( (\Lambda_1\Lambda_2)\Lambda_3 = \Lambda_1(\Lambda_2\Lambda_3) \) when evaluating the products via the expansion Eq. (1.60).

The relation (1.65) can easily be verified explicitly for the structure constants (1.61)–(1.63) of the Lorentz group using the identity for the \( \epsilon \)-tensor
\[ \epsilon_{ijl}\epsilon_{km} + \epsilon_{jkl}\epsilon_{lim} + \epsilon_{kil}\epsilon_{jlm} = 0. \]

The Jacobi identity implies that the \( r \) matrices with \( r \times r \) elements
\[ (F_r)_{st} = -if_{rst} \]
satisfy the commutation rules (1.64). They are the generators of the so-called \textit{adjoint representation} of the Lie algebra. The matrix in the spatial block of Eq. (1.43) for \( L_i \) is precisely of this type.

In terms of the matrices \( F_r \) of the adjoint representation, the commutation rules can also be written as
\[ [G_r, G_s] = -(F_t)_{ab}G_t. \]

Inserting for \( G_r \) the generators (1.68), we reobtain the relation (1.65).

Continuing the expansion in terms of commutators in the exponent of (1.60), all commutators can be evaluated successively and one remains at the end with an expression
\[ \Lambda_{12} = e^{-i\alpha_1^2(\alpha_1, \alpha_2)}G_r, \]
in which the parameters of the product $\alpha_{12}^r$ are completely determined from those of the factor, $\alpha_1^r, \alpha_2^r$. The result depends only on the structure constants $f_{abc}$, not on the representation.

If we employ the tensor notation $L^{ab}$ for $L_i, M_i$ of Eqs. (1.53), (1.53), and perform multiplication covariantly, so that products $L^{ab} L^{cd}$ have the matrix elements $(L^{ab})_{\sigma \tau} (L^{cd})^\tau_\delta$, the commutators (1.61)–(1.63) can be written as

$$[L^{ab}, L^{cd}] = -i(g^{ac} L^{bd} - g^{ad} L^{bc} + g^{bd} L^{ac} - g^{bc} L^{ad}).$$

(1.71)

Due to the antisymmetry in $a \leftrightarrow b$ and $c \leftrightarrow d$ it is sufficient to specify only the simpler commutators

$$[L^{ab}, L^{ac}] = -i g^{aa} L^{bc}, \text{ no sum over } a.$$  

(1.72)

This list of commutators omits only those commutation rules of (1.71) which vanishes, since none of the indices $ab$ is equal to one of the indices $cd$.

### 1.4 Vectors, Tensors, Scalars

We shall frequently consider four-component physical quantities $v^a$ which, under Lorentz transformation, change in the same way as the coordinates $x^a$:

$$v'^a = \Lambda^a_b v^b.$$  

(1.73)

This transformation property defines a Lorentz vector of four-vector. In addition to such vectors, there are quantities with more indices $v^{ab}, v^{abc}, \ldots$ which transform like products of vectors:

$$v'^{ab} = \Lambda^a_c \Lambda^b_d v^{cd},$$  

(1.74)

$$v'^{abc} = \Lambda^a_d \Lambda^b_e \Lambda^c_f v^{def}.$$  

(1.75)

These are the transformation properties of Lorentz tensors of rank two, three, . . . .

Given any two four-vectors $u^a$ and $v^a$, we define their scalar product in the same way as before in (1.30) for two coordinate vectors $x^a$ and $y^a$:

$$uv = u^a g_{ab} v^b,$$  

(1.76)

Scalar products are, of course, invariant under Lorentz transformations due to their pseudo-orthogonality (1.28).

If $v^a, v^{ab}, v^{abc}, \ldots$ are functions of $x$, they are called vector and tensor fields. Derivatives with respect to $x$ of such field obey vector and tensor transformation laws. Indeed, since

$$x'^a = \Lambda^a_b x^b,$$  

(1.77)

we see that the derivative $\partial/\partial x^b$ satisfies

$$\frac{\partial}{\partial x'^a} = (\Lambda^{T-1})_a^b \frac{\partial}{\partial x^b},$$  

(1.78)
i.e., it transforms with the inverse of the transposed Lorentz matrix $\Lambda^a_b$. Using the pseudo-orthogonality relation (1.28), we can also write

$$\frac{\partial}{\partial x^a} = \left( g \Lambda g^{-1} \right)^b_a \frac{\partial}{\partial x^b}. \quad (1.79)$$

In general, any four-component quantity $v_a$ which transforms like the derivatives

$$v'_a = \left( g \Lambda g^{-1} \right)^b_a v_b \quad (1.80)$$

is called a covariant Lorentz vector or four-vector, as opposed to the vector $v^a$ transforming like $x^a$ itself, which is called contravariant vector.

A covariant vector $v_a$ can be produced from a contravariant one $v^b$ by multiplication with the metric tensor:

$$v_a = g_{ab} v^b. \quad (1.81)$$

This operation is called lowering the index. The operation can be inverted to what is called raising the index:

$$v^a = g^{ab} v_b, \quad (1.82)$$

where $g^{ab}$ are the matrix elements of the inverse metric

$$g^{ab} \equiv \left( g^{-1} \right)^{ab}. \quad (1.83)$$

With Einstein’s summation convention, the inverse metric $g^{ab} \equiv \left( g^{-1} \right)^{ab}$ satisfies the equation

$$g^{ab} g_{bc} = \delta^a_c \quad (1.84)$$

The sum over a common upper and lower index is called contraction.

In Minkowski space, the matrices $g$ and $g^{-1}$ happen to be the same and so are the matrix elements $g_{ab}$ and $g^{ab}$, both being equal to (1.29). This will no longer be true in the general geometries of gravitational physics. For this reason it will be useful to keep separate symbols for the metric $g$ and its inverse $g^{-1}$, and for their matrix elements $g_{ab}$ and $g^{ab}$.

The contraction of a covariant vector with a contravariant vector is a scalar product, as is obvious if we rewrite the scalar product (1.76) as

$$uv = u^a g_{ab} v^b = u^a v_a = u^a v^a. \quad (1.85)$$

Of course, we can form also the scalar product of two covariant vectors with the help of the inverse metric $g^{-1}$:

$$uv = u_a g^{ab} v_b. \quad (1.86)$$

The invariance under Lorentz transformations (1.80) is easily verified using the pseudo-orthogonality property (1.28):

$$u'_a g^{ab} v'_b = u'^T g^{-1} v' = u'^T g^{-1} \Lambda^T g g^{-1} g \Lambda g^{-1} v = u'^T g^{-1} v = u_a g^{ab} v_b. \quad (1.87)$$
Since $\partial/\partial x^a$ transforms like a covariant vector, it is useful to emphasize this behavior by the notation
\[ \partial_a \equiv \frac{\partial}{\partial x^a}. \] (1.88)

Extending the definition of covariant vectors one defines covariant tensors of rank two $t_{ab}$, three $t_{abc}$, etc. as quantities transforming like
\[ t'_{ab} = (g\Lambda g^{-1})_a^c (g\Lambda g^{-1})_b^d t_{cd}, \]
\[ t'_{abc} = (g\Lambda g^{-1})_a^c (g\Lambda g^{-1})_b^f (g\Lambda g^{-1})_c^g t_{efg}. \] (1.89)

Co- and contravariant vectors, tensors, can always be multiplied with each other to form new co- and contravariant quantities if the indices to be contracted are raised and lowered appropriately. If no uncontracted indices are left, one obtains an invariant, a Lorentz scalar.

It is useful to introduce a contravariant version of the covariant derivative vector
\[ \partial^a \equiv g^{ab}\partial_b, \] (1.90)
and covariant versions of the contravariant coordinate vector
\[ x_a \equiv g_{ab}x^b. \] (1.91)

The invariance of Maxwell’s equations (1.20) is a direct consequence of these contraction rules since the differential operator on the left-hand side can be written covariantly as
\[ \frac{1}{c^2} \partial^2 - \nabla^2 = \frac{\partial}{\partial x^a}g^{ab}\frac{\partial}{\partial x^b} = \partial_a g^{ab}\partial_b = \partial^a\partial_a = \partial^2. \] (1.92)

The right-hand side is obviously a Lorentz scalar.

### 1.4.1 Discrete Lorentz Transformations

The Lorentz group can be extended to include space reflections in any of the four spacetime directions
\[ x^a \rightarrow -x^a, \] (1.93)
without destroying the defining property (1.28). The determinant of $\Lambda$, however, is then negative. If only $x^0$ is reversed, the reflection is also called time reversal and denoted by
\[ T = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}. \] (1.94)
The simultaneous reflection of the three spatial coordinates is called \textit{parity transformation} and denoted by the \(4 \times 4\)-matrix \(P\), i.e.,

\[
P = \begin{pmatrix}
1 & -1 \\
-1 & 1 \\
-1 & 1 \\
-1 & 1
\end{pmatrix}.
\] (1.95)

After this extension, the entire Lorentz group can no longer be obtained from the neighborhood of the identity by a product of infinitesimal transformations, i.e., by an exponential of the Lie algebra in Eq. (1.57). It consists of four topologically disjoint pieces which can be obtained by a product of infinitesimal transformations multiplied with \(1\), \(P\), \(T\), and \(PT\). The four pieces of the group are

\[
e^{-i\frac{1}{2}\omega_{ab}L^{ab}}, \quad e^{-i\frac{1}{2}\omega_{ab}L^{ab}}P, \quad e^{-i\frac{1}{2}\omega_{ab}L^{ab}}T, \quad e^{-i\frac{1}{2}\omega_{ab}L^{ab}}PT.
\] (1.96)

The Lorentz transformations \(\Lambda\) of the pieces associated with \(P\) and \(T\) have a negative determinant. This leads to the definition of \textit{pseudotensors} which transform like a tensor, but with an additional factor \(\det\Lambda\). A vector with this property is also called \textit{axial vector}. In three dimensions, the angular momentum \(L = x \times p\) is an axial vector since it does not change sign under space reflections, as the vector \(x\), but remains invariant.

### 1.4.2 Poincaré group

Just as the Galilei transformations, the Lorentz transformations can be extended by the group of spacetime translations

\[
x^a = x^a - a^a
\] (1.97)

to form the \textit{inhomogeneous Lorentz group} or \textit{Poincaré group}.

Inertial frames may be defined as all those frames in which Maxwell’s equations are valid. They differ from each other by Poincaré transformations.

\[
x'^a = \Lambda^a_b x^b - a^a.
\] (1.98)

### 1.5 Differential Operators for Lorentz Transformations

The physical laws in four-dimensional spacetime are formulated in terms of Lorentz-invariant field theories. The fields depend on the spacetime coordinates \(x^a\). In order to perform transformations of the Lorentz group we need differential operators for the generators of this group.

For Lorentz transformations \(\Lambda\) with small rotation angles and rapidities, we can approximate the exponential in (1.57) as

\[
\Lambda \equiv 1 - i \frac{1}{2} \omega_{ab}L^{ab}.
\] (1.99)
Then the Lorentz transformation of the coordinates

\[ x \xrightarrow{\Lambda} x' = \Lambda x \]  

is conveniently characterized by the infinitesimal change

\[ \delta_{\Lambda} x = x' - x = -i \frac{1}{2} \omega_{ab} L^{ab} x. \]  

Inserting the 4 × 4 matrix generators (1.51), this becomes more explicitly

\[ \delta_{\Lambda} x^a = \omega^a_{\ b} x^b. \]  

We now observe that this can be expressed in terms of the differential operators

\[ \hat{L}^{ab} \equiv i (x^a \partial^b - x^b \partial^a) = -\hat{L}^{ba}, \]  

as a commutator

\[ \delta_{\Lambda} x = i \frac{1}{2} \omega_{ab} [\hat{L}^{ab}, x]. \]  

The differential operators (1.103) satisfy the same commutation relations (1.71), (1.72) as the 4 × 4 -generators \( L^{ab} \) of the Lorentz group. They form a representation of the Lie algebra (1.71), (1.72). By exponentiation we can thus form the operator representation of finite Lorentz transformations

\[ \hat{D}(\Lambda) \equiv e^{-i \frac{1}{2} \omega_{ab} \hat{L}^{ab}}, \]  

which satisfy the same group multiplication rules as the 4 × 4-matrices \( \Lambda \).

The relation between the finite Lorentz transformations (1.100) and the operator version (1.105) is

\[ x' = \Lambda x = e^{-i \frac{1}{2} \omega_{ab} \hat{L}^{ab} x} x = e^{i \frac{1}{2} \omega_{ab} L^{ab} x} e^{-i \frac{1}{2} \omega_{ab} L^{ab} x} = \hat{D}^{-1}(\Lambda) x \hat{D}(\Lambda). \]  

This is proved by expanding, on the left-hand side, \( e^{-i \frac{1}{2} \omega_{ab} \hat{L}^{ab} x} x \) in powers of \( \omega_{ab} \), and doing the same thing on the right-hand with \( e^{i \frac{1}{2} \omega_{ab} L^{ab} x} e^{-i \frac{1}{2} \omega_{ab} L^{ab} x} \) with the help of Lie’s expansion formula

\[ e^{-iA} \hat{B} e^{iA} = 1 - i [\hat{A}, \hat{B}] + \frac{1}{2!} [\hat{A}, [\hat{A}, \hat{B}]] + \ldots . \]  

This operator representation (1.105) can be used to generate Lorentz transformations on the spacetime argument of any function of \( x \):

\[ f'(x) \equiv f(\Lambda^{-1}x) = f \left( \hat{D}(\Lambda) x \hat{D}^{-1}(\Lambda) \right) = \hat{D}(\Lambda) f (x) \hat{D}^{-1}(\Lambda). \]  

The latter step follows from a power series expansion of \( f(x) \). Take, for example an expansion term \( f_{a,b} x^a x^b \) of \( f(x) \). In the transformed function \( f'(x) \), this becomes

\[ f_{a,b} \hat{D}(\Lambda) x^a \hat{D}^{-1}(\Lambda) \hat{D}(\Lambda) x^b \hat{D}^{-1}(\Lambda) = \hat{D}(\Lambda) \left( f_{a,b} x^a x^b \right) \hat{D}^{-1}(\Lambda). \]
1.6 Vector and Tensor Operators

In working out the commutation rules among the differential operators \( \hat{L}^{ab} \) one conveniently uses the commutation rules between \( \hat{L}^{ab} \) and \( x^c, p^c \):

\[
[\hat{L}^{ab}, x^c] = -i(g^{ac}x^b - g^{bc}x^a) = -(L^{ab})^c_d x^d, \quad (1.110)
\]
\[
[\hat{L}^{ab}, p^c] = -i(g^{ac}p^b - g^{bc}p^a) = -(L^{ab})^c_d p^d. \quad (1.111)
\]

These commutation rules identify \( x^c \) and \( p^c \) as vector operators.

In general, an operator \( \hat{t}^{c_1, \ldots, c_n} \) is said to be a tensor operator of rank \( n \) if each of its tensor indices is transformed under commutation with \( L^{ab} \) like the index of \( x^a \) or \( p^a \) in (1.110) and (1.111):

\[
[\hat{L}^{ab}, \hat{t}^{c_1, \ldots, c_n}] = -i\left(g^{ac_1}t^{b, \ldots, c_n} - g^{bc_1}t^{a, \ldots, c_n}\right) + \ldots + (g^{ac_n}t^{c_1, \ldots, b} - g^{bc_n}t^{c_1, \ldots, a})
\]
\[
= -(L^{ab})^{c_1}_d t^{d, c_2, \ldots, c_n} - (L^{ab})^{c_2}_d t^{d, c_1, \ldots, c_n} - \ldots - (L^{ab})^{c_n}_d t^{d, c_1, \ldots, c_n}. \quad (1.112)
\]

The commutators (1.71) between the generators imply that these are themselves tensor operators.

The simplest examples for such tensor operators are the direct products of vectors such as \( \hat{t}^{c_1, \ldots, c_n} = x^{c_1} \ldots x^{c_n} \) or \( \hat{t}^{c_1, \ldots, c_n} = p^{c_1} \ldots p^{c_n} \). In fact, the right-hand side can be found for such direct products using the commutation rules between products of operators

\[
[a, \hat{b}c] = [\hat{a}, \hat{b}][c] + [\hat{a}, \hat{c}][\hat{b}], \quad [\hat{a}b, \hat{c}] = \hat{a}[\hat{b}, \hat{c}] + [\hat{a}, \hat{c}][\hat{b}]. \quad (1.113)
\]

These are the analogs of the Leibnitz chain rule for derivatives

\[
\partial(fg) = (\partial f)g + f(\partial g). \quad (1.114)
\]

1.7 Finite Operator Lorentz Transformations

Let us apply such a finite operator representation (1.105) to the vector \( x^c \) to form

\[
\hat{D}(\Lambda)x^c \hat{D}^{-1}(\Lambda). \quad (1.115)
\]

We do this separately for rotations and Lorentz transformations, first for rotations.

1.7.1 Rotations

An arbitrary three-vector \((x^1, x^2, x^3)\) is rotated around the 3-axis by the operator \( \hat{D}(R_3(\varphi)) = e^{-i\varphi \hat{L}_3} \) with \( \hat{L}_3 = -i(x^1 \partial_2 - x^2 \partial_1) \) by the operation

\[
\hat{D}(R_3(\varphi))x^i \hat{D}^{-1}(R_3(\varphi)) = e^{-i\varphi \hat{L}_3}x^i e^{i\varphi \hat{L}_3}. \quad (1.116)
\]

Since \( \hat{L}_3 \) commutes with \( x^3 \), this component is invariant under the operation (1.116):

\[
\hat{D}(R_3(\varphi))x^3 \hat{D}^{-1}(R_3(\varphi)) = e^{-i\varphi \hat{L}_3}x^3 e^{i\varphi \hat{L}_3} = x^3. \quad (1.117)
\]
For \( x^1 \) and \( x^2 \), the Lie expansion of (1.115) contains the commutators
\[
-i[L_3, x^1] = x^2, \quad -i[L_3, x^2] = -x^1.
\] (1.118)
Thus, the first-order expansion term on the right-hand side of (1.116) transforms the two-dimensional vector \((x^1, x^2)\) into \((x^2, -x^1)\). The second-order term is obtained by commuting the operator \(-i\hat{L}_3\) with \((x^2, -x^1)\), yielding \(-(x^1, x^2)\). To third-order, this is again transformed into \(-(x^2, -x^1)\), and so on. Obviously, all even orders reproduce the initial two-dimensional vector \((x^1, x^2)\) with an alternating sign, while all odd powers are proportional to \((x^2, -x^1)\). Thus we obtain the expansion
\[
e^{-i\varphi \hat{L}_3}(x^1, x^2)e^{i\varphi \hat{L}_3} = \left(1 - \frac{1}{2!}\varphi^2 + \frac{1}{4!}\varphi^4 + \ldots\right)(x^1, x^2) + \left(\varphi - \frac{1}{3!}\varphi^3 + \frac{1}{5!}\varphi^5 + \ldots\right)(x^2, -x^1).
\] (1.119)
The even and odd powers can be summed up to a cosine and a sine, respectively, resulting in
\[
e^{-i\varphi \hat{L}_3}(x^1, x^2)e^{i\varphi \hat{L}_3} = \cos \varphi (x^1, x^2) + \sin \varphi (x^2, -x^1).
\] (1.120)
Together with the invariant \( x^3 \) in (1.117), the right-hand side forms a vector arising from \( x^1 \) by an inverse rotation (1.37). Thus
\[
\hat{D}(R_3(\varphi))x^i\hat{D}^{-1}(R_3(\varphi)) = e^{-i\varphi \hat{L}_3}x^i e^{i\varphi \hat{L}_3} = (e^{i\varphi \hat{L}_3})^i_j x^j = R_3^{-1}(\varphi)^i_j x^j.
\] (1.121)
By performing successively rotations around the three axes we can generate in this way any inverse rotation:
\[
\hat{D}(R(\varphi))x^i\hat{D}^{-1}(R(\varphi)) = e^{-i\varphi \hat{L}_3}x^i e^{i\varphi \hat{L}_3} = (e^{i\varphi \hat{L}_3})^i_j x^j = R^{-1}(\varphi)^i_j x^j.
\] (1.122)
This is the finite transformation law associated with the commutation relation
\[
[L_i, x_k] = x_j (L_i)_{jk},
\] (1.123)
which characterizes the vector operator nature of \( x^i \). Thus (1.122) holds for any vector operator \( \hat{v}^i \).

The time component \( x^0 \) is obviously unchanged by a rotation since \( \hat{L}_3 \) commutes with \( x^0 \). Hence we can extend (1.122) trivially to a four-vector, replacing \( \hat{D}(R(\varphi)) \) by \( \hat{D}(\Lambda(R(\varphi))) \) [recall (1.44)].

### 1.7.2 Lorentz Boosts

A similar calculation may be done for Lorentz boosts. Here we first consider a boost in the 3-direction \( B_3(\zeta) = e^{-i\zeta \hat{M}_3} \) generated by \( \hat{M}_3 = \hat{L}^{03} = -i(x^0 \partial_3 + x^3 \partial_0) \) [recall (1.57), (1.53), and (1.103)]. Note the positive relative sign of the two terms in the
generator \( \hat{L}^{03} \) caused by the fact that \( \partial_i = -\partial^i \) in contrast to \( \partial_0 = \partial^0 \). Thus we form

\[
\hat{D}(B_3(\xi)) x^i \hat{D}^{-1}(B_3(\xi)) = e^{-i\xi \hat{M}_3} x^i e^{i\xi \hat{M}_3}.
\] (1.124)

The Lie expansion of the right-hand side involves the commutators

\[-i[M_3, x^0] = -x^3, \quad -i[M_3, x^1] = -x^0, \quad -i[M_3, x^2] = 0, \quad -i[M_3, x^3] = 0. \] (1.125)

Here the two-vector \((x^1, x^2)\) is unchanged, while the two-vector \((x^0, x^3)\) is transformed into \(-(x^3, x^0)\). In the second expansion term, the latter becomes \((x^0, x^3)\), and so on, yielding

\[
e^{-i\xi \hat{M}_3}(x^0, x^3)e^{i\xi \hat{M}_3} = \left(1 + \frac{1}{2!}\xi^2 + \frac{1}{4!}\xi^4 + \ldots\right)(x^0, x^3)
- \left(\xi + \frac{1}{3!}\xi^3 + \frac{1}{5!}\xi^5 + \ldots\right)(x^3, x^0).
\] (1.126)

The right-hand sides can be summed up to hyperbolic cosines and sines:

\[
e^{-i\xi \hat{M}_3}(x^0, x^3)e^{i\xi \hat{M}_3} = \cosh \xi (x^0, x^3) - \sinh \xi (x^3, x^0).
\] (1.127)

Together with the invariance of \((x^1, x^2)\), this corresponds precisely to the inverse of the boost transformation (1.35):

\[
\hat{D}(B_3(\xi)) x^a \hat{D}^{-1}(B_3(\xi)) = e^{-i\xi \hat{M}_3} x^a e^{i\xi \hat{M}_3} = (e^{i\xi \hat{M}_3})^a_b x^b = B_3^{-1}(\xi)^a_b x^b.
\] (1.128)

### 1.7.3 Lorentz Group

By performing successively rotations and boosts in all directions we find all Lorentz transformations

\[
\hat{D}(\Lambda) x^c \hat{D}^{-1}(\Lambda) = e^{-i\frac{1}{2}\omega_{ab}\hat{L}^{ab}} x^c e^{i\frac{1}{2}\omega_{ab}\hat{L}^{ab}} = (e^{i\frac{1}{2}\omega_{ab}\hat{L}^{ab}})^c_d x^d = (\Lambda^{-1})^c_d x^d,
\] (1.129)

where \(\omega_{ab}\) are the parameters (1.55) and (1.56). In the last term on the right-hand side we have expressed the \(4 \times 4\)-matrix \(\Lambda\) as an exponential of its generators as well, to emphasize the one-to-one correspondence between the generators \(\hat{L}^{ab}\) and their differential-operator representation \(\hat{L}^{ab}\).

At first it may seem surprising that the group transformations appearing as a left-hand factor of the two sides of these equations are inverse to each other. However, we may easily convince ourselves this is necessary to guarantee the correct group multiplication law. Indeed, if we perform two transformations after each other they appear in opposite order on the right- and left-hand sides:

\[
\hat{D}(\Lambda_2\Lambda_1) x^c \hat{D}^{-1}(\Lambda_2\Lambda_1) = \hat{D}(\Lambda_2) \hat{D}(\Lambda_1) x^c \hat{D}^{-1}(\Lambda_1) \hat{D}^{-1}(\Lambda_2) = (\Lambda_2^{-1})^c_d \hat{D}(\Lambda_2) x^d \hat{D}^{-1}(\Lambda_2) = (\Lambda_2^{-1})^c_d (\Lambda_2^{-1})^d_{c'} x^{c'} = [(\Lambda_2^{-1})^c_d x^d]_{c'}.
\] (1.130)

If the right-hand side of (1.129) would contain \(\Lambda\) instead of \(\Lambda^{-1}\), the order of the factors in \(\Lambda_2\Lambda_1\) on the right-hand side of (1.130) would be opposite to the order in \(\hat{D}(\Lambda_2\Lambda_1)\) on the left-hand side.
A straightforward extension of the operation (1.129) yields the transformation law for a tensor $\hat{t}^{c_1,\ldots,c_n} = x^{c_1} \cdots x^{c_n}$:

$$\hat{D}(\Lambda) \hat{t}^{c_1,\ldots,c_n} \hat{D}^{-1}(\Lambda) = e^{-i\frac{1}{2}\omega_{ab} L^{ab}} \hat{t}^{c_1,\ldots,c_n} e^{i\frac{1}{2}\omega_{ab} L^{ab}}$$

$$= (\Lambda^{-1})_{c_1}^{c'_1} \cdots (\Lambda^{-1})_{c_n}^{c'_n} \hat{t}^{c'_1,\ldots,c'_n}$$

$$= (e^{i\frac{1}{2}\omega_{ab} L^{ab}})_{c_1}^{c'_1} \cdots (e^{i\frac{1}{2}\omega_{ab} L^{ab}})_{c_n}^{c'_n} \hat{t}^{c'_1,\ldots,c'_n}.$$  \hspace{1cm} (1.131)

This follows directly by inserting in the product $x^{c_1} \cdots x^{c_n}$, an auxiliary unit factor $1 = \hat{D}(\Lambda) \hat{D}^{-1}(\Lambda) = e^{-i\frac{1}{2}\omega_{ab} L^{ab}} e^{i\frac{1}{2}\omega_{ab} L^{ab}}$ between neighboring factors $x^{c_i}$ and performing the operation (1.131) on each of them. The last term in (1.131) can also be written as

$$[e^{i\frac{1}{2}\omega_{ab} (L^{ab} \times 1 \times 1 \cdots 1 + \cdots + 1 \times L^{ab} \times 1 \cdots 1)}]^{c_1,\ldots,c_n} \hat{t}^{c'_1,\ldots,c'_n}. \hspace{1cm} (1.132)$$

Since the commutation relations (1.112) determine the result completely, the transformation formula (1.131) is true for any tensor operator $\hat{t}^{c_1,\ldots,c_n}$ not only those composed from a product of vectors $x^{c_i}$.

The result can easily be extended to an exponential function $e^{-ipx}$ and further to any function $f(x)$ which possesses a Fourier representation

$$\hat{D}(\Lambda) f(x) \hat{D}^{-1}(\Lambda) = f(\Lambda^{-1} x) = e^{-i\frac{1}{2}\omega_{ab} L^{ab}} f(x) e^{i\frac{1}{2}\omega_{ab} L^{ab}}.$$  \hspace{1cm} (1.133)

Since the last differential operator has nothing to act on, it can also be omitted and we can also write

$$\hat{D}(\Lambda) f(x) \hat{D}^{-1}(\Lambda) = f(\Lambda^{-1} x) = e^{-i\frac{1}{2}\omega_{ab} L^{ab}} f(x). \hspace{1cm} (1.134)$$

1.8 Relativistic Point Mechanics

The Lorentz invariance of the Maxwell equations explains the observed invariance of the light velocity in different inertial frames. It is, however, incompatible with Newton’s mechanics. There exists a modification of Newton’s laws which makes them Lorentz-invariant as well, while differing very little from Newton’s equations in their description of slow macroscopic bodies, for which Newton’s equations were originally designed. Let us introduce the Poincaré-invariant distance measure in spacetime

$$ds \equiv \sqrt{dx^2} = \left( g_{ab} dx^a dx^b \right)^{1/2}. \hspace{1cm} (1.135)$$

At a fixed coordinate point of an inertial frame, $ds$ is equal to $c$ times the elapsed time:

$$ds = \sqrt{g_{00} dx^0 dx^0} = dx^0 = c dt. \hspace{1cm} (1.136)$$

Einstein called the quantity

$$d\tau \equiv \frac{1}{c} ds \hspace{1cm} (1.137)$$
the proper time.

When going from one inertial frame to another, two simultaneous events at different points in the first frame will take place at different times in the other frame. Their invariant distance, however, remains the same, due to pseudo-orthogonality relation (1.28) which ensures that

$$ds' = \left( g_{ab} dx'^a dx'^b \right)^{1/2} = \left( g_{ab} dx^a dx^b \right)^{1/2} = ds.$$ \hspace{1cm} (1.138)

A particle moving with a constant velocity along a trajectory \(x(t)\) in one Minkowski frame remains at rest in another frame moving with velocity \(v = \dot{x}(t)\) with respect to the first. Its proper time is then related to the coordinate time in the first frame by the Lorentz transformation

$$cd\tau = ds = \sqrt{c^2 dt^2 - dx^2} = cdt \sqrt{1 - \frac{1}{c^2} \left( \frac{dx}{dt} \right)^2} = cdt \sqrt{1 - \frac{v^2}{c^2}} = \frac{cdt}{\gamma}.$$ \hspace{1cm} (1.139)

This is the famous Einstein relation implying that a moving particle lives longer by a factor \(\gamma\). There exists direct experimental evidence for this phenomenon. For example, the meson \(\pi^+\) lives on the average \(\tau_a \approx 2.60 \times 10^{-8}\) sec, after which it decays into a muon and a neutrino. If the pion is observed in a bubble chamber with a velocity equal to 10% of the light velocity \(c \equiv 299 792 458\) m/sec, it leaves trace of an average length \(l \approx \tau_a \times c \times 0.1 / \sqrt{1 - 0.1^2} \approx 0.78\) cm. A very fast muon moving with 90% of the light velocity, on the other hand, leaves a trace which is longer by a factor \((0.9/0.1) \times \sqrt{1 - 0.10^2} / \sqrt{1 - 0.9^2} \approx 20.6\). Massless particles move with light velocity and have \(d\tau = 0\), i.e., the proper time stands still along their paths. This implies that massless particles can never decay—they are necessarily stable particles.

Another place to see this time dilation effect is by observing the spectral lines of a moving atom, say a hydrogen atom. If the atom is at rest, the frequency of the line is given by

$$\nu = -Ry \left( \frac{1}{n^2} - \frac{1}{m^2} \right)$$ \hspace{1cm} (1.140)

where \(Ry\) is the Rydberg constant (\(\approx 13.6\) eV), and \(n\) and \(m\) are the principal quantum numbers of initial and final electron orbits. If the atom emits a light quantum while moving through the laboratory frame of reference with velocity \(v\) orthogonal to the direction of observation, this frequency is lowered by a factor \(1/\gamma\):

$$\nu_{\text{obs}} = \frac{1}{\gamma} = \sqrt{1 - \frac{v^2}{c^2}}.$$ \hspace{1cm} (1.141)

If the atom runs away from the observer or towards him, the frequency is further changed by the Doppler shift. Due to the growing or decreasing distance, the wave trains arrive with a smaller of higher frequency given by

$$\frac{\nu_{\text{obs}}}{\nu} = \left( 1 \pm \frac{v}{c} \right)^{-1} \frac{1}{\gamma} = \frac{1 \mp v/c}{1 \pm v/c}.$$ \hspace{1cm} (1.142)
In the first case the observer sees an additional red shift, in the second a violet shift of the spectral lines.

Without external forces, the trajectories of free particles are straight lines in four-dimensional spacetime. If the particle positions are parametrized by the proper time $\tau$, they satisfy the equation of motion

$$\frac{d^2}{d\tau^2}q^a(\tau) = \frac{d}{d\tau}p^a(\tau) = 0.$$  \hspace{1cm} (1.143)

The first derivative of $q^a(\tau)$ is the relativistic four-vector of momentum $p^a(\tau)$, briefly called four-momentum:

$$p^a(\tau) \equiv m \frac{d}{d\tau}q^a(\tau) \equiv mu^a(\tau).$$  \hspace{1cm} (1.144)

On the right-hand side we have introduced the relativistic four-vector of velocity $u^a(\tau)$, or four-velocity. Inserting (1.139) into (1.144) we identify the components of $u^a(\tau)$ as

$$u^a = \begin{pmatrix} \gamma c \\ \gamma v^a \end{pmatrix},$$  \hspace{1cm} (1.145)

and see that $u^a(\tau)$ is normalized to the light-velocity:

$$u^a(\tau)u_a(\tau) = c^2.$$  \hspace{1cm} (1.146)

The time and space components of (1.144) are

$$p^0 = m\gamma c = mu^0, \quad p^i = m\gamma v^i = mu^i.$$  \hspace{1cm} (1.147)

This shows that the time dilation factor $\gamma$ is equal to $p^0/mc$, and the same factor increases the spatial momentum with respect to the nonrelativistic momentum $mv^i$. This correction becomes important for particles moving near the velocity of light, which are called relativistic. The light particle has $m = 0$ and $v = c$. It is ultra-relativistic.

Note that by Eq. (1.147), the hyperbolic functions of the rapidity in Eq. (1.31) are related to the four velocity and to energy and momentum by

$$\cosh \zeta = u^0/c = p^0/mc, \quad \sinh \zeta = |u|/c = |p|/mc.$$  \hspace{1cm} (1.148)

Under a Lorentz transformation of space and time, the four-momenta $p^a$ transform in exactly the same way as the coordinate four-vectors $x^a$. This is, of course, due to the Lorentz invariance of the proper time $\tau$ in Eq. (1.144). Indeed, from Eq. (1.147) we derive the important relation

$$p^{02} - p^2 = m^2c^2,$$  \hspace{1cm} (1.149)
which shows that the square of the four-momentum taken with the Minkowski metric is an invariant:

\[ p^2 \equiv p^a g_{ab} p^b = m^2 c^2. \]  

(1.150)

Since both \( x^a \) and \( p^a \) are Lorentz vectors, the scalar product of them,

\[ xp \equiv g_{ab} x^a p^b, \]  

(1.151)

is an invariant. In the canonical formalism, the momentum \( p^i \) is the conjugate variable to the space coordinate \( x^i \). Equation (1.151) suggests that the quantity \( cp^0 \) is conjugate to \( x^0/c = t \). As such it must be the energy of the particle:

\[ E = cp^0. \]  

(1.152)

From relation (1.149), we calculate the energy as a function of the momentum of a relativistic particle:

\[ E = c\sqrt{p^2 + m^2 c^2}. \]  

(1.153)

For small velocities, this can be expanded as

\[ E = mc^2 + \frac{m}{2}v^2 + \ldots. \]  

(1.154)

The first term gives a non-vanishing rest energy which is unobservable in non-relativistic physics. The second term is Newton’s kinetic energy.

The first term has dramatic observable effects. Particles can be produced and disappear in collision processes. In the latter case, their rest energy \( mc^2 \) can be transformed into kinetic energy of other particles. The large factor \( c \) makes unstable particles a source of immense energy, which led to the atomic disaster of Hiroshima and Nagasaki in 1945.

### 1.9 Quantum Mechanics

In quantum mechanics, free spinless particles of momentum \( p \) are described by plane waves of the form

\[ \psi_p(x) = \mathcal{N} e^{-ipx/\hbar}, \]  

(1.155)

where \( \mathcal{N} \) is some normalization factor. The momentum components are the eigenvalue of the differential operators

\[ \hat{p}_a = i\hbar \frac{\partial}{\partial x^a}, \]  

(1.156)

which satisfy with \( x^b \) the commutation rules

\[ [\hat{p}_a, x^b] = i\hbar \delta_a^b. \]  

(1.157)

In terms of these, the generators (1.103) can be rewritten as

\[ \hat{L}^{\alpha\beta} \equiv \frac{1}{\hbar} (x^a \hat{p}^\beta - x^\beta \hat{p}^a). \]  

(1.158)
Apart from the factor \(1/\hbar\), this is the tensor version of the four-dimensional angular momentum.

It is worth observing that the differential operators (1.158) can also be expressed as a sandwich of the \(4 \times 4\) matrix generators (1.51) between \(x^c\) and \(\hat{p}^d\):

\[
\hat{L}^{ab} = -\frac{i}{\hbar} (L^{ab})_{cd} x^c \hat{p}^d = -\frac{i}{\hbar} x^T L^{ab} \hat{p} = i\hat{p}^T L^{ab} x. \quad (1.159)
\]

This way of forming operator representations of the \(4 \times 4\) Lie algebra (1.71) is a special application of a general construction technique of higher representations of a defining matrix representations. In fact, the procedure of second quantization is based on this construction, which extends the single-particle Schrödinger operators to the Fock space of many-particle states.

In general, one may always introduce vectors of creation and annihilation operators \(\hat{a}_{c}^\dagger\) and \(\hat{a}^d\) with the commutation rules

\[
[\hat{a}^c, \hat{a}^d] = [\hat{a}_{c}^\dagger, \hat{a}_{d}^\dagger] = 0; \quad [\hat{a}^c, \hat{a}_{d}^\dagger] = \delta^c_d, \quad (1.160)
\]

and form sandwich operators

\[
\hat{L}^{ab} = \hat{a}_{c}^\dagger (L^{ab})^c_d \hat{a}^d. \quad (1.161)
\]

These satisfy the same commutation rules as the sandwiched matrices due to the Leibnitz chain rule (1.113). Since \(-i\hat{p}_{a}/\hbar\) and \(x^a\) commute in the same way as \(\hat{a}\) and \(\hat{a}_{c}^\dagger\), the commutation rules of the matrices go directly over to the sandwich operators (1.159). The higher representations generated by them lie in the Hilbert space of square-integrable functions.

Under a Lorentz transformation, the momentum of the particle described by the wave function (1.155) goes over into \(p' = \Lambda p\), so that the wave function transforms as follows:

\[
\phi_p(x) \overset{\Lambda}{\rightarrow} \phi'_p(x) \equiv \phi_p'(x) = \mathcal{N}e^{-i(\Lambda p)x} = \mathcal{N}e^{-ip\Lambda^{-1}x} = \phi_p(\Lambda^{-1}x). \quad (1.162)
\]

This can also be written as \(\phi'_p(x') = \phi_p(x)\). An arbitrary superposition of such waves transforms like

\[
\phi(x) \overset{\Lambda}{\rightarrow} \phi'(x) = \phi(\Lambda^{-1}x), \quad (1.163)
\]

which is the defining relation for a scalar field.

The transformation (1.163) can also be generated by the differential-operator representation of the Lorentz group (1.134) as follows:

\[
\phi(x) \overset{\Lambda}{\rightarrow} \phi'(x) = \hat{D}(\Lambda)\phi(x). \quad (1.164)
\]
1.10 Relativistic Particles with Electromagnetic Interactions

Lorentz and Einstein formulated a theory of relativistic massive particles with electromagnetic interactions called the *Maxwell-Lorentz theory*. It is invariant under the Poincaré group and describes the dynamical properties of charged particles such as electrons moving with nonrelativistic and relativistic speeds.

The motion for a particle of charge $e$ and mass $m$ in an electromagnetic field is governed by the *Lorentz equations*

$$\frac{dp^a(\tau)}{d\tau} = m \frac{d^2x^a(\tau)}{d\tau^2} = f^a(\tau),$$

(1.165)

where the four-vector $f^a$ is the *Lorentz force*

$$f^a = e \frac{e}{c} F^a_{\ b} \frac{dx^b}{d\tau} = \frac{e}{cm} F^a_{\ b}(x(\tau)) p^b(\tau),$$

(1.166)

and $F^a_{\ b}(x)$ is a $4 \times 4$ combination of electric and magnetic fields with the components

$$F^i_{\ j} = e^{ijk} B^k; \quad F^0_{\ i} = E^i.$$

(1.167)

By raising the second index of $F^a_{\ b}$ one obtains the tensor

$$F^{ac} = g^{cb} F^a_{\ b}$$

(1.168)

associated with the antisymmetric matrix of the six electromagnetic fields

$$F^{ab} = \begin{pmatrix}
    0 & -E^1 & -E^2 & -E^3 \\
    E^1 & 0 & B^3 & -B^2 \\
    E^2 & -B^3 & 0 & B^1 \\
    E^3 & B^2 & -B^1 & 0
\end{pmatrix}. \quad (1.169)$$

This tensor notation is useful since $F^{ab}$ transforms under the Lorentz group in the same way as the direct product $x^a x^b$, which goes over into $x^a x^b = \Lambda^a_c \Lambda^b_d x^c x^d$. In $F^{ab}(x)$, also the arguments must be transformed as in the scalar field in Eq. (1.163), so that we find the generic transformation behavior of a *tensor field*:

$$F^{ab}(x) \xrightarrow{\Lambda} F'^{ab}(x) = \Lambda^a_c \Lambda^b_d F^{cd}(\Lambda^{-1} x).$$

(1.170)

Recalling the exponential representation (1.132) of the direct product of the Lorentz transformations and the differential operator generation (1.134) of the transformation of the argument $x$, this can also be written as

$$F^{ab}(x) \xrightarrow{\Lambda} F'^{ab}(x) = [e^{-i\frac{1}{2} \omega_{ab} j^{ab}} F]^{ab}(\Lambda^{-1} x),$$

(1.171)
where

\[ \hat{J}^{cd} \equiv L^{cd} \times 1 + 1 \times L^{cd} \quad (1.172) \]

are the generators of the total four-dimensional angular momentum of the tensor field. The factors in the direct products apply successively to the representation spaces associated with the two Lorentz indices and the spacetime coordinates. The generators \( \hat{J}^{ab} \) obey the same commutation rules (1.71) and (1.72) as \( L^{ab} \) and \( \hat{L}^{ab} \).

In order to verify the transformation law (1.170), we recall the basic result of electromagnetism, that under a change to a coordinate frame \( x \rightarrow x' = \Lambda x \) moving with a velocity \( v \), the electric and magnetic fields change as follows

\[ E'_\parallel(x') = E_\parallel(x), \quad E'_\perp(x') = \gamma \left[ E_\perp(x) + \frac{1}{c} v \times B(x) \right], \quad (1.173) \]
\[ B'_\parallel(x') = B_\parallel(x), \quad B'_\perp(x') = \gamma \left[ B_\perp(x) - \frac{1}{c} v \times E(x) \right], \quad (1.174) \]

where the subscripts \( \parallel \) and \( \perp \) denote the components parallel and orthogonal to \( v \).

Recalling the matrices (1.27) we see that (1.173) and (1.174) correspond precisely to the transformation law (1.170) of a tensor field.

The field tensor in the electromagnetic force of the equation of motion (1.165) transforms accordingly:

\[ F^a_b(x(\tau)) \xrightarrow{\Lambda} F'^a_b(x'(\tau)) = \Lambda^c_a \Lambda^d_b F^{ae} dF^e_d(\Lambda^{-1}x(\tau)). \quad (1.175) \]

This can be verified by rewriting \( F^a_b(x(\tau)) \) as

\[ F^a_b(x(\tau)) = \int d^4x \, F^a_b(x) \, \delta(4)(x - x(\tau)), \quad (1.176) \]

and applying the transformation (1.170).

Separating time and space components of the Lorentz force (1.166) we find

\[ \frac{d}{d\tau} p^0 = f^0 = \frac{e}{M} \mathbf{E} \cdot \mathbf{p}, \quad (1.177) \]
\[ \frac{d}{d\tau} \mathbf{p} = \mathbf{f} = \frac{e}{M} \left( \mathbf{E} p^0 + \mathbf{p} \times \mathbf{B} \right). \quad (1.178) \]

The Lorentz force can also be stated in terms of velocity as

\[ f^a = \frac{e}{c} F^a_b \frac{dx^b}{d\tau} = \gamma \left( \frac{e}{c} \mathbf{v} \cdot \mathbf{E} c\mathbf{E}^i + \frac{1}{c} (\mathbf{v} \times \mathbf{B})^i \right). \quad (1.179) \]

The above equations rule the movement of charged point particles in a given external field. The moving particles will, however, also give rise to additional electromagnetic fields. These are calculated by solving the Maxwell equations in the presence of charge and current densities \( \rho \) and \( j \), respectively:

\[ \nabla \cdot \mathbf{E} = \rho \quad \text{(Coulomb’s law)}, \quad (1.180) \]
\[ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c} \mathbf{j} \quad \text{(Ampre’s law)}, \]
\[ \nabla \cdot \mathbf{B} = 0 \quad \text{(absence of magnetic monopoles)}, \]
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{(Faraday’s law)}. \]

In a dielectric and paramagnetic medium with dielectric constant \( \epsilon \) and magnetic permeability \( \mu \) one defines the displacement field \( \mathbf{D}(x) \) and the magnetic field \( \mathbf{H}(x) \) by the relations
\[ \mathbf{D}(x) = \epsilon \mathbf{E}(x), \quad \mathbf{B}(x) = \mu \mathbf{B}(x), \]
and the Maxwell equations become
\[ \nabla \cdot \mathbf{D} = \rho \quad \text{(Coulomb’s law)}, \]
\[ \nabla \times \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{1}{c} \mathbf{j} \quad \text{(Ampre’s law)}, \]
\[ \nabla \cdot \mathbf{B} = 0 \quad \text{(absence of magnetic monopoles)}, \]
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{(Faraday’s law)}. \]

On the right-hand sides of (1.180), (1.181) and (1.185), (1.186) we have omitted factors \( 4\pi \), for convenience. This makes the charge of the electron equal to \( -e = -\sqrt{\frac{\alpha}{4\pi}} \approx \frac{1}{137.035989} \), where
\[ \alpha \approx 1/137.035989 \]
is the fine-structure constant.

In the vacuum, the two inhomogeneous Maxwell equations (1.180) and (1.181) can be combined to a single equation
\[ \partial_a F^{ab} = -\frac{1}{c} j^a, \]
where \( j^a \) is the four-component current density
\[ j^a(x) = \left( \begin{array}{c} \rho(x, t) \\ \mathbf{j}(x, t) \end{array} \right). \]

Indeed, the zeroth component of (1.190) is equal to (1.180):
\[ \partial_0 F^{0i} = -\nabla \cdot \mathbf{E} = -\rho, \]
whereas the spatial components with \( a = i \) reduce to Eq. (1.181):
\[ \partial_0 F^{i0} + \partial_j F^{ij} = \partial_j \epsilon^{ijk} \mathbf{B}^k + \frac{1}{c} \frac{\partial E^i}{\partial t} = -\left( \nabla \times \mathbf{B} \right)^i + \frac{1}{c} \frac{\partial E^i}{\partial t} = -\frac{1}{c} j^i. \]

The remaining homogeneous Maxwell equations (1.182) and (1.183) can also be rephrased in tensor form as
\[ \partial_a \mathbf{F}^{ab} = 0. \]
Here \( \tilde{F}^{ab} \) is the so-called dual field tensor defined by
\[
\tilde{F}^{ab} = \epsilon^{abcd} F_{cd}.
\] (1.195)
where \( \epsilon^{abcd} \) is the totally antisymmetric unit tensor with \( \epsilon^{0123} = 1 \).

The antisymmetry of \( F^{ab} \) in (1.190) implies the vanishing of the four-divergence of the current density:
\[
\partial_a j^a(x) = 0.
\] (1.196)
This is the four-dimensional way of expressing the local conservation law of charges. Written out in space and time components it reads
\[
\partial_t \rho(x, t) + \nabla \cdot j(x, t) = 0.
\] (1.197)
Integrating this over a finite volume gives
\[
\partial_t \left[ \int d^3 x \rho(x, t) \right] = -\int d^3 x \nabla \cdot j(x, t) = 0.
\] (1.198)
The right-hand side vanishes by the Gauss divergence theorem, according to which the volume integral over the divergence of a current density is equal to the surface integral over the flux through the boundary of the volume. This vanishes if currents do not leave a finite spatial volume, which is usually true for an infinite system. Thus we find that, as a consequence of local conservation law (1.196), the charge of the system
\[
Q(t) \equiv \int d^3 \rho(x, t) \equiv \frac{1}{c} \int d^3 x j^0(x)
\] (1.199)
satisfies the global conservation law that the charge is time-independent
\[
Q(t) \equiv Q.
\] (1.200)

For a set of point particles of charges \( e_n \), the charge and current densities are
\[
\rho(x, t) = \sum_n e_n \delta^{(3)}(x - x_n(t)),
\] (1.201)
\[
j(x, t) = \sum_n e_n \dot{x}_n(t) \delta^{(3)}(x - x_n(t)).
\] (1.202)
Combining these expressions to a four-component current density (1.191), we can easily verify that \( j^a(x) \) transforms like a vector field [compare with the behaviors (1.163) of scalar field and (1.170) of tensor fields]:
\[
j^a(x) \xrightarrow{\Lambda} j'^a(x) = \Lambda^a_b j^b(\Lambda^{-1} x).
\] (1.203)
To verify this we note that \( \delta^{(3)}(x - x(t)) \) can also be written as an integral along the path of the particle parametrized with the help of the proper time \( \tau \). This is done with the help of the identity
\[
\int_{-\infty}^{\infty} d\tau \delta^{(4)}(x - x(\tau)) = \int_{-\infty}^{\infty} d\tau \delta(x^0 - x^0(\tau)) \delta^{(3)}(x - x(\tau))
\] 
\[= \frac{dx^0}{d\tau} \delta^{(3)}(x - x(t)) = \frac{1}{c\gamma} \delta^{(3)}(x - x(t)).
\] (1.204)
This allows us to rewrite (1.201) and (1.202) as
\[
cp(x, t) = c \sum_n \int_{-\infty}^{\infty} d\tau_n e_n \gamma_n c \delta^{(4)}(x - x_n(\tau)),
\]
(1.205)
\[
j(x, t) = c \sum_n \int_{-\infty}^{\infty} d\tau_n e_n \gamma_n \nu_n \delta^{(4)}(x - x_n(\tau)).
\]
(1.206)

These equations can be combined in a single four-vector equation
\[
j^a(x) = c \sum_n \int_{-\infty}^{\infty} d\tau_n e_n \dot{x}^a_n(\tau) \delta^{(4)}(x - x_n(\tau)),
\]
(1.207)
which makes the transformation behavior (1.203) an obvious consequence of the vector nature of \( \dot{x}^a_n(\tau) \).

In terms of the four-dimensional current density, the inhomogeneous Maxwell equation (1.190) becomes the Maxwell-Lorentz equation
\[
\partial_b F^{ab} = -\frac{1}{c} j^a = -\sum_n \int_{-\infty}^{\infty} d\tau_n e_n \dot{x}^a_n(\tau) \delta^{(4)}(x - x_n(\tau)).
\]
(1.208)

It is instructive to verify the conservation law (1.196) for the current density (1.207). Applying the derivative \( \partial_a \) to the \( \delta \)-function gives \( \partial_a \delta^{(4)}(x - x_n(\tau)) = -\partial \dot{x}^a_n \delta^{(4)}(x - x_n(\tau)) \), and therefore
\[
\partial_a j^a(x) = -c \sum_n \int_{-\infty}^{\infty} d\tau_n e_n \frac{dx^a_n(\tau)}{d\tau} \partial x^a_n(\tau) \delta^{(4)}(x - x_n(\tau))
\]
\[
= -c \sum_n \int_{-\infty}^{\infty} d\tau_n e_n \partial_x \delta^{(4)}(x - x_n(\tau)).
\]
(1.209)

If the particle orbits \( x(\tau) \) are stable, they are either closed in spacetime, or come from negative infinite \( x_0 \) and run to positive infinite \( x_0 \). Then the right-hand side vanishes in any finite volume so that the current density is indeed conserved.

We end this section by remarking, that the vector transformation law (1.203) can also be written by analogy with the tensor law (1.170) as
\[
j^a(x) \xrightarrow{\Lambda} j'^a(x) = [e^{-i\frac{\omega_{ab}}{2} J^{ab}} j^a]^{\Lambda^{-1}}(x),
\]
(1.210)
where
\[
\hat{J}^{cd} \equiv L^{cd} \times \hat{1} + 1 \times \hat{L}^{cd}
\]
(1.211)
are the generators of the total four-dimensional angular momentum of the vector field. As in (1.172), the factors in the direct products apply separately to the representation spaces associated with the Lorentz index and the spacetime coordinates, and the generators \( \hat{J}^{ab} \) obey the same commutation rules (1.71) and (1.72) as \( L_{ab} \) and \( \hat{L}_{ab} \).
1.11 Dirac Field

The observable matter of the universe consists mainly of electrons and nucleons, the latter being predominantly bound states of three quarks. Electrons and quarks are spin-1/2 particles which may be described by four-component Dirac fields $\psi(x)$. These obey the Dirac equation

$$(i\gamma^a \partial_a - m) \psi(x) = 0,$$

where $\gamma^a$ are the $4 \times 4$ Dirac matrices

$$\gamma^a = \begin{pmatrix} 0 & \sigma^a \\ \bar{\sigma}^a & 0 \end{pmatrix},$$

in which the $2 \times 2$ submatrices $\sigma^a$ and $\bar{\sigma}^a$ with $a = 0, \ldots, 3$ form the four-vectors of Pauli matrices

$$\sigma^a \equiv (\sigma^0, \sigma^i), \quad \bar{\sigma}^a \equiv (\sigma^0, -\sigma^i).$$

The spatial components $\sigma^i$ are the ordinary Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

while the zeroth component $\sigma^0$ is defined as the $2 \times 2$ unit matrix:

$$\sigma^0 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

From the algebraic properties of these matrices

$$(\sigma^a)^2 = 1, \quad \sigma^a \sigma^b + \sigma^b \sigma^a = 2g^{ab},$$

we deduce that the Dirac matrices $\gamma^a$ satisfy the anticommutation rules

$$\{ \gamma^a, \gamma^b \} = 2g^{ab}.$$

Under Lorentz transformations, the Dirac field transforms according to the spinor representation of the Lorentz group

$$\psi_A(x) \xrightarrow{\Lambda} \psi'_A(x) = D_A^B(\Lambda)\psi_B(\Lambda^{-1}x),$$

by analogy with the transformation law (1.203) of a vector field. The $4 \times 4$-matrices $\Lambda$ of the defining representation of the Lorentz group in (1.203) are replaced by the $4 \times 4$-matrices $D(\Lambda)$ representing the Lorentz group in spinor space.

It is easy to find these matrices. If we denote the spinor representation of the Lie algebra (1.72) by $4 \times 4$-matrices $\Sigma^{ab}$, these have to satisfy the commutation rules

$$[\Sigma^{ab}, \Sigma^{ac}] = -ig^{ac}\Sigma^{bc}, \quad \text{no sum over } a.$$
These can be solved by the matrices
\[ \Sigma^{ab} \equiv \frac{1}{2} \sigma^{ab}, \]

where \( \sigma^{ab} \) is the antisymmetric tensor of matrices
\[ \sigma^{ab} \equiv \frac{i}{2} [\gamma^a, \gamma^b]. \]

The representation matrices of finite Lorentz transformations may now be expressed as exponentials of the form (1.54):
\[ D(\Lambda) = e^{-i \frac{1}{2} \omega_{ab} \Sigma^{ab}}, \]
where \( \omega_{ab} \) is the same antisymmetric matrix as in (1.54), containing the rotation and boost parameters as specified in (1.55) and (1.56). Comparison with (1.57) shows that pure rotations and pure Lorentz transformations are generated by the spinor representations of \( L^{ab} \) in (1.57):
\[ \Sigma^{ij} = \epsilon_{ijk} \frac{1}{2} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \quad \Sigma^{0i} = \frac{i}{2} \begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \]

The generators of the rotation group \( \Sigma^i = \frac{1}{2} \epsilon_{ijk} \Sigma^{jk} \) corresponding to \( L_i \) in (1.53) consist of a direct sum of two Pauli matrices, the \( 4 \times 4 \)-spin matrix:
\[ \Sigma \equiv \frac{1}{2} \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}. \]

The generators \( \Sigma^{0i} \) of the pure Lorentz transformations corresponding to \( M_i \) in (1.53) can also be expressed as \( \Sigma^{0i} = i \alpha^i / 2 \) with the vector of \( 4 \times 4 \)-matrices
\[ \alpha = \begin{pmatrix} -\sigma & 0 \\ 0 & \sigma \end{pmatrix}. \]

In terms of \( \Sigma \) and \( \alpha \), the representation matrices (1.223) for pure rotations and pure Lorentz transformations are seen to have the explicit form
\[ D(R) = e^{-i \varphi \Sigma} = \begin{pmatrix} e^{-i \varphi \sigma/2} & 0 \\ 0 & e^{-i \varphi \sigma/2} \end{pmatrix}, \quad D(B) = e^{i \alpha \Sigma} = \begin{pmatrix} e^{-\zeta \sigma/2} & 0 \\ 0 & e^{\zeta \sigma/2} \end{pmatrix}. \]

The commutation relations (1.220) are a direct consequence of the commutation relations of the generators \( \Sigma^{ab} \) with the gamma matrices:
\[ [\Sigma^{ab}, \gamma^c] = - (L^{ab})^c_d \gamma^d = -i (\epsilon^{ac} \gamma^b - \epsilon^{bc} \gamma^a). \]

Comparison with (1.110) and (1.111) shows that the matrices \( \gamma^a \) transform like \( x^a \), i.e., they form a vector operator. The commutation rules (1.220) follow directly from (1.228) upon using the Leibnitz chain rule (1.113).
For global transformations, the vector property (1.228) implies that $\gamma^a$ behaves like the vector $x^a$ in Eq. (1.129):

$$\hat{D}(\Lambda)\gamma^c \hat{D}^{-1}(\Lambda) = e^{-i\frac{1}{2}\omega_{ab}^{\Sigma} \Sigma^{ab}} = (e^{i\frac{1}{2}\omega_{ab}^{\Sigma} \Sigma^{ab}})^c = (\Lambda^{-1})^c e^{i\frac{1}{2}\omega_{ab}^{\Sigma} \Sigma^{ab}} = (\Lambda^{-1})^c \gamma^c = (\Lambda^a_b x^b).$$

(1.229)

In terms of the generators $\Sigma^{ab}$, we can write the field transformation law (1.219) more explicitly as [compare with the behavior of scalar (1.163), tensor (1.170), and vector fields (1.203)]:

$$\psi(x) \xrightarrow{\Lambda} \psi'(x) = D(\Lambda) \psi(\Lambda^{-1} x) = e^{-i\frac{1}{2}\omega_{ab}^{\Sigma} \Sigma^{ab}} \psi(\Lambda^{-1} x),$$

(1.230)

It is useful to re-express the transformation of the spacetime argument on the right-hand side in terms of the differential operator of four-dimensional angular momentum and rewrite (1.230) as in (1.172) and (1.211) as

$$\psi(x) \xrightarrow{\Lambda} \psi'(x) = \hat{D}(\Lambda) \times D(\Lambda) \psi(x) = e^{-i\frac{1}{2}\omega_{ab}^{\Sigma} \Sigma^{ab}} \psi(x),$$

(1.231)

where

$$\hat{J}^{cd} \equiv \Sigma^{cd} \times \hat{1} + 1 \times \hat{L}^{cd}.$$

(1.232)

are the generators of the total four-dimensional angular momentum of the Dirac field.

### 1.12 Spacetime-Dependent Lorentz Transformations

The theory of gravitation to be developed in this book will not only be Lorentz-invariant, but also invariant under local Lorentz transformations.

$$x'^a = \Lambda^a_b(x^b).$$

(1.233)

As a preparation for dealing with such theories let us derive a group-theoretic formula which is useful for many purposes.

#### 1.12.1 Angular Velocities

Consider a time-dependent $3 \times 3$-rotation matrix $R(\varphi(t)) = e^{-i\varphi(t) L}$ with the generators $(L_i)_{jk} = -i\epsilon_{ijk}$ [compare (1.43)]. As time proceeds, the rotation angles change with an angular velocity $\omega(t)$ defined by the following relation

$$R^{-1}(\varphi(t)) \dot{R}(\varphi(t)) = -i\omega(t) \cdot L.$$ 

(1.234)

The components of $\omega(t)$ can be specified more explicitly by parametrizing the rotations in terms of Euler angles $\alpha$, $\beta$, $\gamma$:

$$R(\alpha, \beta, \gamma) = R_3(\alpha) R_2(\beta) R_3(\gamma),$$

(1.235)
where \( R_1(\alpha) \), \( R_2(\gamma) \) are rotations around the \( z \)-axis by angles \( \alpha \), \( \gamma \), respectively, and \( R_2(\beta) \) is a rotation around the \( y \)-axis by \( \beta \), i.e.,

\[
R(\alpha, \beta, \gamma) \equiv e^{-i\alpha L_3} e^{-i\beta L_2} e^{-i\gamma L_3}.
\]

The relations between the vector \( \varphi \) of rotation angles in (1.57) and the Euler angles \( \alpha, \beta, \gamma \) can be found by purely geometric considerations. Most easily, we equate the \( 2 \times 2 \) representation of the rotations \( R(\varphi) \),

\[
R(\varphi) = \cos \frac{\varphi}{2} - i \sigma \cdot \varphi \sin \frac{\varphi}{2},
\]

with the \( 2 \times 2 \) representation of the Euler decomposition (1.236):

\[
R(\alpha, \beta, \gamma) = \left( \cos \frac{\alpha}{2} - i \sigma_3 \sin \frac{\alpha}{2} \right) \left( \cos \frac{\beta}{2} - i \sigma_2 \sin \frac{\beta}{2} \right) \left( \cos \frac{\gamma}{2} - i \sigma_3 \sin \frac{\gamma}{2} \right).
\]

The desired relations follow directly from the multiplication rules for the Pauli matrices (1.217).

In the Euler decomposition, we may calculate the derivatives:

\[
\begin{align*}
\frac{\hbar}{i} \partial_\alpha R &= R \left[ \cos \beta L_3 - \sin \beta (\cos \gamma L_1 - \sin \gamma L_2) \right], \\
\frac{\hbar}{i} \partial_\beta R &= R (\cos \gamma L_2 + \sin \gamma L_1), \\
\frac{\hbar}{i} \partial_\gamma R &= RL_3.
\end{align*}
\]

The third equation is trivial, the second follows from the rotation of the generator

\[
e^{i\gamma L_3/\hbar} L_2 e^{-i\gamma L_3/\hbar} = \cos \alpha L_2 + \sin \gamma L_1,
\]

which is a consequence of Lie’s expansion formula

\[
e^{iA}Be^{-iA} = 1 + i[A, B] + \frac{i^2}{2!}[A, [A, B]] + \ldots,
\]

and the commutation rules (1.61) of the \( 3 \times 3 \) matrices \( L_i \). The derivation of the first equation (1.239) requires, in addition, the rotation

\[
e^{i\beta L_2/\hbar} L_3 e^{-i\beta L_2/\hbar} = \cos \beta L_3 - \sin \beta L_1.
\]

We may now calculate the time derivative of \( R(\alpha, \beta, \gamma) \) using Eqs. (1.239)–(1.241) and the chain rule of differentiation, and find the right-hand side of (1.234) with the angular velocities

\[
\begin{align*}
\omega_1 &= \dot{\beta} \sin \gamma - \dot{\alpha} \sin \beta \cos \gamma, \\
\omega_2 &= \dot{\beta} \cos \gamma + \dot{\alpha} \sin \beta \sin \gamma, \\
\omega_3 &= \dot{\alpha} \cos \beta + \dot{\gamma}.
\end{align*}
\]

Only commutation relations have been used to derive (1.239)–(1.241), so that the formulas (1.245)–(1.247) hold for all representations of the rotation group.
1.12.2 Angular Gradients

The concept of angular velocities can be generalized to spacetime-dependent Euler angles $\alpha(x)$, $\beta(x)$, $\gamma(x)$, replacing (1.234) by angular gradients

$$R^{-1}(\varphi(x)) \partial_a R(\varphi(x)) = -i \omega_a(x) \cdot L,$$  \hspace{1cm} (1.248)

with the generalization of the vector of angular velocity

$$\omega_{a;1} = \partial_a \beta \sin \gamma - \partial_a \alpha \sin \beta \gamma,$$

$$\omega_{a;2} = \partial_a \beta \cos \gamma + \partial_a \alpha \sin \beta \sin \gamma,$$

$$\omega_{a;3} = \partial_a \alpha \cos \beta + \partial_a \gamma.$$  \hspace{1cm} (1.249-1.251)

The derivatives $\partial_a$ act only upon the functions right after it. These equations are again valid if $R(\varphi(x))$ and $L$ in (1.248) are replaced by any representation of the rotation group and its generators.

A relation of type (1.248) exists also for the Lorentz group where $\Lambda(\omega_{ab}(x)) = e^{-\frac{i}{2} \omega_{ab}(x) L^{ab}}$ [recall (1.57)], where the generalized angular velocities are defined by

$$\Lambda^{-1}(\omega_{ab}(x)) \partial_c \Lambda(\omega_{ab}(x)) = -i \frac{1}{2} \omega_{c;ab}(x) L^{ab}.$$  \hspace{1cm} (1.252)

Inserting the explicit $4 \times 4$-generators (1.51) on the right-hand side, we find for the matrix elements the relation

$$[\Lambda^{-1}(\omega_{ab}(x)) \partial_c \Lambda(\omega_{ab}(x))]_{ef} = \omega_{c;ef}(x).$$  \hspace{1cm} (1.253)

As before, the matrices $\Lambda(\omega_{ab}(x))$ and $L^{ab}$ in (1.252) can be replaced by any representations of the Lorentz group and its generators, in particular in the spinor representation (1.223) where

$$D^{-1}(\Lambda(\omega_{ab}(x))) \partial_c D(\Lambda(\omega_{ab}(x))) = -i \frac{1}{2} \omega_{c;ab}(x) \Sigma^{ab}.$$  \hspace{1cm} (1.254)

1.13 Energy-Momentum Tensors

The four-dimensional current density $j^a(x)$ contains all information on the electric properties of relativistic particle orbits. It is possible to collect also the mechanical properties in a tensor, the energy-momentum tensor.

1.13.1 Point Particles

The energy density of the particles can be written as

$$E^{\text{part}}(x, t) = \sum_n m_n c^2 \delta^{(3)}(x - x_n(t)).$$  \hspace{1cm} (1.255)
We have previously seen that the energy transforms like a zeroth component of a four-vector [recall (1.147)]. The energy density measures the energy per spatial volume element. An infinitesimal four-volume $d^4x$ is invariant under Lorentz transformations, due to the unit determinant $|Λ^a_b| = 1$ implied by the pseudo-orthogonality relation (1.28), so that indeed

$$d^4x' = \left| \frac{∂x'^a}{∂x^b} \right| d^4x = |Λ^a_b| d^4x = d^4x. \quad (1.256)$$

This shows that $δ^{(3)}(x)$ which transforms like an inverse spatial volume

$$\frac{1}{d^3x} = \frac{dx^0}{d^4x} \quad (1.257)$$

behaves like the zeroth component of a four-vector. The energy density (1.255) can therefore be viewed as a 00-component of a Lorentz tensor called the symmetric energy-momentum tensor. By convention, this is chosen to have the dimension of momentum density, so that we must identify the energy density with $c part T^{ab}$. In fact, using the identity (1.204), we may rewrite (1.255) as

$$part E(x,t) = c \sum_n \int_{-∞}^{∞} dτ_n \frac{1}{m_n} p^0_n(τ)p^0_n(τ)δ^{(4)}(x - x(τ)), \quad (1.258)$$

which is equal to $c$ times the 00-component of the energy-momentum tensor

$$part T^{ab}(x,t) = \sum_n \int_{-∞}^{∞} dτ_n \frac{1}{m_n} p^n_a(τ)p^n_b(τ)δ^{(4)}(x - x(τ)). \quad (1.259)$$

The spatial momenta of the particles

$$part P^i(x,t) = \sum_n m_n γ_n \dot{x}^i_n(τ)δ^{(3)}(x - x(τ)) \quad (1.260)$$

are three-vectors. Their densities transform therefore like 0i-components of a Lorentz tensor. Indeed, using once more the identity (1.204), we may rewrite (1.260) as

$$part P^i(x,t) = T^{0i}(x,t) = \sum_n \int_{-∞}^{∞} dτ_n \frac{1}{m_n} p^0_n(τ)p^i_n(τ)δ^{(4)}(x - x(τ)), \quad (1.261)$$

which shows precisely the tensor character. The four-vector of the total energy-momentum of the many-particle system is given by the integrals over the 0a-components

$$part P^a(t) ≡ \int d^3x \; part T^{0a}(x,t). \quad (1.262)$$

Inserting here (1.258) and (1.261) we obtain the sum over all four-momenta

$$part P^a(t) = \sum_n p^a_n(τ). \quad (1.263)$$
By analogy with the four-dimensional current density \( j^a(x) \), let us calculate the four-divergence \( \partial_b T^{ab} \). A partial integration yields

\[
\sum_n \int_{-\infty}^{\infty} d\tau_n p_n^a(\tau) \partial_b \delta^4(x - x(\tau)) = -\sum_n \int_{-\infty}^{\infty} d\tau_n p_n^a(\tau) \partial_b \delta^4(x - x(\tau)) - \sum_n \int_{-\infty}^{\infty} d\tau_n \hat{p}_n^a(\tau) \partial_b \delta^4(x - x(\tau)).
\]

The first term on the right-hand side disappears if the particles are stable, i.e., if the orbits are closed or come from negative infinite \( x^0 \) and disappear into positive infinite \( x^0 \). The derivative \( \hat{p}_n^a(\tau) \) in the second term can be made more explicit if only electromagnetic forces act on the particles. Then it is equal to the Lorentz force, i.e., the four-vector \( f^a(\tau) \) of Eq. (1.179), and we obtain

\[
\partial_b T^{ab} = \sum_n \int_{-\infty}^{\infty} d\tau_n f_n^a(\tau) \delta^4(x - x(\tau))
\]

Expressed in terms of the current four-vector (1.207), this reads

\[
\partial_b T^{ab}(x) = \frac{1}{c^2} F^a_b(x) j^b(x).
\]

In the absence of electromagnetic fields, the energy-momentum tensor of the particles is conserved.

Integrating (1.262) over the spatial coordinates gives the time change of the total four-momentum

\[
\partial_t \text{part} P^a(t) = c \partial_0 \left[ \int d^3x \text{part} T^{a0} \right] = c \int d^3x \partial_b \text{part} T^{ab} - c \int d^3x \partial_i \text{part} T^{ai}
\]

This agrees, of course, with the Lorentz equations (1.165) since by (1.263)

\[
\partial_t \text{part} P^a(t) = \partial_t \sum_n p_n^a(\tau) = \sum_n \dot{p}_n^a(\tau) \gamma_n.
\]

If there are no electromagnetic forces, then \( P^a \) is time-independent.

### 1.13.2 Electromagnetic Field

The electromagnetic field possesses an energy-momentum tensor of its own. The energy density is well-known:

\[
\mathcal{E}(x) = \frac{1}{2} \left[ E^2(x) + B^2(x) \right] = c \frac{\varepsilon \text{EM}}{c^4} T^{00}(x).
\]
The momentum density \( T^{0i}(x) \) is given by the Poynting vector
\[
S(x) = c \mathbf{E}(x) \times \mathbf{B}(x) : \quad (1.270)
\]
and the relation is
\[
S^i(x) = c^2 T^{0i}. \quad (1.271)
\]
The densities (1.269) and (1.271) can be combined to the energy-momentum tensor
\[
T^{ab}(x) = \frac{1}{c} \left[ -F^a_c F^{bc} + \frac{1}{4} g^{ab} F^{cd} F_{cd} \right]. \quad (1.272)
\]
The four-divergence of this is
\[
\partial_b T^{ab}(x) = \frac{1}{c} \left[ -F^a_c \partial_b F^{bc} - (\partial_b F^{a}_c) F^{bc} + \frac{1}{4} \partial^a \left( F^{cd} F_{cd} \right) \right]. \quad (1.273)
\]
The second and third terms cancel each other, due to the homogeneous Maxwell equations (1.182) and (1.183). In order to see this, take the trivial identity
\[
\partial_b \epsilon_{acdefg} F^{efg} = 0,
\]
we find
\[
-F^{cd} \partial_c F_{cd} - F^{db} \partial_b F_{ed} - F^{bc} \partial_b F_{ce} + F^{dc} \partial_d F_{cd} + F^{ec} \partial_e F_{ce} + F^{bd} \partial_b F_{ed} = 0. \quad (1.275)
\]
Due to the antisymmetry of \( F_{ab} \), this gives
\[
-\partial_e \left( F^{cd} F_{cd} \right) + 4 F^{bd} \partial_b F_{bd} = 0, \quad (1.276)
\]
so that we obtain the conservation law
\[
\partial_b T^{ab}(x) = -\frac{1}{c} \left[ F^a_c(x) \partial_b F^{bc}(x) \right] = 0. \quad (1.277)
\]
In the last step we have used Maxwell’s equation Eq. (1.190) with zero currents.

The timelike component of the conservation law (1.277) reads
\[
\partial_t T^{0i}(x) + c \partial_i T^{0i}(x) = 0, \quad (1.278)
\]
which can be rewritten with (1.269) and (1.271) as the well-known Poynting law of energy flow:
\[
\partial_t \mathcal{E}(x) + \nabla \cdot \mathbf{S}(x) = 0. \quad (1.279)
\]
If currents are present, the Maxwell equation (1.190) changes the conservation law (1.277) to
\[ c \partial_b \, T^{ab}(x) = -\frac{1}{c} F^a\epsilon(x) j^c(x) = 0, \] (1.280)
which modifies (1.279) to
\[ \partial_t \mathcal{E}(x) + \nabla \cdot \mathbf{S}(x) = -\mathbf{j}(x) \cdot \mathbf{E}(x). \] (1.281)

A current parallel to the electric field reduces the field energy.

In a medium, the energy density and Poynting vector become
\[ \mathcal{E}(x) \equiv \frac{1}{2} \left[ \mathbf{E}(x) \cdot \mathbf{D}(x) + \mathbf{B}(x) \cdot \mathbf{H}(x) \right], \quad \mathbf{S}(x) \equiv c \mathbf{E}(x) \times \mathbf{H}(x), \] (1.282)
and the conservation law can easily be verified using the Maxwell equations (1.186) and (1.188):
\[
\nabla \cdot \mathbf{S}(x) = c \nabla \cdot \left[ \mathbf{E}(x) \times \mathbf{H}(x) \right] = c \left[ \nabla \times \mathbf{E}(x) \cdot \mathbf{H}(x) - c \mathbf{E}(x) \cdot \nabla \times \mathbf{B}(x) \right] = \{ \partial_i \mathbf{B}(x) \cdot \mathbf{H}(x) + \mathbf{E}(x) \cdot [\partial_i \mathbf{D}(x) + j(x)] \} = \partial_t \mathcal{E}(x) + j(x) \cdot \mathbf{E}(x). \] (1.283)

We now observe that the force on the right-hand side of (1.280) is precisely the opposite of the right-hand side of (1.266), as required by Newton’s third axiom of actio = reactio. Thus, the total energy-momentum tensor of the combined system of particles and electromagnetic fields
\[ T^{ab}(x) = T^{ab}_{\text{part}}(x) + T^{ab}_{\text{em}}(x) \] (1.284)
has a vanishing four-divergence,
\[ \partial_b T^{ab}(x) = 0 \] (1.285)
implying that the total four-momentum \( P^a \equiv \int d^3 x \, T^{0a} \) is a conserved quantity
\[ \partial_t P^a(t) = 0. \] (1.286)

### 1.14 Angular Momentum and Spin

Similar considerations apply to the total angular momentum of particles and fields. Since \( T^{0a}(x) \) is a momentum density, we may calculate the spatial tensor of total angular momentum from the integral
\[ J^{ij}(t) = \int d^3 x \left[ x^i T^{j0}(x) - x^j T^{i0}(x) \right]. \] (1.287)

In three space dimensions one describes the angular momentum by a vector \( J^i = \frac{1}{2} \epsilon^{ijk} J^k \). The angular momentum (1.287) may be viewed as the integral
\[ J^{ij}(t) = \int d^3 x \, J^{ij,0}(x). \] (1.288)
over the $i, j, 0$-component of the Lorentz tensor

$$J^{ab,c}(x) = x^a T_{bc}(x) - x^b T_{ac}(x). \quad (1.289)$$

It is easy to see that due to (1.285) and the symmetry of the energy-momentum tensor, the Lorentz tensor $J^{ab,c}(x)$ is divergenceless in the index $c$

$$\partial_c J^{ab,c}(x) = 0. \quad (1.290)$$

As a consequence, the spatial integral

$$J^{ab}(t) = \int d^3 x J^{ab,0}(x) \quad (1.291)$$

is a conserved quantity. This is the four-dimensional extension of the conserved total angular momentum. The conservation of the components $J^{0i}$ is the center-of-mass theorem.

A set of point particles with the energy-momentum tensor (1.259) possesses four-dimensional angular momentum

$$J^{ab}(\tau) = \sum_n \left[ x^a_n(\tau) p^b_n(\tau) - x^b_n(\tau) p^a_n(\tau) \right]. \quad (1.292)$$

In the absence of electromagnetic fields, this is conserved, otherwise the $\tau$-dependence is important.

The spin of a particle is defined by its total angular momentum in its rest frame. It is the intrinsic angular momentum of the particle. Electrons, protons, neutrons, and neutrinos have spin 1/2. For nuclei and atoms, the spin can take much larger values.

There exists a four-vector $S^a(\tau)$ along the orbit of a particle whose spatial part reduces to the angular momentum in the rest frame. It is defined by a combination of the angular momentum (1.292) and the four-velocity $u^d(\tau)$ [recall (1.145)]

$$S^a(\tau) \equiv \frac{1}{2c} \epsilon^{abcd} \sum_n \left[ x^a_n(\tau) p^b_n(\tau) - x^b_n(\tau) p^a_n(\tau) \right]. \quad (1.293)$$

In the rest frame where

$$u^a_R = (c, 0, 0, 0), \quad (1.294)$$

this reduces indeed to the three-vector of total angular momentum

$$S^a_R(\tau) = (0, \sum_n \left[ x^a_n(\tau) p^n(\tau) - x^n(\tau) p^a(\tau) \right]). \quad (1.295)$$

For a free particle we find, due to conservation of momentum and total angular momentum

$$\frac{d}{d\tau} u^a_d(\tau) = 0, \quad \frac{d}{d\tau} J^{ab,0}(\tau) = 0, \quad (1.296)$$

so that $S^a(\tau)$ is conserved:

$$\frac{d}{d\tau} S^a(\tau) = 0. \quad (1.297)$$
The spin four-vector is useful to understand an important phenomenon in atomic physics called the Thomas precession of the electron spin in an atom. It explains why the observed fine structure of atomic physics is compatible with the gyromagnetic ratio close to two of the magnetic moment of the electron.

The relation between the spin vector and the four-vector is displayed by applying the pure Lorentz transformation matrix (1.27) to (1.295) yielding

$$S^i = S_R^i + \frac{v^i v^j}{c^2} S_R^j, \quad S^0 = \frac{v^i}{c} S_R^i$$  \hspace{1cm} (1.298)

Note that $S^0$ and $S^i$ satisfy

$$u^a S_a = 0.$$  \hspace{1cm} (1.299)

The inverse of the transformation (1.298) is found with the help of the identity $v^2/c^2 = (\gamma^2 - 1)/\gamma^2$ as follows:

$$S_R^i = S^i - \frac{\gamma}{\gamma + 1} \frac{v^i v^j}{c^2} S^j = S^i - \frac{1}{\gamma} \frac{v^i v^j}{v^2} S^j.$$  \hspace{1cm} (1.300)

If external forces act on the system, the spin vector starts moving. This movement is called precession. If the point article moves in an orbit under the influence of a central force (for example, an electron around a nucleus in an atom), there is no torque on the particle so that the total angular momentum in its rest frame is conserved. Hence $dS_R^i(\tau)/d\tau = 0$, which is expressed covariantly as $dS^a(\tau)/d\tau \propto u^a(\tau)$. In the rest frame of the atom, however, the spin shows precession. Let us calculate its rate. From the definition (1.293) we have

$$\frac{dS_a}{d\tau} = \frac{1}{2} \epsilon_{abcd} \text{part} J_{bc} \frac{du^d}{d\tau},$$  \hspace{1cm} (1.301)

There is no contribution from

$$\frac{d}{d\tau} \text{part} J_{bc} = x^a(\tau) p^b(\tau) - x^b(\tau) p^a(\tau),$$  \hspace{1cm} (1.302)

since $\dot{p} = m \dot{\mu}$, and the $\epsilon$-tensor is antisymmetric.

The right-hand side of (1.301) can be simplified by multiplying it with the trivial expression

$$g_{st} u^s u^t = c^2,$$  \hspace{1cm} (1.303)

and using the identity for the $\epsilon$-tensor

$$\epsilon^{abcd} g_{st} = \epsilon^{dabc} g^{td} + \epsilon^{absc} g^{ct} + \epsilon^{ascd} g^{bt} + \epsilon^{sabcd} g^{at},$$  \hspace{1cm} (1.304)

which can easily be verified using its antisymmetry and choosing $a, b, c, d$ to be equal to 0, 1, 2, 3, respectively, in particular for $abcd = 0123$. After this, the right-hand side of (1.301) becomes a sum of the four terms

$$\frac{1}{2} \left( \epsilon^{abcd} \text{part} J_{bc} u^s u^d u^d + \epsilon_{abcd} \text{part} J^{bc} u_c u^s u^d + \epsilon_{ascd} \text{part} J^{bc} u_d u^a u^d + \epsilon_{sabcd} \text{part} J^{bc} u^s u^a u^d \right).$$
The first term vanishes, since \( u^d \dot{u}_d = (1/2) du^2/d\tau = (1/2) dc^2/d\tau = 0 \). The last term is equal to \(-S^a d\dot{u}_a/c^2\). Inserting the identity (1.304) into the second and third terms, we obtain twice the left-hand side of (1.301). Taking this to the left-hand side, we find the equation of motion

\[
\frac{dS_a}{d\tau} = \frac{1}{c^2} S_c \frac{du_c}{d\tau} u_a.
\]

(1.305)

Note that on account of this equation, the time derivative \( dS_a/d\tau \) points in the direction of \( u^a \), in accordance with the initial assumption of a torque-free force.

We are now prepared to calculate the rate of the Thomas precession. Denoting in the remainder of this section the derivatives with respect to the physical time \( t = \gamma \tau \) by a dot, we can rewrite (1.305) as

\[
\dot{S} \equiv \frac{dS}{dt} = \frac{1}{\gamma} \frac{dS}{d\tau} = -\frac{1}{c^2} \left( S^0 \dot{u}^0 + S \cdot \dot{u} \right) u = \frac{\gamma^2}{c^2} (S \cdot \dot{v}) v,
\]

(1.306)

\[
\dot{S}_0 \equiv \frac{dS_0}{dt} = \frac{1}{c} \frac{dS}{d\tau} = \frac{\gamma^2}{c^2} (S \cdot \dot{v}).
\]

(1.307)

We now differentiate Eq. (1.300) with respect to the time using the relation \( \dot{\gamma} = \frac{\gamma}{\gamma + 1} \frac{1}{c^2} v \times \dot{v} \), and find

\[
\dot{S}_R = \dot{S} - \frac{\gamma}{\gamma + 1} \frac{1}{c^2} S^0 \dot{v} - \frac{\gamma}{\gamma + 1} \frac{1}{c^2} S \cdot \dot{v} - \frac{\gamma^3}{(\gamma + 1)^2} \frac{1}{c^2} (v \cdot \dot{v}) S^0 v.
\]

(1.308)

Inserting here Eqs. (1.306) and (1.307), we obtain

\[
\dot{S}_R = \frac{\gamma^2}{\gamma + 1} \frac{1}{c^2} (S \cdot \dot{v}) v - \frac{\gamma}{\gamma + 1} \frac{1}{c^2} S^0 \dot{v} - \frac{\gamma^3}{(\gamma + 1)^2} (v \cdot \dot{v}) S^0 v.
\]

(1.309)

On the right-hand side we return to the spin vector \( S_R \) using Eqs. (1.298), and find

\[
\dot{S}_R = \frac{\gamma^2}{\gamma + 1} \frac{1}{c^2} [(S_R \cdot \dot{v}) v - (S_R \cdot v) \dot{v}] = \Omega_T \times S_R,
\]

(1.310)

with the Thomas precession frequency

\[
\Omega_T = -\frac{\gamma^2}{(\gamma + 1) c^2} v \times \dot{v}.
\]

(1.311)

This is a purely kinematic effect. If an electromagnetic field is present, there will be an additional dynamic precession. For slow particles, it is given by

\[
\dot{S} \equiv -S \times \Omega_{em} \approx \mu \times \left( B - \frac{v}{c} \times E \right),
\]

(1.312)

where \( \mu \) is the magnetic moment

\[
\mu = g \mu_B \frac{S}{2Mc} = \frac{eg}{2Mc} S,
\]

(1.313)
and \( g \) the dimensionless gyromagnetic ratio, also called Land factor. Recall the value of the Bohr magneton

\[
\mu_B \equiv \frac{e\hbar}{2Mc} \approx 3.094 \times 10^{-30} \text{ C cm} \approx 0.927 \times 10^{-20} \text{ erg gauss} \approx 5.788 \times 10^{-8} \text{ eV gauss}.
\]

(1.314)

If the electron moves fast, we transform the electromagnetic field to the electron rest frame by a Lorentz transformation (1.173), (1.174), and obtain an equation of motion for the spin becomes

\[
\dot{\mathbf{S}}_R = \mathbf{\mu} \times \mathbf{B}' = \mathbf{\mu} \times \left[ \gamma \left( \mathbf{B} - \frac{\mathbf{v}}{c} \times \mathbf{E} \right) - \frac{\gamma^2}{\gamma + 1} \frac{\mathbf{v}}{c} \left( \frac{\mathbf{v}}{c} \cdot \mathbf{B} \right) \right].
\]

(1.315)

Expressing \( \mathbf{\mu} \) via Eq. (1.313), this becomes

\[
\dot{\mathbf{S}}_R \equiv -\mathbf{S}_R \times \mathbf{\Omega}_{em} = \frac{eg}{2mc} \mathbf{S}_R \times \left[ \left( \mathbf{B} - \frac{\mathbf{v}}{c} \times \mathbf{E} \right) - \frac{\gamma}{\gamma + 1} \frac{\mathbf{v}}{c} \left( \frac{\mathbf{v}}{c} \cdot \mathbf{B} \right) \right],
\]

(1.316)

which is the relativistic generalization of Eq. (1.312). It is easy to see that the associated fully covariant equation is

\[
S''^a = \frac{g}{2mc} \left[ eF^{ab}S_b + \frac{1}{mc} p^a S_c \frac{d}{dt} p^c \right] = \frac{eg}{2mc} \left[ F^{ab}S_b + \frac{1}{mc^2} \frac{d}{dt} p^a S_c F^{ck} p_k \right].
\]

(1.317)

On the right-hand side we have inserted the relativistic equation of motion (1.165) of a point particle in an external electromagnetic field.

If we add to this the torque-free Thomas precession rate (1.305), we obtain the covariant Bargmann-Michel-Telegdi equation\(^\text{1}\)

\[
S''^a = \frac{1}{2mc} \left[ egF^{ab}S_b + \frac{g - 2}{mc} p^a S_c \frac{d}{dt} p^c \right] = \frac{e}{2mc} \left[ gF^{ab}S_b + \frac{g - 2}{mc^2} \frac{d}{dt} p^a S_c F^{ck} p_k \right].
\]

(1.318)

For the spin vector \( \mathbf{S}_R \) in the electron rest frame this implies a change in the electromagnetic precession rate in Eq. (1.316) to\(^\text{2}\)

\[
\frac{d\mathbf{S}}{dt} = \mathbf{\Omega}_{emT} \times \mathbf{S} \equiv (\mathbf{\Omega}_{em} + \mathbf{\Omega}_T) \times \mathbf{S}
\]

(1.319)

with a frequency given by the Thomas equation

\[
\mathbf{\Omega}_{emT} = -\frac{e}{mc} \left[ \left( \frac{g}{2} - 1 + \frac{1}{\gamma} \right) \mathbf{B} - \left( \frac{g}{2} - 1 \right) \frac{\gamma}{\gamma + 1} \frac{\mathbf{v}}{c} \left( \frac{\mathbf{v}}{c} \cdot \mathbf{B} \right) \frac{\mathbf{v}}{c} - \left( \frac{g}{2} - \frac{\gamma}{\gamma + 1} \right) \frac{\mathbf{v}}{c} \times \mathbf{E} \right].
\]

(1.320)


The contribution of the Thomas precession is the part without the gyromagnetic factor $g$:

$$\Omega_T = -\frac{e}{mc} \left[ - \left( 1 - \frac{1}{\gamma} \right) \mathbf{B} + \frac{\gamma}{\gamma+1} \left( \mathbf{v} \cdot \mathbf{B} \right) \mathbf{v} + \frac{\gamma}{\gamma+1} \mathbf{v} \times \mathbf{E} \right]. \quad (1.321)$$

This is agrees with the Thomas frequency (1.311) if we insert there the acceleration

$$\dot{\mathbf{v}}(t) = \frac{d}{dt} \frac{\mathbf{P}}{m} = \frac{e}{\gamma m} \left[ \mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} - \frac{\mathbf{v}}{c} \left( \mathbf{v} \cdot \mathbf{E} \right) \right], \quad (1.322)$$

which follows directly from (1.177) and (1.178).

The Thomas equation (1.320) can be used to calculate the time dependence of the helicity $h \equiv \mathbf{S} \cdot \hat{\mathbf{v}}$ of an electron, i.e., its component of the spin in the direction of motion. Using the chain rule of differentiation,

$$\frac{d}{dt} (\mathbf{S} \cdot \mathbf{v}) = \dot{\mathbf{S}} \cdot \mathbf{v} + \frac{1}{v} \left[ \mathbf{S} - (\mathbf{v} \cdot \mathbf{S}) \mathbf{v} \right] \frac{d}{dt} \mathbf{v} \quad (1.323)$$

and inserting (1.319) as well as the equation for the acceleration (1.322), we obtain

$$\frac{dh}{dt} = -\frac{e}{mc} \mathbf{S}_{R\perp} \cdot \left[ \left( \frac{g}{2} - 1 \right) \mathbf{v} \times \mathbf{B} + \left( \frac{g v}{2c} - \frac{e}{v} \right) \mathbf{E} \right], \quad (1.324)$$

where $\mathbf{S}_{R\perp}$ is the component of the spin vector orthogonal to $\mathbf{v}$. This equation shows that for a Dirac electron which has $g = 2$ the helicity remains constant in a purely magnetic field. Moreover, if the electron moves ultra-relativistically ($v \approx c$), the value $g = 2$ makes the last term extremely small, $\approx (e/mc)g^{-2}\mathbf{S}_{R\perp} \cdot \mathbf{E}$, so that the helicity is almost unaffected by an electric field. The anomalous magnetic moment of the electron $a \equiv (g - 2)/2$, however, changes this to a finite value $\approx -(e/mc)a\mathbf{S}_{R\perp} \cdot \mathbf{E}$. This drastic effect was used to measure the experimental values of $a$ for electrons, positrons, and muons:

$$a(e^-) = (115 965.77 \pm 0.35) \times 10^{-8}, \quad (1.325)$$
$$a(e^+) = (116 030 \pm 120) \times 10^{-8}, \quad (1.326)$$
$$a(\mu^\pm) = (116 616 \pm 31) \times 10^{-8}. \quad (1.327)$$

### 1.15 Energy-Momentum Tensor of Perfect Fluid

A perfect fluid is defined as an idealized uniform material medium moving with velocity $\mathbf{v}(x,t)$. The uniformity is an acceptable approximation as long as the microscopic mean free paths are short with respect to the length scale recognizable by the observer. Consider such a fluid at rest. Then the energy-momentum tensor has no momentum density so that

$$T^{0i}_{\text{fluid}} = 0. \quad (1.328)$$
The energy density is given by
\[ c^2 \rho. \quad (1.329) \]
where \( \rho \) is the mass density.

Due to the isotropy, the purely spatial part of the energy-momentum tensor must be diagonal:
\[ T_{ij} = \frac{p}{c} \delta_{ij}, \quad (1.330) \]
where \( p \) is the pressure of the fluid. We can now calculate the energy-momentum tensor of a moving perfect fluid by performing a Lorentz transformation on the energy-momentum tensor at rest:
\[ T_{cd} \quad (1.331) \]
Applying to this the Lorentz boosts from rest to momentum \( p \) of Eq. (1.34), and expressing the hyperbolic functions in terms of energy and momentum according to Eq. (1.148), we obtain
\[ \frac{1}{c} \left[ \left( \frac{p}{c^2} + \rho \right) u^a u^b - pg^{ab} \right], \quad (1.332) \]
where \( u^a \) is the four-velocity (1.145) of the fluid with \( u^a u_a = c^2 \).

**Appendix 1A Tensor Identities**

In the tensor calculus of Euclidean as well as Minkowski space in \( D \) spacetime dimensions, a special role is played by the contravariant Levi-Civita tensor
\[ \epsilon^{a_1 a_2 \ldots a_D}, \quad a_i = 0, 1, \ldots, D - 1. \quad (1A.1) \]
It is the totally antisymmetric unit tensor with the normalization
\[ \epsilon^{012\ldots(D-1)} = 1. \quad (1A.2) \]
It vanishes if any two indices coincide, and is equal to \( \pm 1 \) if they differ from the natural ordering \( 0, 1, \ldots, (D - 1) \) by an even or odd perturbation. The Levi-Civita tensor serves to calculate a determinant of a tensor \( t_{ab} \) as follows
\[ \det(t_{ab}) = \frac{1}{D!} \epsilon^{a_1 a_2 \ldots a_D} \epsilon^{b_1 b_2 \ldots b_D} t_{a_1 b_1} \cdots t_{a_D b_D}. \quad (1A.3) \]
In order to see this it is useful to introduce also the covariant version of \( \epsilon^{a_1 \ldots a_D} \) defined by
\[ \epsilon_{a_1 a_2 \ldots a_D} \equiv g_{a_1 b_1} g_{a_2 b_2} \ldots g_{a_D b_D} \epsilon^{b_1 b_2 \ldots b_D}. \quad (1A.4) \]
It is again a totally antisymmetric unit tensor with
\[ \epsilon_{012\ldots(D-1)} = -1. \] (1A.5)

The contraction of the two is easily seen to be
\[ \epsilon_{a_1\ldots a_D} \epsilon^{a_1\ldots a_D} = -D!. \] (1A.6)

Now, by definition, a determinant is a totally antisymmetric sum
\[ \det(t_{ab}) = \epsilon^{a_1\ldots a_D} t_{a_10} \cdots t_{a_D(D-1)}. \] (1A.7)

We may also write
\[ \det(t_{ab}) \epsilon_{b_1\ldots b_D} = -\epsilon^{a_1\ldots a_D} t_{a_1 b_1} \cdots t_{a_D b_D}. \] (1A.8)

By contracting with \( \epsilon^{b_1\ldots b_D} \) and using (1A.6) we find
\[ \det(t_{ab}) = -\frac{1}{D!} \epsilon^{a_1\ldots a_D} \epsilon^{b_1\ldots b_D} t_{a_1 b_1} \cdots t_{a_D b_D}, \] (1A.9)

which agrees with (1A.7).

In the same way we can derive the formula
\[ \det(t^{a\ b}) = -\frac{1}{D!} \epsilon^{a_1\ldots a_D} \epsilon_{b_1\ldots b_D} t^{a_1 b_1} \cdots t^{a_D b_D}. \] (1A.10)

Under mirror reflection, the Levi-Civita tensor behaves like a pseudotensor. Indeed, if we subject it to a Lorentz transformation \( \Lambda^a_b \), we obtain
\[ \epsilon'_{a_1\ldots a_D} = \Lambda^a_{b_1} \cdots \Lambda^a_{b_D} \epsilon^{b_1\ldots b_D} = \det(\Lambda) \epsilon^{a_1\ldots a_D}. \] (1A.11)

As long as \( \det \Lambda = 1 \), the tensor \( \epsilon^{a_1\ldots a_D} \) is covariant under Lorentz transformations. If space or time inversion are included, then \( \det \Lambda = -1 \), and (1A.11) exhibits the pseudotensor nature of \( \epsilon^{a_1\ldots a_D} \).

We now collect a set of useful identities of the Levi-Civita tensor which will be needed in this text.

### 1A.1 Product Formulas

a) \( D = 2 \) Euclidean space with \( g_{ij} = \delta_{ij} \).

The antisymmetric Levi-Civita tensor \( \epsilon_{ij} \) with the normalization \( \epsilon_{12} = 1 \) satisfies the identities

\[ \epsilon_{ij} \epsilon_{kl} = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}, \] (1A.12)
\[ \epsilon_{ij} \epsilon_{ik} = \delta_{jk}, \] (1A.13)
\[ \epsilon_{ij} \epsilon_{ij} = 2, \] (1A.14)
\[ \epsilon_{ij} \delta_{kl} = \epsilon_{ik} \delta_{jl} + \epsilon_{kj} \delta_{il}. \] (1A.15)
b) $D = 3$ Euclidean space with $g_{ij} = \delta_{ij}$.

The antisymmetric Levi-Civita tensor $\epsilon_{ijk}$ with the normalization $\epsilon_{123} = 1$ satisfies the identities

\begin{align*}
\epsilon_{ijk}\epsilon_{lmn} &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jm}\delta_{kl}, \\
\epsilon_{ijk}\epsilon_{lmn} &= \delta_{jm}\delta_{kn} - \delta_{in}\delta_{km}, \\
\epsilon_{ijk}\epsilon_{ijn} &= 2\delta_{kn}, \\
\epsilon_{ijk}\epsilon_{ijk} &= 6, \\
\epsilon_{ijk}\delta_{lm} &= \epsilon_{ijl}\delta_{km} + \epsilon_{ikl}\delta_{jm} + \epsilon_{jkl}\delta_{im},
\end{align*}

(1A.16) (1A.17) (1A.18) (1A.19) (1A.20)

c) $D = 4$ Minkowski Space with metric

\[ g_{ab} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} . \]

(1A.21)

The antisymmetric Levi-Civita tensor with the normalization $\epsilon^{0123} = -\epsilon_{0123} = 1$ satisfies the product identities

\begin{align*}
\epsilon_{abcd}\epsilon^{efgh} &= -\left( \delta^e_d \delta^f_c \delta^g_b \delta^h_a + \delta^f_d \delta^g_b \delta^h_c \delta^e_a + \delta^g_d \delta^h_b \delta^e_c \delta^f_a + \delta^h_d \delta^e_b \delta^f_c \delta^g_a \\
&+ \delta^e_c \delta^f_d \delta^g_a \delta^h_b + \delta^f_c \delta^g_a \delta^h_d \delta^e_b + \delta^g_c \delta^h_a \delta^e_d \delta^f_b + \delta^h_c \delta^e_a \delta^f_d \delta^g_b \\
&+ \delta^e_d \delta^f_c \delta^g_b \delta^h_a - \delta^f_d \delta^g_c \delta^h_b \delta^e_a + \delta^g_d \delta^h_c \delta^e_b \delta^f_a - \delta^h_d \delta^e_c \delta^f_b \delta^g_a \\
&- \delta^e_d \delta^g_c \delta^h_b \delta^f_a + \delta^f_d \delta^g_c \delta^h_a \delta^e_b - \delta^g_d \delta^h_c \delta^f_a \delta^e_b + \delta^h_d \delta^f_c \delta^g_a \delta^e_b \\
&- \delta^e_d \delta^f_c \delta^g_b \delta^h_a - \delta^f_d \delta^g_c \delta^h_a \delta^e_b + \delta^g_d \delta^h_c \delta^e_b \delta^f_a - \delta^h_d \delta^e_c \delta^f_b \delta^g_a \\
&+ \delta^e_d \delta^f_b \delta^g_c \delta^h_a - \delta^f_d \delta^g_b \delta^h_c \delta^e_a + \delta^g_d \delta^h_b \delta^e_c \delta^f_a - \delta^h_d \delta^e_b \delta^f_c \delta^g_a \right), \quad (1A.22)
\end{align*}

\[ \epsilon_{abcd}\epsilon^{efgh} = -\left( \delta^f_b \delta^g_c \delta^d_e \delta^h_a + \delta^g_b \delta^d_e \delta^h_c \delta^f_a + \delta^d_b \delta^h_c \delta^e_f \delta^g_a - \delta^h_b \delta^e_c \delta^f_d \delta^g_a \right), \quad (1A.23) \]

\[ \epsilon_{abcd}\epsilon^{abgh} = -2\left( \delta^g_b \delta^h_c \delta^e_d - \delta^h_b \delta^g_c \delta^e_d \right), \quad (1A.24) \]

\[ \epsilon_{abcd}\epsilon^{abcd} = -6\delta^e_b , \quad (1A.25) \]

\[ \epsilon_{abcd}\epsilon^{abcd} = -24 , \quad (1A.26) \]

\[ \epsilon_{abcd}g_{ef} = \epsilon_{abce}g_{df} + \epsilon_{abce}g_{df} + \epsilon_{acde}g_{bf} + \epsilon_{acde}g_{bf} . \quad (1A.27) \]

### 1A.2 Determinants

a) $D = 2$ Euclidean:

\[ g = \det(g_{ij}) = \frac{1}{2!}\epsilon_{ik}^{}\epsilon_{lj}^{}g_{kl} = \frac{1}{2}g_{ij}^{}C_{ij}^{} , \quad (1A.28) \]

\[ C_{ij}^{} = \epsilon_{ik}^{}\epsilon_{lj}^{}g_{kl} = \text{cofactor}, \]

\[ g_{ij}^{} = \frac{1}{g}C_{ij}^{} = \text{inverse of } g_{ij} , \]
b) $D = 3$ Euclidean:

$$g = \det(g_{ij}) = \frac{1}{3!} \epsilon_{ijkl} g_{ij} g_{km} g_{ln} = g_{ij} C^{ij}.$$  \hspace{1cm} (1A.29)

$$C^{ij} = \frac{1}{2!} \epsilon_{ijkl} g_{km} g_{ln} = \text{cofactor},$$

$$g^{ij} = \frac{1}{g} C^{ij} = \text{inverse of } g_{ij}.$$  

c) $D = 4$ Minkowski:

$$g = \det(g_{ab}) = -\frac{1}{4!} \epsilon^{abcd} \epsilon^{efgh} g_{ac} g_{bf} g_{eg} g_{dh} = \frac{1}{4} g^{ae} C_{ae},$$  \hspace{1cm} (1A.30)

$$C_{ae} = -\frac{1}{3!} \epsilon^{abcd} \epsilon^{efgh} g_{bf} g_{eg} g_{dh} = \text{cofactor},$$

$$g^{ab} = \frac{1}{g} C_{ab} = \text{inverse of } g_{ab}.$$  

1A.3 Expansion of Determinants

From Formulas (1A.28)–(1A.30) together with (1A.12), (1A.16), (1A.22), we find

$$D=2: \text{det}(g_{ij}) = \frac{1}{2!} [(trg)^2 - tr(g^2)],$$

$$D=3: \text{det}(g_{ij}) = \frac{1}{3!} [(trg)^3 + 2 tr(g^3) - 3 tr(g) tr(g^2)],$$  \hspace{1cm} (1A.31)

$$D=4: \text{det}(g_{ab}) = \frac{1}{24} [(trg)^4 - 6 (trg)^2 tr(g^2) + 3 [tr(g^2)]^2 + 8 tr(g) tr(g^3) - 6 tr(g^4)].$$  

Notes and References


2

Action Approach

The most efficient way of describing the physical properties of a system is based on its action $\mathcal{A}$. The extrema of $\mathcal{A}$ yield the equations of motion, and the sum over all paths of the phase factors $e^{i\mathcal{A}/\hbar}$ renders the quantum-mechanical time evolution amplitude $[1, 2]$. The sum over all paths is called path integral. Historically, the action approach was introduced in classical mechanics to economize Newton’s procedure of setting up equations of motion, and to make it applicable to a large variety of mechanical problems. In quantum mechanics, the sum over all paths with phase factors $e^{i\mathcal{A}/\hbar}$ replaced and generalized the Schrödinger theory. The path integral runs over all position and momentum variables at each time and specifies what are called quantum fluctuations. Their size is controlled by Planck’s quantum $\hbar$, and they are somewhat similar to thermal fluctuations whose size is controlled by the temperature $T$. In the limit $\hbar \to 0$, the paths with highest amplitude will run along the extrema of the action, thus explaining the emergence of classical mechanics from quantum mechanics.

The pleasant property of the action approach is that it can be generalized directly to field theory. Classical fields were discovered by Maxwell as a useful concept to describe the phenomena of electromagnetism. In particular, his equations allow us to study the propagation of free electromagnetic waves without considering the sources. In the last century, Einstein constructed his theory of gravity by assuming the metric of spacetime to become a spacetime-dependent field, which can propagate in the form of gravitational waves. In condensed matter physics, fields were introduced in many systems, and Landau made them a universal tool for understanding phase transitions [3]. Such fields are called order fields. A more recently discovered domain of applications of fields is in the statistical mechanics of grand-canonical ensembles of line-like objects, such as vortex lines in superfluids and superconductors [4], or defect lines in crystals [5]. In this context, they are known as disorder fields.
2.1 General Particle Dynamics

Given an arbitrary classical system with generalized coordinates $q_n(t)$ and velocities $\dot{q}_n(t)$, the typical action has the form

$$A[q_k] = \int_{t_a}^{t_b} dt \ L(q_k(t), \dot{q}_k(t), t),$$

(2.1)

where $L(q_k(t), \dot{q}_k(t), t)$ is called the Lagrangian of the system, which is quadratic in the velocities $\dot{q}_k(t)$. A Lagrangian with this property is called local in time. If a theory is governed by a local Lagrangian, the action and the entire theory are also called local. The quadratic dependence on $\dot{q}(t)$ may emerge only after an integration by parts in the action. For example, $-\int dt q(t)\ddot{q}(t)$ is a local term in the Lagrangian since it turns into $\int dt \dot{q}_k^2(t)$ after a partial integration in the action (2.1).

The physical trajectories of the system are found from the extremal principle. One compares the action for one orbit $q_k(t)$ connecting the end points

$$q_k(t_a) = q_{k,a}, \quad q_k(t_b) = q_{k,b},$$

(2.2)

with that of an infinitesimally different orbit $q'_k(t) \equiv q_k(t) + \delta q_k(t)$ connecting the same end points, where $\delta q_k(t)$ is called the variation of the orbit. Since the endpoint of $q_k(t) + \delta q_k(t)$ are the same as those of $q_k(t)$, the variations of the orbit vanish at the end points:

$$\delta q(t_a) = 0, \quad \delta q(t_b) = 0.$$  

(2.3)

The associated variation of the action is

$$\delta A \equiv A[q_k + \delta q_k] - A[q_k] = \int_{t_a}^{t_b} dt \sum_k \left( \frac{\partial L}{\partial q_k(t)} \delta q_k(t) + \frac{\partial L}{\partial \dot{q}_k(t)} \delta \dot{q}_k(t) \right).$$

(2.4)

After an integration by parts, this becomes

$$\delta A = \int_{t_a}^{t_b} dt \sum_k \left( \frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k(t) + \sum_k \left. \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k(t) \right|_{t_a}^{t_b}. \tag{2.5}$$

In going from (2.4) to (2.5) one uses the fact that by definition of $\delta q_k(t)$ the variation of the time derivative is equal to the time derivative of the variation:

$$\delta \dot{q}_k(t) = \dot{\delta q}_k(t) - \dot{q}_k(t) = \frac{d}{dt} [q_k(t) + \delta q_k(t)] - q_k(t) = \frac{d}{dt} \delta q_k(t). \tag{2.6}$$

Expressed more formally, the time derivative commutes with the variations of the orbit:

$$\delta \frac{d}{dt} q_k(t) \equiv \frac{d}{dt} \delta q_k(t). \tag{2.7}$$
Using the property (2.3), the boundary term on the right-hand side of (2.5) vanishes. Since the action is extremal for a classical orbit, $\delta A$ must vanish for all variations $\delta q_k(t)$, implying that $q_k(t)$ satisfies the Euler-Lagrange equation

$$\frac{\partial L}{\partial q_k(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k(t)} = 0,$$

which is the equation of motion of the system. It is a second-order differential equation for the orbit $q_k(t)$.

The local Lagrangian of a set of gravitating mass points is

$$L(x(t), \dot{x}(t)) = \sum_k m_k \dot{x}_k^2(t) + G_N \sum_{k \neq k'} m_k m_{k'} \left| x_k(t) - x_{k'}(t) \right|.$$  

If we identify the $3N$ coordinates $x_n^i$ ($n = 1, \ldots, N$) with $3N$ generalized coordinates $q_k$ ($k = 1, \ldots, 3N$), the Euler-Lagrange equations (2.8) reduce precisely to Newton’s equations (1.2).

The energy of a general Lagrangian system is found from the Lagrangian by forming the so-called Hamiltonian. It is defined by the Legendre transform

$$H = \sum_k p_k \dot{q}_k - L,$$  

where

$$p_k \equiv \frac{\partial L}{\partial \dot{q}_k},$$

are called canonical momenta. The energy (2.10) forms the basis of the Hamiltonian formalism. If expressed in terms of $p_k, q_k$, the equations of motion are

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}.$$  

For the Lagrangian (2.9), the generalized momenta are equal to the physical momenta $p_n = m_n \dot{x}_n$, and the Hamiltonian is given by Newton’s expression

$$H = T + V \equiv \sum_k \frac{m_k}{2} \dot{x}_k^2 - G_N \sum_{k \neq k'} \frac{m_k m_{k'}}{|x_k(t) - x_{k'}(t)|}.$$  

The first term is the kinetic energy, the second the potential energy of the system.

### 2.2 Single Relativistic Particle

For a single relativistic massive point particles, the mechanical action reads

$$\mathcal{A} = \int_{t_a}^{t_b} dt \int \frac{m}{L} = -mc^2 \int_{t_a}^{t_b} dt \sqrt{1 - \frac{\dot{x}^2(t)}{c^2}}.$$
The canonical momenta (2.11) yield directly the spatial momenta of the particle:

\[ p(t) = \frac{\partial L}{\partial \dot{x}(t)}. \tag{2.15} \]

In a relativistic notation, the derivative with respect to the contravariant vectors \( \frac{\partial m}{\partial \dot{x}^i} \) is a covariant vector with a lower index \( i \). To ensure the nonrelativistic identification (2.15) and maintain the relativistic notation we must therefore identify

\[ p_i \equiv -\frac{\partial m}{\partial \dot{x}^i} = m\gamma \dot{x}_i. \tag{2.16} \]

The energy is obtained from the Legendre transform:

\[ mH = p\dot{x} - L = -p_i \dot{x}^i - L = m\gamma v^2 + mc^2 \sqrt{1 - \frac{v^2}{c^2}} = m\gamma v^2 + mc^2 \frac{1}{\gamma}, \tag{2.17} \]

in agreement with the energy in Eq. (1.152) [recalling (1.147)].

The action (2.14) can be written in a more covariant form by observing that

\[ \int_{t_a}^{t_b} dt \sqrt{1 - \left(\frac{dx}{dt}\right)^2} = \frac{1}{c} \int_{t_a}^{t_b} d\sigma \sqrt{\left(\frac{dx^0}{d\sigma}\right)^2 - \left(\frac{dx}{d\sigma}\right)^2}. \tag{2.18} \]

In this expression, the infinitesimal time element \( dt \) can be replaced by an arbitrary time-like parameter \( t \to \sigma = f(t) \), so that the action takes the more general form

\[ A_\mu = \int_{\sigma_a}^{\sigma_b} d\sigma \frac{m}{\sqrt{g_{ab} \dot{x}^a(\sigma) \dot{x}^b(\sigma)}}. \tag{2.19} \]

For this action we may define generalized four-momentum with respect to the parameter \( \tau \) by forming the derivatives

\[ p_{\sigma,a} \equiv -\frac{\partial m}{\partial \dot{x}^a(\sigma)} = \frac{mc}{\sqrt{g_{ab} \dot{x}^a(\sigma) \dot{x}^b(\sigma)}} g_{ab} \dot{x}^b(\sigma), \tag{2.20} \]

where the dots denote the derivatives with respect to the argument. Note the minus sign in the definition of the canonical momentum with respect to the nonrelativistic case. This is introduced to make the canonical formalism compatible with the negative signs in the spatial part of the Minkowski metric (1.29). The derivatives with respect to \( \dot{x}^a \) transforms like the covariant components of a vector with a subscripts \( a \), whereas the physical momenta are given by the contravariant components \( p^a \).

If \( \sigma \) is chosen to be the proper time \( \tau \), then the square root in (2.86) becomes \( \tau \)-independent

\[ \sqrt{g_{ab} \dot{x}^a(\tau) \dot{x}^b(\tau)} = c, \tag{2.21} \]
so that
\[ p_{\tau,a} = mg_{ab} \dot{x}^b(\tau) = mu_a(\tau), \quad (2.22) \]
in agreement with the previously defined four-momenta in Eq. (1.147).

In terms of the proper time, the Euler-Lagrange equation reads:
\[ \frac{d}{d\tau} p_{\tau,a} = m \frac{d}{d\tau} g_{ab} \dot{x}^b(\tau) = m \ddot{x}_a(\tau) = 0. \quad (2.23) \]
Free particles run along straight lines in Minkowski space.

Note that the Legendre transform with respect to the momentum \( p_{\sigma,0} \) has nothing to do with the physical energy. In fact, it vanishes identically:
\[ m \tilde{H}_\sigma = -p_{\sigma a} \dot{x}^a(\sigma) - \tilde{L} = -mc \sqrt{\dot{x}_a(\sigma) \dot{x}^a(\sigma)} + mc \sqrt{\dot{x}_a(\sigma) \dot{x}^a(\sigma)} \equiv 0. \quad (2.24) \]
The reason for this is the invariance of the action (2.19) under arbitrary reparametrizations of the time \( \sigma \rightarrow \sigma' = f(\sigma) \). We shall understand this better in Chapter 3 when discussing generators of continuous symmetry transformations in general. See in particular Subsection 3.5.3.

The role of the physical energy is played by \( p_{\tau,0} = mc\gamma \). It is equal to \( 1/c \) times the energy \( H \) in (2.17), as it should.

## 2.3 Scalar Fields

The free classical point particles of the last section are quanta of a relativistic local scalar free-field theory.

### 2.3.1 Locality

Generalizing the concept of temporal locality described after Eq. (2.1), locality in field theory implies that the action is a spacetime integral over the Lagrangian density:
\[ \mathcal{A} = \int_{t_a}^{t_b} dt \int d^4x \mathcal{L}(x) = \frac{1}{c} \int d^4x \mathcal{L}(x) = \frac{1}{c} \int d^4x \mathcal{L}(\phi(x)) \partial_\mu \phi(x)). \quad (2.25) \]
According to the concept of temporal locality in Section 2.1, there should only be a quadratic dependence on the time derivatives of the fields \( \phi(x), \phi^*(x) \). Due to the equal footing of space and time in relativistic theories, the same restriction applies to the space derivatives. A local Lagrangian density can at most be quadratic in the first spacetime derivatives of the fields at the same point. Physically this implies that a field at a point \( x \) interacts at most with the field at the infinitesimally close neighbor point \( x + dx \), just like the mass points on a linear chain with nearest-neighbor spring interactions. If the derivative terms are not of this form, they must
at least be equivalent to it by a partial integration in the action integral (2.25). If the Lagrangian density is local we call also the action and the entire theory local.

A local free-field Lagrangian density is quadratic in both the fields and their derivatives at the same point, so that it reads, for a scalar field,

\[ \mathcal{L}(x) = \frac{1}{2} \left[ \hbar^2 \partial_a \phi(x) \partial^a \phi(x) - m^2 c^2 \phi(x) \phi(x) \right]. \] (2.26)

If the particles are charged, the fields are complex, and the Lagrangian density becomes

\[ \mathcal{L}(x) = \bar{\phi} \partial_a \phi^* \partial^a \phi - m^2 c^2 \phi^* \phi. \] (2.27)

### 2.3.2 Lorenz Invariance

In addition to being local, any relativistic Lagrangian density \( \mathcal{L}(x) \) must be a scalar, i.e., transforms under Lorentz transformations in the same way as the scalar field \( \phi(x) \) in Eq. (1.163):

\[ \mathcal{L}(x) \xrightarrow{\Lambda} \mathcal{L}'(x) = \mathcal{L}(\Lambda^{-1} x). \] (2.28)

We verify this by showing that \( \mathcal{L}'(x') = \mathcal{L}(x) \). By definition, \( \mathcal{L}'(x') \) is equal to

\[ \mathcal{L}'(x') = \hbar^2 \partial'_a \phi^*(x') \partial'^a \phi(x') - m^2 c^2 \phi^* \phi(x'). \] (2.29)

Using the transformation behavior of the scalar field \( \phi(x) \) in Eq. (1.163), we obtain

\[ \mathcal{L}'(x') = \hbar^2 \partial'_a \phi^*(x') \partial'^a \phi(x') - m^2 c^2 \phi^* \phi(x). \] (2.30)

Inserting here

\[ \partial'_a = \Lambda_a^b \partial_b, \quad \partial'^a = \Lambda^a_b \partial^b \] (2.31)

with

\[ \Lambda_a^b \equiv g_{ac} g^{bd} \Lambda_c^d, \] (2.32)

we see that \( \partial'^2 \) is Lorentz-invariant,

\[ \partial'^2 = \partial^2, \] (2.33)

so that the transformed Lagrangian density (2.29) coincides indeed with the original one in (2.27).

As a spacetime integral over a scalar Lagrangian density, the action (2.25) is Lorentz invariant. This follows directly from the Lorentz invariance of the spacetime volume element

\[ dx^0 dx^1 dx^2 dx^3 = d^4x = d^4x', \] (2.34)

proved in Eq. (1.256). This is verified by direct calculation:

\[ \mathcal{A}' = \int d^4x \mathcal{L}'(x) = \int d^4x' \mathcal{L}'(x') = \int d^4x' \mathcal{L}(x) = \int d^4x \mathcal{L}(x) = \mathcal{A}. \] (2.35)
2.3.3 Field Equations

The equation of motion for the scalar field is derived by varying the action (2.25) with respect to the fields $\phi(x), \phi^*(x)$ independently. The independence of the field variables is expressed by the functional differentiation rules

$$\frac{\delta \phi(x)}{\delta \phi(x')} = \delta^{(4)}(x - x'), \quad \frac{\delta \phi^*(x)}{\delta \phi^*(x')} = \delta^{(4)}(x - x'),$$

$$\frac{\delta \phi(x)}{\delta \phi^*(x')} = 0, \quad \frac{\delta \phi^*(x)}{\delta \phi(x')} = 0.$$  \hspace{1cm} (2.36)

With the help of these rules and the Leibnitz chain rule (1.114), we calculate the functional derivative of the action (2.25) as follows:

$$\frac{\delta A}{\delta \phi^*(x)} = \int d^4 x' \left[ \bar{h}^2 \partial^2 \phi^*(x') \partial^a \phi^* - m^2 c^2 \phi^*(x) \right] = 0.$$ \hspace{1cm} (2.37)

Similarly

$$\frac{\delta A}{\delta \phi(x)} = \int d^4 x' \left[ \bar{h}^2 \partial^2 \phi^* \partial^a \partial^a \phi^* - m^2 c^2 \phi^* \delta^{(4)}(x' - x) \right] = \phi^*(x) (-\bar{h}^2 \partial^2 + m^2 c^2) = 0,$$ \hspace{1cm} (2.38)

where the arrow pointing to the left on top of the last derivative indicates that it acts on the field to the left. The second equation is just the complex conjugate of the previous one.

The field equations can also be derived directly from the Lagrangian density (2.27) by forming ordinary partial derivatives of $\mathcal{L}$ with respect to all fields and their derivatives. Indeed, a functional derivative of a local action can be expanded in terms of derivatives of the Lagrangian density according to the general rule

$$\frac{\delta A}{\delta \phi(x)} = \frac{\partial \mathcal{L}(x)}{\partial \phi(x)} - \partial_a \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \phi(x)]} + \partial_a \partial_b \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \partial_b \phi(x)]} + \ldots,$$ \hspace{1cm} (2.39)

and a complex-conjugate expansion for $\phi^*(x)$. These expansions follow directly from the defining relations (2.36). At the extremum of the action, the field satisfies the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}(x)}{\partial \phi(x)} - \partial_a \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \phi(x)]} + \partial_a \partial_b \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \partial_b \phi(x)]} + \ldots = 0.$$ \hspace{1cm} (2.40)

Inserting the Lagrangian density (2.27), we obtain the field equation for $\phi(x)$

$$\frac{\delta A}{\delta \phi^*(x)} = \frac{\partial \mathcal{L}(x)}{\partial \phi^*(x)} - \partial_a \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \phi^*(x)]} = (-\bar{h}^2 \partial^2 + m^2 c^2) \phi(x) = 0,$$ \hspace{1cm} (2.41)
and its complex conjugate for $\phi^*(x)$.

The Euler-Lagrange equations are invariant under partial integrations in the action integral (2.25). Take for example a Lagrangian density which is equivalent to (2.27) by a partial integration:

$$\mathcal{L} = -\hbar^2 \phi^*(x) \partial^2 \phi(x) - m^2 c^2 \phi^*(x) \phi(x). \quad (2.42)$$

Inserted into (2.40), the field equation for $\phi(x)$ becomes particularly simple:

$$\delta A \frac{\delta A}{\delta \phi(x)} = \frac{\partial \mathcal{L}(x)}{\partial \phi^*(x)} = (-\hbar^2 \partial^2 + m^2 c^2) \phi(x) = 0. \quad (2.43)$$

The derivation of the equation for $\phi^*(x)$, on the other hand, becomes more complicated. All derivatives written out in (2.39) have to be evaluated. At the end we simply find the complex-conjugate of Eq. (2.43):

$$\delta A \frac{\delta A}{\delta \phi(x)} = \frac{\partial \mathcal{L}(x)}{\partial \phi^*(x)} - \partial_a \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \phi(x)]} + \partial_a \partial_b \frac{\partial \mathcal{L}(x)}{\partial [\partial_a \partial_b \phi(x)]} = (-\hbar^2 \partial^2 + m^2 c^2) \phi^*(x) = 0. \quad (2.44)$$

### 2.3.4 Plane Waves

The field equations (2.43) and (2.44) are solved by the quantum mechanical plane waves

$$f_p(x) = N e^{-ipx/\hbar}, \quad f_p^*(x) = N e^{ipx/\hbar}, \quad (2.45)$$

where the four-momenta satisfy the so-called mass shell condition

$$p^0 p_a - m^2 c^2 = 0, \quad (2.46)$$

and $N$ is some normalization factor. It is important to realize that the two sets of solutions (2.45) are independent of each other. They differ by the sign of energy

$$i\partial_0 f_p(x) = p^0 f_p(x), \quad i\partial_0 f_p^*(x) = -p^0 f_p^*(x). \quad (2.47)$$

For this reason they will be referred to as positive- and negative-frequency wave functions, respectively. The physical significance of the latter can only be understood after quantizing the field, where they turn out to be associated with antiparticles. Field quantization, however, lies outside the scope of this text. Only at the end, in Subsection 27.2, will its effects on gravity be discussed.

### 2.3.5 Schrödinger Quantum Mechanics as Nonrelativistic Limit

The nonrelativistic limit of the action (2.25) is obtained by removing from the positive frequency part of the field $\phi(x)$ a rapidly oscillating factor corresponding to the rest energy $mc^2$, replacing

$$\phi(x) \rightarrow e^{-imc^2 t/\hbar} \frac{1}{\sqrt{2M}} \psi(x, t). \quad (2.48)$$
For a plane wave $f_p(x)$ in (2.45), the field $\psi(x)$ becomes $N\sqrt{2M}e^{-ip^0c-mc^2t/\hbar}e^{ipx/\hbar}$. In the limit of large $c$, the first exponential becomes $e^{-ip^2t/2M}$ [recall (1.154)]. The result is a plane-wave solution to the Schrödinger equation

$$\left[ i\hbar \partial_t + \frac{\hbar^2}{2M} \partial_x^2 \right] \psi(x, t) = 0.$$  

This is the Euler-Lagrange equation extremizing the nonrelativistic action

$$\mathcal{A} = \int dt d^3x \psi^* (x, t) \left[ i\hbar \partial_t + \frac{\hbar^2}{2M} \partial_x^2 \right] \psi(x, t).$$  

Note that the plane wave $f_p^*(x)$ in (2.45) with negative frequency does not possess a nonrelativistic limit since it turns into $N\sqrt{2M}e^{ip^0c+mc^2t/\hbar}e^{ipx/\hbar}$ which has a temporal prefactor $e^{2imc^2t/\hbar}$. This oscillates infinitely rapidly for $c \to \infty$, and is therefore equivalent to zero by the Riemann-Lebesgue Lemma.\footnote{This statement holds in the sense of distributions. Any integral over an infinitely rapidly oscillating function multiplied by a smooth function yields zero.}

### 2.3.6 Natural Units

The appearance of the constants $\hbar$ and $c$ in all future formulas can be avoided by working with fundamental units $l_0, m_0, t_0, E_0$ different from the ordinary cgs units. They are chosen to give $\hbar$ and $c$ the value 1. Expressed in terms of the conventional length, time, mass, and energy, these natural units are

$$l_0 = \frac{\hbar}{mc}, \quad t_0 = \frac{\hbar}{mc^2}, \quad m_0 = M, \quad E_0 = mc^2.$$  

If, for example, the particle is a proton with mass $m_p$, these units are

$$l_0 = 2.103138 \times 10^{-11}\text{cm} \quad (2.52)$$

Compton wavelength of proton,

$$t_0 = l_0/c = 7.0153141 \times 10^{-22}\text{sec} \quad (2.53)$$

time it takes light to cross the Compton wavelength,

$$m_0 = m_p = 1.6726141 \times 10^{-24}\text{g} \quad (2.54)$$

$$E_0 = 938.2592\text{MeV}.$$  

For any other mass, they can easily be rescaled.

With these natural units we can drop $c$ and $\hbar$ in all formulas and write the action simply as

$$\mathcal{A} = \int d^4x \phi^* (x) (-\partial^2 - m^2)\phi(x).$$  

\text{H. Kleinert, GRAVITY WITH TORSION}
Actually, since we are dealing with relativistic particles, there is no fundamental reason to assume $\phi(x)$ to be a complex field. In nonrelativistic field theory, this was necessary in order to construct a term linear in the time derivative

$$\int dt d^3x \psi^*(x,t) i\hbar \partial_t \psi(x,t)$$

in the action (2.50). For a real field, this would have been a pure surface term and thus not influenced the dynamics of the system. The second-order time derivatives of a relativistic field in (2.56), however, lead to the correct field equation for a real field. As we shall understand better in the next chapter, the complex scalar field describes charged spinless particles, the real field neutral particles.

Thus we may also consider a real scalar field with an action

$$A = \frac{1}{2} \int d^4x \phi(x)(-\partial^2 - m^2)\phi(x).$$

In this case it is customary to use a prefactor $1/2$ to normalize the field.

For either Lagrangian (2.56) or (2.58), the Euler-Lagrange equation (2.40) becomes the *Klein-Gordon equation*

$$(-\partial^2 - m^2)\phi(x) = 0.$$  

### 2.3.7 Hamiltonian Formalism

It is possible to set up a Hamiltonian formalism for the scalar fields. For this we introduce an appropriate generalization of the canonical momentum (2.11). The labels $k$ in that equation are now replaced by the continuous spatial labels $x$, so that we define a density of field momentum:

$$\pi(x) \equiv \frac{\partial L}{\partial \dot{\phi}(x)} = \partial_0 \phi(x), \quad \pi^*(x) \equiv \frac{\partial L}{\partial \dot{\phi}^*(x)} = \partial_0 \phi(x),$$

and a *Hamiltonian density*

$$\mathcal{H}(x) = \pi(x) \partial_0 \phi(x) + \partial_0 \phi(x) \pi^*(x) - L(x) = \pi^*(x) \partial_0 \phi(x) + \nabla \phi^*(x) \nabla \phi(x) + m^2 \phi^*(x) \phi(x).$$

For a real field, we simply drop the complex conjugation symbols. The spatial integral over $\mathcal{H}(x)$ the field Hamiltonian

$$H = \int d^3x \mathcal{H}(x).$$

### 2.3.8 Conserved Current

For a complex field, there exists an important local conservation law. We define the four-vector of probability current density

$$j_a(x) = \frac{i}{2} \phi^* \vec{\partial}_a \phi,$$
which described the probability flow of the charged scalar particle. The double arrow on top of the derivative is defined by $\overrightarrow{\partial} \equiv \overrightarrow{\partial} - \overleftarrow{\partial}$ as in the nonrelativistic Eq. (2.70). It is easy to verify that on behalf of the Klein-Gordon equation (2.59), the current density has no four-divergence:

$$\partial_a j^a(x) = 0. \quad (2.64)$$

This current conservation law permits us to couple electromagnetism to the field and identify $ej^a(x)$ as the electromagnetic current of the charged scalar particles.

The deeper reason for the existence of a conserved current will be understood in Chapter 3, where we shall see that it is intimately connected with an invariance of the action (2.56) under arbitrary changes of the phase of the field

$$\phi(x) \rightarrow e^{-i\alpha} \phi(x). \quad (2.65)$$

The zeroth component of $j^a(x)$,

$$\rho(x) = j^0(x), \quad (2.66)$$

describes the charge density. The spatial integral over $\rho(x)$ is the total probability. It measures the total charge in natural units:

$$Q(t) = \int d^3x \, j^0(x). \quad (2.67)$$

Because of the local conservation law (2.64), the total charge does not depend on time. This is seen by rewriting

$$\dot{Q}(t) = \int d^3x \, \partial_0 j^0(x) = \int d^3x \, \partial_a j^a(x) - \int d^3x \, \partial_j j^j(x) = - \int d^3x \, \partial_0 j^0(x). \quad (2.68)$$

The right-hand side vanishes due to the Gauss divergence theorem, assuming the currents to vanish at spatial infinity [compare (1.198)].

For the solutions $\psi(x, t)$ of the Schrödinger equation (2.49), the probability density is

$$\rho(x, t) \equiv \psi^*(x, t) \psi(x, t) \quad (2.69)$$

and the particle current density

$$j(x, t) \equiv \frac{i}{2m} \psi^*(x, t) (\overrightarrow{\nabla} - \overleftarrow{\nabla}) \psi(x, t) \equiv \frac{i}{2m} \psi^*(x, t) \overrightarrow{\nabla} \psi(x, t). \quad (2.70)$$

They satisfy the conservation law

$$\partial_t \rho(x, t) = - \nabla \cdot j(x, t). \quad (2.71)$$

It is this property which permits normalizing the Schrödinger field $\psi(x, t)$ to unity for all times, since

$$\partial_t \int d^3x \, \psi^*(x, t) \psi(x, t) = \int d^3x \, \partial_t \rho(x, t) = - \int d^3x \, \nabla \cdot j(x, t) = 0. \quad (2.72)$$
2.4 Maxwell’s Equation from Extremum of Field Action

The above action approach is easily generalized to apply to electromagnetic fields. By setting up an appropriate action, Maxwell’s field equations can be derived by extremization. The relevant fields are the Coulomb potential \( A^0(x, t) \) and the vector potential \( A(x, t) \). Recall that electric and magnetic fields \( E(x) \) and \( B(x) \) can be written as derivatives of the Coulomb potential \( A^0(x, t) \) and the vector potential \( A(x, t) \) as

\[
\begin{align*}
E(x) &= -\nabla A^0(x) - \frac{1}{c} \dot{A}(x), \\
B(x) &= \nabla \times A(x),
\end{align*}
\]

with the components

\[
\begin{align*}
E^i(x) &= -\partial_i A^0(x) - \frac{1}{c} \partial_t A^i(x), \\
B^i(x) &= \epsilon^{ijk} \partial_j A^k(x).
\end{align*}
\]

With the identifications (1.167) of electric and magnetic fields with the components \( F^{i0} \) and \(-F^{jk}\) of the covariant field tensor \( F^{ab} \), we can also write

\[
\begin{align*}
F^{i0}(x) &= \partial^i A^0(x) - \frac{1}{c} \partial_t A^i(x), \\
F^{jk}(x) &= \partial^j A^k(x) - \partial^k A^j(x),
\end{align*}
\]

where \( \partial^i = -\partial_i \). This suggests combining the Coulomb potential and the vector potential into a four-component vector potential

\[
A^a(x) = \begin{pmatrix} A^0(x, t) \\ A^i(x, t) \end{pmatrix},
\]

in terms of which the field tensor is simply the four-dimensional curl:

\[
F^{ab}(x) = \partial^a A^b(x) - \partial^b A^a(x).
\]

The field \( A^a(x) \) transforms in the same way as the vector field \( j^a(x) \) in Eq. (1.203):

\[
A^a(x) \xrightarrow{\Lambda} A'^a(x) = \Lambda^a_b A^b(\Lambda^{-1} x).
\]

2.4.1 Electromagnetic Field Action

Maxwell’s equations can be derived from the electromagnetic field action

\[
\mathcal{A}_\text{em} = \frac{1}{c} \int d^4x \mathcal{L}_\text{em}(x),
\]
where the temporal integral runs from $t_a$ to $t_b$, as in (2.1) and (2.25), and the Lagrangian density reads

$$
\mathcal{L}^m(x) \equiv \mathcal{L} \left( A^a(x), \partial^b A^a(x) \right) = -\frac{1}{4} F^{ab}(x) F_{ab}(x) - \frac{1}{c^2} j^a(x) A_a(x).
$$

(2.83)

It depends quadratically on the fields $A^a(x)$ and its derivatives, thus defining a local field theory [recall Subsection 2.3.1]. All Lorentz indices are fully contracted.

If (2.83) is decomposed into electric and magnetic parts using Eqs. (2.73) and (2.74), it reads

$$
\mathcal{L}^m(x) = \frac{1}{2} \left[ E^2(x) - B^2(x) \right] - \rho(x) A^0(x) + \frac{1}{c} j(x) A(x).
$$

(2.84)

From the transformation laws (1.170), (1.203), and (2.81) it follows that (2.83) transforms like a scalar field under Lorentz transformations as in (2.28). Together with (2.34), this implies that the action is Lorentz-invariant.

The field equations are obtained from the Euler-Lagrange equation (2.40) with the field $A^0(x)$ replaced by the four-vector potential $A^a(x)$, so that it reads

$$
\partial_b F_{ab} = -\frac{1}{c} j^a,
$$

(2.86)

Inserting the Lagrangian density (2.83), we obtain

$$
\partial_b \mathcal{L} = \partial_b \left\{ \frac{1}{4} F^{ab}(x) \partial^b \partial_a \mathcal{L} - \frac{1}{c} \partial^b j^a \partial_a \mathcal{L} \right\} = 0.
$$

(2.85)

In this sense, the homogeneous Maxwell equation (2.87) is a Bianchi identity, since it follows directly from the commuting derivatives of $A^a(x)$ in Eq. (2.88).
2.4.2 Alternative Action for Electromagnetic Field

There exists an alternative form of the electromagnetic Lagrangian density (2.83) due to Schwinger which contains directly the field tensor as independent variables and uses the vector potential only as Lagrange multipliers to enforce the inhomogeneous Maxwell equations (2.86):

\[
\mathcal{L}_{\text{em}}(x) = \mathcal{L}(A^a(x), F_{ab}(x)) = -\frac{1}{4} F^{ab}(x) F_{ab}(x) - \frac{1}{c} \left[ j^a(x) + \partial_b F^{ab}(x) \right] A_a(x). \tag{2.89}
\]

Extremizing this with respect to \( F_{ab} \) show that \( F_{ab} \) is a four-curl of the vector potential, as in Eq. (2.80). As a consequence, \( F_{ab} \) satisfies the Bianchi identity (1.195).

If (2.89) is decomposed into electric and magnetic parts, it reads

\[
\mathcal{L}_{\text{em}}(x) = \mathcal{L}(A^0(x), A(x), E(x), B(x)) = \frac{1}{4} \left[ E^2 - B^2 \right] + \left[ \nabla \cdot E(x) - \rho(x) \right] A^0(x) - \left[ \nabla \times B(x) - \frac{1}{c} \partial_t E(x) - \frac{1}{c} j(x) \right] \cdot A(x), \tag{2.90}
\]

where the Lagrange multipliers \( A^0(x) \) and \( A(x) \) enforce directly the Coulomb law (1.180) and the Ampère law (1.181).

The above equations hold only in the vacuum. In homogeneous materials with nonzero dielectric constant \( \varepsilon \) and magnetic permeability \( \mu \) determining the displacement fields \( D = \varepsilon E \) and the magnetic fields \( H = B/\mu \), the Lagrangian density (2.89) reads

\[
\mathcal{L}_{\text{em}}(x) = \frac{1}{4} \left[ E \cdot D - B \cdot H \right] + \left[ \nabla \cdot D(x) - \rho(x) \right] A^0(x) - \left[ \nabla \times H(x) - \frac{1}{c} \partial_t D(x) - \frac{1}{c} j(x) \right] \cdot A(x). \tag{2.91}
\]

Now variation with respect to the Lagrange multipliers \( A^0(x) \) and \( A(x) \) yields the Coulomb and Ampère laws in a medium Eqs. (1.185) and (1.186):

\[
\nabla \cdot D(x) = \rho(x), \quad \nabla \times H(x) - \frac{1}{c} \partial_t D(x) = \frac{1}{c} j(x). \tag{2.92}
\]

Variation with respect to \( D(x) \) and \( H(x) \) yields the same curl equations (2.73) and (2.74) as in the vacuum, so that the homogeneous Maxwell equations (1.182) and (1.183), i.e., the Bianchi identities (1.195), are unaffected by the medium.

2.4.3 Hamiltonian of Electromagnetic Fields

As in Eqs. (2.60)–(2.62), we can find a Hamiltonian for the electromagnetic fields, by defining a density of field momentum:

\[
\pi_a(x) = \frac{\partial \mathcal{L}_{\text{em}}}{\partial \partial_t A^a(x)} = -F_{0a}(x), \tag{2.93}
\]
and a Hamiltonian density
\[ \mathcal{H}^{em}(x) = \pi_a(x) \partial^0 A^a(x) - \mathcal{L}^{em}(x). \] (2.94)

It is important to realize that \[ \partial^0 \mathcal{L}^{em} \partial^0 A^a \]
vanishes, so that \( A^a \) possesses no conjugate field momentum. This implies that it is not a proper dynamical variable. Indeed, by inserting (2.83) and (2.93) into (2.94) we find
\[ \mathcal{H}^{em} = -F_{0a} \partial^0 A^a - \mathcal{L} = -\frac{1}{c} F_{0a} F^{0a} - \mathcal{L} - F_{0a} \partial^0 A^a. \] (2.95)

Integrating this over all space gives
\[ \mathcal{H}^{em} = c \int d^3 x \mathcal{H}^{em} = \int d^3 x \left[ \frac{1}{2} (E^2 + B^2) + E \cdot \nabla A^0 + \frac{1}{c} j^a A_a. \right]. \] (2.96)

The result is the well-known energy of the electromagnetic field in the presence of external currents [6]. To obtain this expression from (2.95), an integration by part is necessary, in which the surface terms at spatial infinity is neglected, where the charge density \( \rho(x) \) is always assumed to be zero. After this, Coulomb’s law (1.180) leads directly to (2.96).

At first sight, one may wonder why the electrostatic energy does not show up explicitly in (2.96). The answer is that it is contained in the \( E^2 \)-term which, by Coulomb’s law (1.180), satisfies
\[ \nabla \cdot E = -\nabla^2 A^0 - \frac{1}{c} \partial_t \nabla \cdot A = \rho. \] (2.97)

Splitting \( E \) into transverse and longitudinal parts
\[ E = E_t + E_l, \] (2.98)
which satisfy \( \nabla \cdot E_t = 0 \) and \( \nabla \times E_t = 0 \), respectively, we see that (2.97) implies
\[ \nabla \cdot E_t = \rho. \] (2.99)

The longitudinal part can be written as a derivative of some scalar potential \( \phi' \),
\[ E_l = \nabla \phi' \] (2.100)
which, due to (2.99), can be calculated from the equation
\[ \phi'(x) = \frac{1}{\nabla^2} \rho(x) = - \int d^3 x' \frac{1}{4\pi |x - x'|} \rho(x', t). \] (2.101)

Using this we see that
\[ \frac{1}{2} \int d^3 x E^2 = \frac{1}{2} \int d^3 x \left( E_t^2 + E_l^2 \right) = \frac{1}{2} \int d^3 x \left[ E_t^2 + \left( \partial_i \frac{1}{\nabla^2} \phi' \right) \left( \partial_i \frac{1}{\nabla^2} \phi' \right) \right] \]
\[ = \frac{1}{2} \int d^3 x E_t^2 + \frac{1}{2} \int d^3 x d^3 x' \rho(x, t) \frac{1}{4\pi |x - x'|} \rho(x', t). \] (2.102)

The last term is the Coulomb energy associated with the charge density \( \rho(x, t) \).
2.4.4 Gauge Invariance of Maxwell’s Theory

The four-dimensional curl (2.80) is manifestly invariant under the gauge transformations

\[ A_a(x) \rightarrow A'_a(x) = A_a(x) + \partial_a \Lambda(x), \]  

(2.103)

where \( \Lambda(x) \) is any smooth field which satisfies the integrability condition

\[ (\partial_a \partial_b - \partial_b \partial_a) \Lambda(x) = 0. \]  

(2.104)

Gauge invariance implies that a scalar field degree of freedom contained in \( A^a(x) \) does not contribute to the physically observable electromagnetic fields \( E(x) \) and \( B(x) \). This degree of freedom can be removed by fixing a gauge. One way of doing this is to require the vector potential to satisfy the Lorentz gauge condition

\[ \partial_a A^a(x) = 0. \]  

(2.105)

For such a vector field, the field equations (2.86) decouple and each of the four components of the vector potential \( A^a(x) \) satisfies the massless Klein-Gordon equation:

\[ -\partial^2 A_b(x) = 0. \]  

(2.106)

If a vector potential \( A^a(x) \) does not satisfy the Lorentz gauge condition (2.105), one may always perform a gauge transformation (2.103) to a new field \( A'^a(x) \) that has no four-divergence. We merely have to choose a gauge function \( \Lambda(x) \) in (2.103) which solves the inhomogeneous differential equation

\[ -\partial^2 \Lambda(x) = \partial_a A^a(x). \]  

(2.107)

Then \( A'^a(x) \) will satisfy \( \partial_a A'^a(x) = 0 \).

There are infinitely many solutions to equation (2.107). Given one solution \( \Lambda(x) \) which leads to the Lorentz gauge, one can add any solution of the homogenous Klein-Gordon equation without changing the four-divergence of \( A^a(x) \). The associated gauge transformations

\[ A_a(x) \rightarrow A_a(x) + \partial_a \Lambda'(x), \quad \partial^2 \Lambda'(x) = 0, \]  

(2.108)

are called restricted gauge transformations or gauge transformation of the second kind. If a vector potential \( A^a(x) \) in the Lorentz gauge solves the field equations (2.86), the gauge transformations of the second kind can be used to remove its spatial divergence \( \nabla \cdot A(x, t) \). Under (2.108), the components \( A^0(x, t) \) and \( A(x, t) \) go over into

\[ A^0(x) \rightarrow A'^0(x, t) = A^0(x, t) + \partial_0 \Lambda'(x, t), \]

\[ A(x) \rightarrow A'(x, t) = A(x, t) - \nabla \Lambda'(x, t). \]  

(2.109)

Thus, if we choose the gauge function

\[ \Lambda'(x, t) = -\int d^3x' \frac{1}{4\pi|x - x'|} \nabla \cdot A(x', t), \]  

(2.110)
then
\[ \nabla^2 \Lambda'(x, t) = \nabla \cdot A(x, t) \] (2.111)

makes the gauge-transformed field \( A'(x, t) \) divergence-free:
\[ \nabla \cdot A'(x, t) = \nabla \cdot [A(x, t) - \nabla \Lambda(x, t)] = 0. \] (2.112)

The condition
\[ \nabla \cdot A'(x, t) = 0 \] (2.113)
is known as the Coulomb gauge or radiation gauge.

The solution (2.110) to the differential equation (2.111) is still undetermined up to an arbitrary solution \( \Lambda''(x) \) of the homogeneous Poisson equation
\[ \nabla^2 \Lambda''(x, t) = 0. \] (2.114)

Together with the property \( \partial^2 \Lambda''(x, t) = 0 \) implied by (2.108), one also has
\[ \partial_t^2 \Lambda''(x, t) = 0. \] (2.115)

This leaves only trivial linear functions \( \Lambda''(x, t) \) of \( x \) and \( t \) which contribute constants to (2.109). These, in turn, are zero since the fields \( A^a(x) \) are always assumed to vanish at infinity before and after the gauge transformation.

Another possible gauge is obtained by removing the zeroth component of the vector potential \( A^a(x) \) in the field equation (2.86). This is obtained by performing the gauge transformation (2.103) with a gauge function
\[ \Lambda(x, t) = - \int^t dt' A_0(x, t'), \] (2.116)

instead of (2.110). The new field \( A'^a(x) \) has no zeroth component:
\[ A'^0(x) = 0. \] (2.117)

This is called the axial gauge. The solutions of Eqs. (2.116) are determined up to a trivial constant, so that no more gauge freedom is left.

For free fields, the Coulomb gauge and the axial gauge coincide. This is a consequence of the charge-free Coulomb law \( \nabla \cdot E = 0 \) in Eq. (2.97). By expressing \( E(x) \) explicitly in terms of the space- and time-like components of the vector potential as
\[ E(x) = -\partial_0 A(x) - \nabla A^0(x), \] (2.118)

Coulomb’s law reads
\[ \nabla^2 A^0(x, t) = -\nabla \cdot \dot{A}(x, t). \] (2.119)

This shows that if \( \nabla \cdot A(x) = 0 \), also \( A^0(x) = 0 \) (assuming zero boundary values at infinity), and vice versa.
The differential equation (2.119) can be integrated to
\[ A_0^0(x, t) = \frac{1}{4\pi} \int d^3 x' \frac{1}{|x' - x|} \nabla \cdot \dot{A}(x', t). \] (2.120)

In an infinite volume with asymptotically vanishing fields there is no freedom of adding to the left-hand side a nontrivial solution of the homogeneous Poisson equation
\[ \nabla^2 A_0^0(x, t) = 0. \] (2.121)

In the presence of charges, Coulomb’s law has a source term [see Eq. (2.97)]:
\[ \nabla \cdot E(x, t) = \rho(x, t), \] (2.122)
where \( \rho(x, t) \) is the electric charge density. In this case the vanishing of \( \nabla \cdot A(x, t) \) no longer implies \( A_0^0(x, t) \equiv 0 \). Then one has the possibility of choosing \( \Lambda(x, t) \) either to satisfy the Coulomb gauge
\[ \nabla \cdot A(x, t) \equiv 0, \] (2.123)
or the axial gauge
\[ A_0^0(x, t) \equiv 0. \] (2.124)

Only for free fields the two gauges coincide.

In a fixed gauge, the vector potential \( A^a(x) \) does not, in general, transform as a four-vector field under Lorentz transformations, which according to (1.203) and (1.210) would imply
\[ A^a(x) \xrightarrow{\Lambda} A'^a(x) = \Lambda^a_b A^b(\Lambda^{-1} x) = [c^{-i\frac{1}{2} \omega_{ab} j^{ab} A}]^a(\Lambda^{-1} x). \] (2.125)
This is only true if the gauge is fixed in a Lorentz-invariant way, for instance by the Lorentz gauge condition (2.105). In the Coulomb gauge, the right-hand side of (2.125) will be modified by an additional gauge transformation depending on \( \Lambda \) which ensures the Coulomb gauge for the transformed vector potential.

### 2.5 Maxwell-Lorentz Action for Charged Point Particles

Consider now charged relativistic massive particles interacting with electromagnetic fields and derive the Maxwell-Lorentz equations of Section 1.10 from the action approach. A single particle of charge \( e \) carries a current
\[ j^a(x) = e c \int_{-\infty}^{\infty} d\tau \dot{q}^a(\tau) \delta^{(4)}(x - q(\tau)), \] (2.126)
and the total action in an external field is given by the sum of (16.20) and (2.19):
\[ \mathcal{A} = \mathcal{A}^{\text{em}} + \mathcal{A} = -\frac{1}{4} \int d^4 x F_{ab}(x) F^{ab}(x) - m c^2 \int_{\tau_a}^{\tau_b} d\tau - \frac{1}{c} \int d^4 x j^a(x) A_a(x). \] (2.127)
In terms of the physical time $t$, the last two terms can be separated into spatial and time-like components as follows:

$$-mc^2 \int_{t_a}^{t_b} dt \sqrt{1 - \frac{\dot{q}^2}{c^2}} - e \int_{t_a}^{t_b} dt A^0(q(t), t) + \frac{e}{c} \int_{t_a}^{t_b} dt v \cdot A(q(t), t). \quad (2.128)$$

The equations of motion are obtained by writing the free-particle action in the form (2.19), and extremizing (2.127) with respect to variations $\delta q^a(\tau)$. This yields the Maxwell-Lorentz equations (1.165):

$$m \frac{d^2 q^a}{d\tau^2} = e \left[ -\frac{\partial}{\partial \tau} A^a + \frac{dq^b}{d\tau} \partial^a A_b \right] = e \left[ -\frac{dq^b}{d\tau} \partial_b A^a + \frac{dq^b}{d\tau} \partial^a A_b \right] = e F_{ab} \frac{dq^b}{d\tau}. \quad (2.129)$$

On the right-hand side we recognize the Lorentz force (1.179).

Note that in the presence of electromagnetic fields, the canonical momenta (2.11) are no longer equal to the physical momenta as in (2.15), but receive a contribution from the vector potential:

$$P_i = -\frac{\partial L}{\partial \dot{q}^i} = -(m \gamma \dot{q}^i + \frac{e}{c} A^i) = p_i + \frac{e}{c} A_i. \quad (2.130)$$

Including the zeroth component, the canonical four-momentum is

$$P_a = p_a + \frac{e}{c} A_a. \quad (2.131)$$

The zeroth component of $P_a$ coincides with $1/c$ times the energy defined by the Legendre transform [recall (2.96)]:

$$P_0 = \frac{1}{c} (H + eA^0) = -\frac{1}{c} (P_i q^i - L). \quad (2.132)$$

### 2.6 Scalar Field with Electromagnetic Interaction

The spacetime derivatives of a plane wave such as (1.155) yields the energy-momentum of the particle whose probability amplitude is described by the wave:

$$i\hbar \partial_a \phi_p(x) = p_a \phi_p(x). \quad (2.133)$$

In the presence of electromagnetism, the role of the momentum four-vector is taken over by the momenta (2.131). In the Lagrangian density (2.27) of the scalar field, this is accounted for by the so-called minimal replacement of the derivatives by the covariant derivatives:

$$\partial_a \phi(x) \rightarrow D_a \phi(x) \equiv \left[ \partial_a + i \frac{e}{c\hbar} A_a(x) \right] \phi(x). \quad (2.134)$$
The Lagrangian density of a scalar field with electromagnetic interactions is therefore

$$\mathcal{L}(x) = \frac{\hbar^2}{2} [D_a \phi(x)]^* D^a \phi(x) - m^2 \phi^* \phi(x) - \frac{1}{4} \int d^4 x \, F^{ab}(x) F_{ab}(x). \quad (2.135)$$

It governs the so-called scalar electrodynamics.

This expression is invariant under local gauge transformations (2.103) of the electromagnetic field, if we simultaneously multiply the scalar field by an $x$-dependent phase factor:

$$\varphi(x) \to e^{i e A(x)/c} \varphi(x). \quad (2.136)$$

By extremization of the action in natural units $A = \int d^4 x \, \mathcal{L}(x)$ we find the Euler-Lagrange equation and its conjugate

$$\frac{\delta A}{\delta \varphi}(x) = (-D^2 - m^2) \varphi(x), \quad \frac{\delta A}{\delta \varphi^*}(x) = (-D'^2 - m^2) \varphi^*(x). \quad (2.137)$$

In the presence of the electromagnetic field, the particle current density (2.63) turns into the charge current density

$$j_a(x) = e \frac{i}{2} \phi^* D_a \phi + \text{c.c.} = e \frac{i}{2} \phi^* \partial_a \phi - \frac{e^2}{c} A_a(x) \phi^* \phi. \quad (2.138)$$

This satisfies the same conservation law (2.64) as the current density of the free scalar field, as we can verify by a short calculation:

$$\partial_a j^a = \partial_a \left[ e \frac{i}{2} \phi^* D^a \phi \right] + \text{c.c.} = e \frac{i}{2} \partial_a \phi^* D^a \phi + e \frac{i}{2} \phi^* \partial_a D^a \phi + \text{c.c.} \quad (2.139)$$

$$= e \frac{i}{2} \partial_a \phi^* D^a \phi + e \frac{i}{2} \phi^* D^2 \phi - e \frac{i}{2} A_a \phi^* D_a \phi + \text{c.c.} = e \frac{i}{2} D_a \phi^* D^a \phi - m^2 e \frac{i}{2} \phi^* \phi + \text{c.c.} = 0.$$  

### 2.7 Dirac Fields

An action whose extremum yields the Dirac equation (1.212) is, in natural units,

$$\mathcal{D} \mathcal{A} = \int d^4 x \, \mathcal{D} \mathcal{L}(x) = \int \int d^4 x \, \bar{\psi}(x) \left( i \gamma^a \partial_a \right) \psi(x) \quad (2.140)$$

where

$$\bar{\psi}(x) \equiv \psi^\dagger(x) \gamma^0, \quad (2.141)$$

and the matrices $\gamma^a$ satisfy the anticommutation rules (1.218). The Dirac equation and its conjugate are obtained from the extremal principle

$$\frac{\delta \mathcal{D} \mathcal{A}}{\delta \bar{\psi}(x)} = (i \gamma^a \partial_a - m) \psi(x) = 0, \quad \frac{\delta \mathcal{D} \mathcal{A}}{\delta \psi(x)} = \bar{\psi}(x) \left( -i \gamma^a \partial_a - m \right) \psi(x) = 0. \quad (2.142)$$

The action (2.140) is invariant under the Lorentz transformations of spinors (1.230). For the mass term this follows from the fact that

$$D^\dagger(\Lambda) \gamma^0 D(\Lambda) = \gamma^0. \quad (2.143)$$
This equation is easily verified by inserting the explicit matrices (1.213) and (1.227). If we define
\[ \bar{D} \equiv \gamma^0 D^1 \gamma^0, \] (2.144)
this implies
\[ \bar{D}(\Lambda) D(\Lambda) = 1, \] (2.145)
so that the mass term in the Lagrangian density transforms like a scalar field in (1.163):
\[ \bar{\psi}(x) \psi(x) \xrightarrow{\Lambda} \bar{\psi}(x) \Lambda^a \psi(x) = \bar{\psi}(\Lambda^{-1}x) \psi(\Lambda^{-1}x). \] (2.146)

Consider now the gradient term in the action (2.140). Its invariance is a consequence of the vector property of the Dirac matrices under the spin representation of the Lorentz group derived in Eq. (1.229), which can be rewritten, due to (2.145), as
\[ \bar{D}(\Lambda) \gamma^a D(\Lambda) = D^{-1}(\Lambda) \gamma^b D(\Lambda) = \Lambda^a_b \gamma^b. \] (2.147)

From this we derive at once that
\[ \bar{\psi}(x) \gamma^a \psi(x) \xrightarrow{\Lambda} \bar{\psi}(x) \gamma^a \psi(x) = \Lambda^a_b \bar{\psi}(\Lambda^{-1}x) \gamma^b \psi(\Lambda^{-1}x), \] (2.148)
and
\[ \bar{\psi}(x) \gamma^a \partial_a \psi(x) \xrightarrow{\Lambda} \bar{\psi}(x) \gamma^a \partial_a \psi(x) = [\bar{\psi} \gamma^b \partial_a \psi](\Lambda^{-1}x). \] (2.149)

Thus also the gradient term in the Dirac Lagrangian density transforms like a scalar field, and so does the full Lagrangian density as in (2.28), which makes the action (2.140) invariant under Lorentz transformations, due to (2.34).

After the discussion in Section 2.6 we know how to couple the Dirac field to electromagnetism. We simply have to replace the derivative in the Lagrangian density by the covariant derivative (2.134), and obtain the gauge-invariant Lagrangian density of the electrodynamics
\[ \mathcal{L}(x) = \bar{\psi}(x) \left( i \gamma^a D_a - m \right) \psi(x) - \frac{1}{4} \int d^4 x F^{ab}(x) F_{ab}(x). \] (2.150)

This is invariant under local gauge transformations (2.103), if we simultaneously multiply the Dirac field by an \( x \)-dependent phase factor
\[ \psi(x) \rightarrow e^{ie\Lambda(x)/c} \psi(x). \] (2.151)

The interaction term in this Lagrangian density comes entirely from the covariant derivative and reads, more explicitly,
\[ \mathcal{L}^{\text{int}}(x) = -\frac{1}{c} \int d^4 x A_a(x) j^a(x), \] (2.152)
where
\[ j^a(x) \equiv e \bar{\psi}(x) \gamma^a \psi(x) \] (2.153)
is the current density of the electrons.
By extremizing the action $\mathcal{A} = \int d^4x \ L (x)$ we now find the Euler-Lagrange equation and its conjugate
\begin{equation}
\frac{\delta \mathcal{A}}{\delta \psi(x)} = (i\gamma^a D_a - m)\psi(x) = 0, \quad \frac{\delta \mathcal{A}}{\delta \psi^*(x)} = \overline{\psi}(x)(-i\gamma^a \overline{D}_a^* - m)\psi(x) = 0.
\end{equation}

For classical fields obeying these equation, the current density (2.153) satisfies the same local conservation law as the scalar field in Eq. (2.139),
\begin{equation}
\partial_a j^a(x) = 0,
\end{equation}
as can be verified by the much simpler calculation than in (2.139):
\begin{equation}
\partial_a j^a = e\partial_a(\overline{\psi}\gamma^a\psi) = e\overline{\psi}\gamma^a\partial_a\psi + e\overline{\psi}\gamma^a\partial_a\psi = e\overline{\psi}\gamma^a\overline{D}_a^*\psi + e\overline{\psi}\gamma^a D_a \psi = 0.
\end{equation}

**Notes and References**


Continuous Symmetries and Conservation Laws.  
Noether’s Theorem

In many physical systems, the action is invariant under some continuous set of transformations. Then there exist local and global conservation laws analogous to current and charge conservation in electrodynamics. With the help of Poisson brackets, the analogs of the charges can be used to generate the symmetry transformation, from which they were derived. After field quantization, the Poisson brackets become commutators of operators associated with these charges.

3.1 Continuous Symmetries and Conservation Laws

Consider first a simple mechanical system with a generic action

\[ \mathcal{A} = \int_{t_a}^{t_b} dt \, L(q(t), \dot{q}(t)), \quad (3.1) \]

and subject it to a continuous set of local transformations of the dynamical variables:

\[ q(t) \rightarrow q'(t) = f(q(t), \dot{q}(t)), \quad (3.2) \]

where \( f(q(t), \dot{q}(t)) \) is some function of \( q(t) \) and \( \dot{q}(t) \). In general, \( q(t) \) will carry various labels as in (2.1) which are suppressed, for brevity. If the transformed action

\[ \mathcal{A}' = \int_{t_a}^{t_b} dt \, L(q'(t), \dot{q}'(t)) \quad (3.3) \]

is the same as \( \mathcal{A} \), up to boundary terms, then (3.2) is called a symmetry transformation.

3.1.1 Group Structure of Symmetry Transformations

For any two symmetry transformations, we may defined a product by performing the transformations successively. The result is certainly again a symmetry transformation. Since all transformations can be undone, they possess an inverse. Thus,
symmetry transformations form a group called the symmetry group of the system. It is important that the equations of motion are not used when showing that the action $A'$ is equal to $A$, up to boundary terms.

### 3.1.2 Substantial Variations

For infinitesimal symmetry transformations (3.2), the difference

$$\delta_s q(t) \equiv q'(t) - q(t)$$

will be called a symmetry variation. It has the general form

$$\delta_s q(t) = \epsilon \Delta(q(t), \dot{q}(t)),$$

where $\epsilon$ is a small parameter. Symmetry variations must not be confused with the variations $\delta q(t)$ used in Section 2.1 to derive the Euler-Lagrange equations (2.8), which always vanish at the ends, $\delta q(t_b) = \delta q(t_a) = 0$ [recall (1.4)]. This is usually not true for symmetry variation $\delta_s q(t)$.

Another name for the symmetry variation (3.5) is substantial variation. It is defined for any function of spacetime $f(x)$ as the difference between $f(x)$ and a transformed function $f'(x)$ when evaluated at the same numerical values of the coordinates $x$ (which usually correspond to two different points in space):

$$\delta_s f(x) = f(x) - f'(x).$$

### 3.1.3 Conservation Laws

Let us calculate the change of the action under a substantial variation (3.5). Using the chain rule of differentiation and a partial integration we obtain

$$\delta_s A = \int_{t_a}^{t_b} dt \left[ \frac{\partial L}{\partial q(t)} - \partial_t \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \right] \delta_s q(t) + \left. \frac{\partial L}{\partial \dot{q}(t)} \right|_{t_a}^{t_b}. $$

Let us denote the solutions of the Euler-Lagrange equations (2.8) by $q_c(t)$ and call them classical orbits, or briefly orbits. For orbits, only the boundary terms in (3.7) survive, and we are left with

$$\delta_s A = \epsilon \left. \frac{\partial L}{\partial \dot{q}} \Delta(q, \dot{q}) \right|_{t_a}^{t_b}, \quad \text{for} \quad q(t) = q_c(t).$$

By the symmetry assumptions, $\delta_s A$ vanishes or is equal to a surface term. In the first case, the quantity

$$Q(t) \equiv \frac{\partial L}{\partial \dot{q}} \Delta(q, \dot{q}), \quad \text{for} \quad q(t) = q_c(t)$$

is called the conserved quantity.
Continuous Symmetries and Conservation Laws. Noether’s Theorem

is the same at times \( t = t_a \) and \( t = t_b \). Since \( t_b \) is arbitrary, \( Q(t) \) is independent of the time \( t \), i.e., it satisfies

\[
Q(t) \equiv Q. \quad (3.10)
\]

It is a conserved quantity, a constant of motion along the orbit. The expression on the right-hand side of (3.9) is called a Noether charge.

In the second case, \( \delta_q q(t) \) is equal to a boundary term

\[
\delta_q A = \epsilon \Lambda(q, \dot{q}) \bigg|_{t_a}^{t_b}, \quad (3.11)
\]

and the conserved Noether charge becomes

\[
Q(t) = \frac{\partial L}{\partial \dot{q}} \Delta(q, \dot{q}) - \Lambda(q, \dot{q}), \quad \text{for} \quad q(t) = q_c(t). \quad (3.12)
\]

It is possible to derive the constant of motion (3.12) also without invoking the action, starting from the Lagrangian \( L(q, \dot{q}) \). We expand its substantial variation of \( L(q, \dot{q}) \) as follows:

\[
\delta_s L \equiv L(q + \delta_s q, \dot{q} + \delta_s \dot{q}) - L(q, \dot{q}) = \left[ \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}(t)} \right] \delta_q q(t) + \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}(t)} \delta_q q(t) \right]. \quad (3.13)
\]

On account of the Euler-Lagrange equations (2.8), the first term on the right-hand side vanishes as before, and only the last term survives. The assumption of invariance of the action up to a possible surface term in Eq. (3.11) is equivalent to assuming that the substantial variation of the Lagrangian is at most a total time derivative of some function \( \Lambda(q, \dot{q}) \):

\[
\delta_s L(q, \dot{q}, t) = \epsilon \frac{d}{dt} \Lambda(q, \dot{q}). \quad (3.14)
\]

Inserting this into the left-hand side of (3.13), we find the equation

\[
\epsilon \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}} \Delta(q, \dot{q}) - \Lambda(q, \dot{q}) \right] = 0, \quad \text{for} \quad q(t) = q_c(t) \quad (3.15)
\]

thus recovering again the conserved Noether charge (3.12).

3.1.4 Alternative Derivation of Conservation Laws

Let us subject the action (3.1) to an arbitrary variation \( \delta q(t) \), which may be nonzero at the boundaries. Along a classical orbit \( q_c(t) \), the first term in (3.7) vanishes, and the action changes at most by the boundary term:

\[
\delta A = \frac{\partial L}{\partial q} \delta_q q \bigg|_{t_a}^{t_b}, \quad \text{for} \quad q(t) = q_c(t). \quad (3.16)
\]
3.1 Continuous Symmetries and Conservation Laws

This observation leads to another derivation of Noether’s theorem. Suppose we perform on \( q(t) \) a so-called local symmetry transformations, which generalizes the previous substantial variations (3.5) to a time-dependent parameter \( \epsilon \):

\[
\delta_t^\epsilon q(t) = \epsilon(t) \Delta(q(t), \dot{q}(t)). \tag{3.17}
\]

The superscript \( t \) on \( \delta_t^\epsilon \) emphasized the extra time dependence in \( \epsilon(t) \). If the variations (3.17) are applied to a classical orbit \( q_c(t) \), the action changes by the boundary term (3.16).

This will now be expressed in a more convenient way. For this purpose we introduce the infinitesimally transformed orbit

\[
q^\epsilon(t) = q(t) + \delta_t^\epsilon q(t) = q(t) + \epsilon(t) \Delta(q(t), \dot{q}(t)), \tag{3.18}
\]

and the transformed Lagrangian

\[
L^\epsilon(t) = L(q^\epsilon(t), \dot{q}^\epsilon(t)). \tag{3.19}
\]

Then the local substantial variation of the action with respect to the time-dependent parameter \( \epsilon(t) \) is

\[
\delta_\epsilon A = \int_{t_a}^{t_b} dt \left[ \frac{\partial L^\epsilon(t)}{\partial \epsilon(t)} - \frac{d}{dt} \frac{\partial L^\epsilon(t)}{\partial \dot{\epsilon}(t)} \right] \epsilon(t) \bigg|_{t_a}^{t_b}. \tag{3.20}
\]

Along a classical orbit, the action is extremal. Hence the infinitesimally transformed action

\[
A^\epsilon \equiv \int_{t_a}^{t_b} dt \, L(q^\epsilon(t), \dot{q}^\epsilon(t)) \tag{3.21}
\]

must satisfy the equation

\[
\frac{\delta A^\epsilon}{\delta \epsilon(t)} = 0. \tag{3.22}
\]

This holds for an arbitrary time dependence of \( \epsilon(t) \), in particular for \( \epsilon(t) \) which vanishes at the ends. In this case, (3.22) leads to an Euler-Lagrange type of equation

\[
\frac{\partial L^\epsilon(t)}{\partial \epsilon(t)} - \frac{d}{dt} \frac{\partial L^\epsilon(t)}{\partial \dot{\epsilon}(t)} = 0, \quad \text{for} \quad q(t) = q_c(t). \tag{3.23}
\]

This can also be checked explicitly differentiating (3.19) according to the chain rule of differentiation:

\[
\frac{\partial L^\epsilon(t)}{\partial \epsilon(t)} = \frac{\partial L}{\partial q(t)} \Delta(q, \dot{q}) + \frac{\partial L}{\partial \dot{q}(t)} \dot{\Delta}(q, \dot{q}), \quad \tag{3.24}
\]

\[
\frac{\partial L^\epsilon(t)}{\partial \dot{\epsilon}(t)} = \frac{\partial L}{\partial q(t)} \Delta(q, \dot{q}). \quad \tag{3.25}
\]
and inserting on the right-hand side the ordinary Euler-Lagrange equations (1.5). Note that (3.25) can also be written as

\[ \frac{\partial L^{(t)}}{\partial \epsilon} = \frac{\partial L}{\partial q(t)} \frac{\delta q(t)}{\epsilon(t)} \tag{3.26} \]

We now invoke the symmetry assumption, that the action is a pure surface term under the time-independent transformations (3.17). This implies that

\[ \frac{\partial L'}{\partial \epsilon} = \frac{\partial L^{(t)}}{\partial \epsilon(t)} = \frac{d}{dt} \Lambda. \tag{3.27} \]

Combining this with (3.23), we derive a conservation law for the charge:

\[ Q = \frac{\partial L^{(t)}}{\partial \epsilon(t)} - \Lambda, \quad \text{for} \quad q(t) = q_c(t). \tag{3.28} \]

Inserting here Eq. (3.25) we find that this is the same charge as the previous (3.12).

### 3.2 Time Translation Invariance and Energy Conservation

As a simple but physically important example consider the case that the Lagrangian does not depend explicitly on time, i.e., that \( L(q, \dot{q}) \equiv L(q, \dot{q}) \). Let us perform a time translation on the system, so that the same events happen at a new time

\[ t' = t - \epsilon. \tag{3.29} \]

The time-translated orbit has the time dependence

\[ q'(t') = q(t), \tag{3.30} \]

i.e., the translated orbit \( q'(t) \) has at the time \( t' \) the same value as the orbit \( q(t) \) had at the original time \( t \). For the Lagrangian, this implies that

\[ L'(t') = L(q'(t'), \dot{q}'(t')) = L(q(t), \dot{q}(t)) = L(t). \tag{3.31} \]

This makes the action (3.3) equal to (3.1), up to boundary terms. Thus time-independent Lagrangians possess time translation symmetry.

The associated substantial variations of the form (3.5) read

\[ \delta_s q(t) = q'(t) - q(t) = q(t' + \epsilon) - q(t) \]

\[ = q(t') + \epsilon \dot{q}(t') - q(t) = \epsilon \dot{q}(t), \tag{3.32} \]

Under these, the Lagrangian changes by

\[ \delta_s L = L(q'(t), \dot{q}'(t)) - L(q(t), \dot{q}(t)) = \frac{\partial L}{\partial q} \delta_s q(t) + \frac{\partial L}{\partial \dot{q}} \delta_s \dot{q}(t). \tag{3.33} \]
Inserting $\delta_s q(t)$ from (3.32) we find, without using the Euler-Lagrange equation,

$$\delta_s L = \epsilon \left( \frac{\partial L}{\partial \dot{q}} \dot{q} + \frac{\partial L}{\partial \ddot{q}} \ddot{q} \right) = \epsilon \frac{d}{dt} L.$$  

This has precisely the derivative form (3.14) with $\Lambda = L$, thus confirming that time translations are symmetry transformations.

According to Eq. (3.12), we find the Noether charge

$$Q = H \equiv \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}), \quad \text{for } q(t) = q_c(t)$$  

(3.35)

to be a constant of motion. This is recognized as the Legendre transform of the Lagrangian, which is the Hamiltonian (2.10) of the system.

Let us briefly check how this Noether charge is obtained from the alternative formula (3.12). The time-dependent substantial variation (3.17) is here

$$\delta_t \delta_s q(t) = \epsilon(t) \dot{q}(t)$$  

(3.36)

under which the Lagrangian is changed by

$$\delta_t \delta_s L = \frac{\partial L}{\partial \dot{q}} \dot{\epsilon} \dot{q} + \frac{\partial L}{\partial \ddot{q}} (\epsilon \ddot{q} + \epsilon \dddot{q}) = \frac{\partial L^e}{\partial \dot{\epsilon}} \dot{\epsilon} + \frac{\partial L^e}{\partial \ddot{\epsilon}} \ddot{\epsilon},$$  

(3.37)

with

$$\frac{\partial L^e}{\partial \dot{\epsilon}} = \frac{\partial L}{\partial \dot{q}} \dot{q}$$  

(3.38)

and

$$\frac{\partial L^e}{\partial \ddot{\epsilon}} = \frac{\partial L}{\partial \ddot{q}} \ddot{q} + \frac{\partial L}{\partial \dddot{q}} \dddot{q} = \frac{d}{dt} L.$$  

(3.39)

The last equation confirms that time translations fulfill the symmetry condition (3.27), and from (3.38) we see that the Noether charge (3.28) coincides with the Hamiltonian found in Eq. (3.12).

### 3.3 Momentum and Angular Momentum

While the conservation law of energy follow from the symmetry of the action under time translations, conservation laws of momentum and angular momentum are found if the action is invariant under translations and rotations, respectively.

Consider a Lagrangian of a point particle in a Euclidean space

$$L = L(q^i(t), \dot{q}^i(t)).$$  

(3.40)

In contrast to the previous discussion of time translation invariance, which was applicable to systems with arbitrary Lagrange coordinates $q(t)$, we denote the coordinates here by $\dot{q}^i$, with the superscripts $i$ emphasizing that we are now dealing with Cartesian coordinates. If the Lagrangian depends only on the velocities $\dot{q}^i$ and
not on the coordinates $q^i$ themselves, the system is *translationally invariant*. If it depends, in addition, only on $\dot{q}^2 = \dot{q}^i \dot{q}^i$, it is also rotationally invariant.

The simplest example is the Lagrangian of a point particle of mass $m$ in Euclidean space:

$$L = \frac{m}{2} \dot{q}^2. \quad (3.41)$$

It exhibits both invariances, leading to conserved Noether charges of momentum and angular momentum, as we shall now demonstrate.

### 3.3.1 Translational Invariance in Space

Under a spatial translation, the coordinates $q^i$ of the particle change to

$$q'^i = q^i + \epsilon^i, \quad (3.42)$$

where $\epsilon^i$ are small numbers. The infinitesimal translations of a particle path are [compare (3.5)]

$$\delta_s q^i(t) = \epsilon^i. \quad (3.43)$$

Under these, the Lagrangian changes by

$$\delta_s L = L(q'^i(t), \dot{q}'^i(t)) - L(q^i(t), \dot{q}^i(t)) = \frac{\partial L}{\partial q^i} \delta_s q^i = \frac{\partial L}{\partial \dot{q}^i} \epsilon^i = 0. \quad (3.44)$$

By assumption, the Lagrangian is independent of $q^i$, so that the right-hand side vanishes. This is to be compared with the substantial variation of the Lagrangian around a classical orbit calculated with the help of the Euler-Lagrange equation:

$$\delta_s L = \left( \frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) \delta_s q^i + \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}^i} \delta q^i \right] = \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}^i} \right] \epsilon^i \quad (3.45)$$

This has the form (3.8), from which we extract a conserved Noether charge (3.9) for each coordinate $q^i$, to be called $p^i$:

$$p^i = \frac{\partial L}{\partial \dot{q}^i}. \quad (3.46)$$

Thus the Noether charges associated with translational invariance are simply the canonical momenta of the point particle.

### 3.3.2 Rotational Invariance

Under rotations, the coordinates $q^i$ of the particle change to

$$q'^i = R^i_{\ j} q^j. \quad (3.47)$$
where \( R^i_j \) are the orthogonal \( 3 \times 3 \) -matrices (1.8). Infinitesimally, these can be written as

\[
R^i_j = \delta^i_j - \varphi_k \epsilon_{kij}
\]  

(3.48)

where \( \varphi \) is the infinitesimal rotation vector in Eq. (1.57). The corresponding rotation of a particle path is

\[
\delta_s q^i(t) = q'^i(t) - q^i(t) = -\varphi^k \epsilon_{kij} q^j(\tau).
\]  

(3.49)

In the antisymmetric tensor notation (1.55) with \( \omega_{ij} = \varphi^k \epsilon_{kij} \), we write

\[
\delta_s q^i = -\omega_{ij} q^j.
\]  

(3.50)

Under this, the substantial variation of the Lagrangian (3.41)

\[
\delta_s L = L(q'^i(t), \dot{q}'^i(t)) - L(q^i(t), \dot{q}^i(t)) = \frac{\partial L}{\partial q^i} \delta_s q^i + \frac{\partial L}{\partial \dot{q}^i} \delta_s \dot{q}^i
\]  

(3.51)

becomes

\[
\delta_s L = -\left( \frac{\partial L}{\partial q^j} \dot{q}^j + \frac{\partial L}{\partial \dot{q}^j} \dot{q}^j \right) \omega_{ij} = 0.
\]  

(3.52)

For any Lagrangian depending only on the rotational invariants \( q^2, \dot{q}^2, q \cdot \dot{q} \) and powers thereof, the right-hand side vanishes on account of the antisymmetry of \( \omega_{ij} \). This ensures the rotational symmetry for the Lagrangian (3.41).

We now calculate the substantial variation of the Lagrangian once more using the Euler-Lagrange equations:

\[
\delta_s L = \left( \frac{\partial L}{\partial \dot{q}^j} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^j} \right) \delta_s q^i + \frac{d}{dt} \left[ \frac{\partial L}{\partial q^i} \delta_s q^i \right]
\]  

(3.53)

The right-hand side yields the conserved Noether charges of the type (3.9), one for each antisymmetric pair \( i, j \):

\[
L_{ij} = q^j \frac{\partial L}{\partial \dot{q}^j} - q^i \frac{\partial L}{\partial \dot{q}^i} = q^j p^j - q^i p^i.
\]  

(3.54)

These are the conserved components of angular momentum for a Cartesian system in any dimension.

In three dimensions, we may prefer working with the original rotation angles \( \varphi^k \), in which case we would have found the angular momentum in the standard form

\[
L_k = \frac{1}{2} \epsilon_{kij} L^{ij} = (q \times p)^k.
\]  

(3.55)
3.3.3 Center-of-Mass Theorem

Let us now study symmetry transformations corresponding to a uniform motion of the coordinate system described by Galilei transformations (1.11), (1.12). Consider a set of free massive point particles in Euclidean space described by the Lagrangian

\[ L(q^i_n) = \sum_n \frac{m_n}{2} \dot{q}^i_n. \]  

(3.56)

The infinitesimal substantial variation associated with the Galilei transformations are

\[ \delta s q^i_n(t) = \dot{q}^i_n(t) - q^i_n(t) = -v^i t, \]

(3.57)

where \( v^i \) is a small relative velocity along the \( i \)th axis. This changes the Lagrangian by

\[ \delta s L = L(q^i_n - v^i t, \dot{q}^i_n - v^i) - L(q^i_n, \dot{q}^i_n). \]

(3.58)

Inserting here (3.56), we find

\[ \delta s L = \sum_n \frac{m_n}{2} \left[ (\dot{q}^i_n - v^i)^2 - (\dot{q}^i_n)^2 \right], \]

(3.59)

which can be written as a total time derivative

\[ \delta s L = \frac{d}{dt} \Lambda = \frac{d}{dt} \sum_n m_n \left[ -\dot{q}^i_n v^i + \frac{v^2}{2} t \right] \]

(3.60)

proving that Galilei transformations are symmetry transformations in the Noether sense. Note that terms quadratic in \( v^i \) are omitted in the last expression since the velocities \( v^i \) in (3.57) are infinitesimal, by assumption.

By calculating \( \delta s L \) once more via the chain rule with the help of the Euler-Lagrange equations, and equating the result with (3.60), we find the conserved Noether charge

\[ Q = \sum_n \frac{\partial L}{\partial \dot{q}^i_n} \delta s q^i_n - \Lambda \]

\[ = \left( -\sum_n m_n \dot{q}^i_n t + \sum_n m_n q^i_n \right) v^i. \]

(3.61)

Since the direction of the velocities \( v^i \) is arbitrary, each component is separately a constant of motion:

\[ N^i = -\sum_n m_n \dot{q}^i t + \sum_n m_n q^i = \text{const}. \]

(3.62)

This is the well-known center-of-mass theorem [1]. Indeed, introducing the center-of-mass coordinates

\[ q^i_{CM} \equiv \frac{\sum_n m_n q^i_n}{\sum_n m_n}, \]

(3.63)
3.3 Momentum and Angular Momentum

and velocities

\[ v^i_{\text{CM}} = \frac{\sum_n m_n \dot{q}^i_n}{\sum_n m_n}, \tag{3.64} \]

the conserved charge (3.62) can be written as

\[ N^i = \sum_n m_n (-v^i_{\text{CM}} t + \dot{q}^i_{\text{CM}}). \tag{3.65} \]

The time-independence of \( N^i \) implies that the center-of-mass moves with uniform velocity according to the law

\[ q^i_{\text{CM}}(t) = q^i_{\text{CM},0} + v^i_{\text{CM}} t, \tag{3.66} \]

where

\[ q^i_{\text{CM},0} = \frac{N^i}{\sum_n m_n} \tag{3.67} \]

is the position of the center of mass at \( t = 0 \).

Note that in non-relativistic physics, the center of mass theorem is a consequence of momentum conservation since momentum \( \equiv \) mass \( \times \) velocity. In relativistic physics, this is no longer true.

### 3.3.4 Conservation Laws from Lorentz Invariance

In relativistic physics, particle orbits are described by functions in Minkowski spacetime \( q^a(\sigma) \), where \( \sigma \) is a Lorentz-invariant length parameter. The action is an integral over some Lagrangian:

\[ \mathcal{A} = \int_{\sigma_0}^{\sigma_f} d\sigma L(q^a(\sigma), \dot{q}^a(\sigma)), \tag{3.68} \]

where the dot denotes the derivative with respect to the parameter \( \sigma \). If the Lagrangian depends only on invariant scalar products \( q^a q^a, \dot{q}^a \dot{q}^a, \dot{q}^a \dot{q}^b \), then it is invariant under Lorentz transformations

\[ q^a \rightarrow q'^a = \Lambda^a_b q^b \tag{3.69} \]

where \( \Lambda^a_b \) are the pseudo-orthogonal \( 4 \times 4 \) -matrices (1.28).

A free massive point particle in spacetime has the Lagrangian [see (2.19)]

\[ L(\dot{q}(\sigma)) = -mc\sqrt{g_{ab} \dot{q}^a \dot{q}^b}, \tag{3.70} \]

so that the action (3.68) is invariant under arbitrary reparametrizations \( \sigma \rightarrow f(\sigma) \). Since the Lagrangian depends only on \( \dot{q}(\sigma) \), it is invariant under arbitrary translations of the coordinates:

\[ \delta_q q^a(\sigma) = q^a(\sigma) - \epsilon^a(\sigma), \tag{3.71} \]
for which \( \delta_s L = 0 \). Calculating this variation once more with the help of the Euler-Lagrange equations, we find

\[
\delta_s L = \int_{\sigma_a}^{\sigma_b} d\sigma \left( \frac{\partial L}{\partial q^a} \delta q^a + \frac{\partial L}{\partial \dot{q}^a} \dot{\delta q}^a \right) = -c^a \int_{\sigma_a}^{\sigma_b} d\sigma \frac{d}{d\sigma} \left( \frac{\partial L}{\partial \dot{q}^a} \right). \tag{3.72}
\]

From this we obtain the conserved Noether charges

\[
p_a \equiv -\frac{\partial L}{\partial \dot{q}^a} = \frac{\dot{q}_a}{\sqrt{g_{ab} \dot{q}^b / c^2}} = m u_a^a, \tag{3.73}
\]

which satisfy the conservation law

\[
\frac{d}{d\sigma} p_a(\sigma) = 0. \tag{3.74}
\]

The Noether charges \( p_a(\sigma) \) are the conserved four-momenta (1.144) of the free relativistic particle, derived in Eq. (2.20) from the canonical formalism. The four-vector

\[
u^a \equiv \frac{\dot{q}^a}{\sqrt{g_{ab} \dot{q}^b / c^2}} \tag{3.75}
\]

is the relativistic four-velocity of the particle. It is the reparametrization-invariant expression for the four-velocity \( \dot{q}_a(\tau) = u_a(\tau) \) in Eqs. (2.22) and (1.144). A sign change is made in Eq. (3.73) to agree with the canonical definition of the covariant momentum components in (2.20). By choosing for \( \sigma \) the physical time \( t = q^0 / c \), we can express \( u^a \) in terms of the physical velocities \( v_i = dq_i / dt \), as in (1.145):

\[
u^a = \gamma (1, v^i / c), \quad \text{with} \quad \gamma \equiv \sqrt{1 - v^2 / c^2}. \tag{3.76}
\]

For small Lorentz transformations near the identity we write

\[
\Lambda^a_b = \delta^a_b + \omega^a_b \tag{3.77}
\]

where

\[
\omega^a_b = g^{ac} \omega_{cb} \tag{3.78}
\]

is an arbitrary infinitesimal antisymmetric matrix. An infinitesimal Lorentz transformation of the particle path is

\[
\delta_s q^a(\sigma) = \dot{q}^a(\sigma) - q^a(\sigma) = \omega^a_b q^b(\sigma). \tag{3.79}
\]

Under it, the substantial variation of a Lorentz-invariant Lagrangian vanishes:

\[
\delta_s L = \left( \frac{\partial L}{\partial q^b} q^b + \frac{\partial L}{\partial \dot{q}^b} \dot{q}^b \right) \omega^a_b = 0. \tag{3.80}
\]
This is to be compared with the substantial variation of the Lagrangian calculated via the chain rule with the help of the Euler-Lagrange equation

\[
\delta s L = \left( \frac{\partial L}{\partial q^a} - \frac{d}{d\sigma} \frac{\partial L}{\partial \dot{q}^a} \right) \delta q^a + \frac{d}{d\sigma} \left[ \frac{\partial L}{\partial \dot{q}^a} \delta q^a \right] = \frac{d}{d\sigma} \left[ \frac{\partial L}{\partial \dot{q}^a} \delta q^a \right] = \frac{1}{2} \omega^{ab} \frac{d}{d\sigma} \left( q^a \frac{\partial L}{\partial \dot{q}^b} - q^b \frac{\partial L}{\partial \dot{q}^a} \right). \tag{3.81}
\]

By equating this with (3.80) we obtain the conserved rotational Noether charges

\[
L^{ab} = -q^a \frac{\partial L}{\partial q^b} + q^b \frac{\partial L}{\partial q^a} = q^a p^b - q^b p^a. \tag{3.82}
\]

They are the four-dimensional generalizations of the angular momenta (3.54).

The Noether charges \( L^{ij} \) coincide with the components (3.54) of angular momentum. The conserved components

\[
L^{0i} = q^0 \dot{p}^i - q^i \dot{p}^0 \equiv M_i \tag{3.83}
\]
yield the relativistic generalization of the center-of-mass theorem (3.62):

\[
M_i = \text{const}. \tag{3.84}
\]

### 3.4 Generating the Symmetries

As mentioned in the introduction to this chapter, there is a second important relation between invariances and conservation laws. The charges associated with continuous symmetry transformations can be used to generate the symmetry transformation from which they it was derived. In the classical theory, this is done with the help of Poisson brackets:

\[
\delta_s \dot{q} = \epsilon \{ \bar{Q}, \dot{q}(t) \}. \tag{3.85}
\]

After canonical quantization, the Poisson brackets turn into \(-i\) times commutators, and the charges become operators, generating the symmetry transformation by the operation

\[
\delta_s \hat{q} = -i \epsilon [ \bar{Q}, \hat{q}(t) ]. \tag{3.86}
\]

The most important example for this quantum-mechanical generation of symmetry transformations is the effect of the Noether charge (3.35) derived in Section 3.2 from the invariance of the system under time displacement. That Noether charge \( Q \) was the Hamiltonian \( H \), whose operator version generates the infinitesimal time displacements (3.32) by the Heisenberg equation of motion

\[
\delta_s q(t) = \epsilon \dot{q}(t) = -i \epsilon [ \hat{H}, \hat{q}(t) ], \tag{3.87}
\]
as a special case of the general Noether relation (3.86).

The canonical quantization is straightforward if the Lagrangian has the standard form

\[ L(q, \dot{q}) = \frac{m}{2} \dot{q}^2 - V(q). \] (3.88)

Then the operator version of the canonical momentum \( p \equiv \dot{q} \) satisfies the equal-time commutation rules

\[ [\hat{p}(t), \hat{q}(t)] = -i, \quad [\hat{p}(t), \hat{p}(t)] = 0, \quad [\hat{q}(t), \hat{q}(t)] = -i. \] (3.89)

The Hamiltonian

\[ H = \frac{p^2}{2m} + V(\hat{q}) \] (3.90)

turns directly into the Hamiltonian operator

\[ \hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}). \] (3.91)

If the Lagrangian does not have the standard form (3.88), quantization is a nontrivial problem, solved in the textbook [2].

Another important example is provided by the charges (3.46) derived in Section 3.3.1 from translational symmetry. After quantization, the commutator (3.86) generating the transformation (3.43) becomes

\[ \epsilon^j = i\epsilon^i [\hat{p}^i(t), \hat{q}^j(t)]. \] (3.92)

This coincides with one of the canonical commutation relations (3.89) in three dimensions.

The relativistic charges (3.73) of spacetime generate translations via

\[ \delta_s \hat{q}^a = e^a = -ie^b [\hat{p}_b(t), \hat{q}^a(\tau)], \] (3.93)

implying the relativistic commutation rules

\[ [\hat{p}_b(t), \hat{q}^a(\tau)] = i\delta^a_b, \] (3.94)

in agreement with the relativistic canonical commutation rules (1.157) (in natural units with \( \hbar = 1 \)).

Note that all commutation rules derived from the Noether charge according to the rule (3.86) hold for the operators in the Heisenberg picture, where they are time-dependent. The commutation rules in the purely algebraic discussion in Chapter 3, on the other hand, apply to the time-independent Schrödinger picture of the operators.

Similarly we find that the quantized versions of the conserved charges \( L_i \) in Eq. (3.55) generate infinitesimal rotations:

\[ \delta_s \hat{q}^i = -\omega^j i_{ijk} \hat{q}^k(t) = i\omega^j [\hat{L}_i, \hat{q}^i(t)], \] (3.95)
whereas the quantized conserved charges $N^i$ of Eq. (3.62) generate infinitesimal Galilei transformations, and that the charges $M^i$ of Eq. (3.83) generate pure Lorentz transformations:

$$
\delta_{s} \hat{q}^i = \epsilon_i \hat{q}^0 = i\epsilon_i [M_i, \hat{q}^0],
$$

$$
\delta_{s} \hat{q}^0 = \epsilon_i \hat{q}^i = i\epsilon_i [M_i, \hat{q}^0].
$$

(3.96)

Since the quantized charges generate the symmetry transformations, they form a representation of the generators of the Lorentz group. As such they must have the same commutation rules between each other as the generators of the symmetry group in Eq. (1.71) or their short version (1.72). This is indeed true, since the operator versions of the Noether charges (3.82) correspond to the operators (1.158) (in natural units).

3.5 Field Theory

A similar relation between continuous symmetries and constants of motion holds in field theories, where the role of the Lagrange coordinates is played by fields $q_x(t) = \varphi(x,t)$.

3.5.1 Continuous Symmetry and Conserved Currents

Let $A$ be the local action of an arbitrary field $\varphi(x) \rightarrow \varphi(x,t)$,

$$
A = \int d^4x \mathcal{L}(\varphi, \partial \varphi, x),
$$

(3.97)

and suppose that a transformation of the field

$$
\delta_s \varphi(x) = \epsilon \Delta(\varphi, \partial \varphi, x)
$$

(3.98)

changes the Lagrangian density $\mathcal{L}$ merely by a total derivative

$$
\delta_s \mathcal{L} = \epsilon \partial_a \Lambda^a,
$$

(3.99)

which makes the change of the action $A$ a surface integral, by Gauss’s divergence theorem:

$$
\delta_s A = \epsilon \int d^4x \partial_a \Lambda^a = \epsilon \int_S ds_a \Lambda^a,
$$

(3.100)

where $S$ is the surface of the total spacetime volume. Then $\delta_s \varphi$ is called a symmetry transformation.

Given such a transformation, we see that the four-dimensional current density

$$
j^a = \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \Delta - \Lambda^a
$$

(3.101)
Continuous Symmetries and Conservation Laws. Noether's Theorem

that has no divergence

$$\partial_a j^a(x) = 0. \quad (3.102)$$

The expression (3.101) is called a Noether current density and (3.102) is a local conservation law, just as in the electromagnetic equation (1.196).

We have seen in Eq. (1.198) that a local conservation law (3.102) always implies a global conservation law of the type (3.9) for the charge, which is now the Noether charge $Q(t)$ defined as in (1.199) by the spatial integral over the zeroth component (here in natural units with $c = 1$)

$$Q(t) = \int d^3 x \, j^0(x, t). \quad (3.103)$$

The proof of the local conservation law (3.102) is just as easy as for the mechanical action (3.1). We calculate the variation of $L$ under infinitesimal symmetry transformations (3.98) in a similar way as in Eq. (3.13), and find

$$\delta_s L = (\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial a} \frac{\partial L}{\partial \varphi}) \delta_s \varphi + \frac{\partial}{\partial a} \left( \frac{\partial L}{\partial \varphi} \delta_s \varphi \right)$$

$$= \epsilon \left( \frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial a} \frac{\partial L}{\partial \varphi} \right) \Delta + \epsilon \frac{\partial}{\partial a} \left( \frac{\partial L}{\partial \varphi} \Delta \right). \quad (3.104)$$

The Euler-Lagrange equation removes the first term and, equating the second term with (3.99), we obtain

$$\partial_a j^a \equiv \partial_a \left( \frac{\partial L}{\partial a \varphi} \Delta - \Lambda^a \right) = 0. \quad (3.105)$$

The relation between continuous symmetries and conservation is called Noether’s theorem [3].

### 3.5.2 Alternative Derivation

There is again an alternative derivative of the conserved current analogous to Eqs. (3.17)–(3.28). It is based on a variation of the fields under symmetry transformations whose parameter $\epsilon$ is made artificially spacetime-dependent $\epsilon(x)$, thus extending (3.17) to

$$\delta^\epsilon_s \varphi(x) = \epsilon(x) \Delta(\varphi(x), \varphi(x)). \quad (3.106)$$

As before in Eq. (3.19), let us calculate the Lagrangian density for a slightly transformed field

$$\varphi^\epsilon(x) \equiv \varphi(x) + \delta^\epsilon_s \varphi(x), \quad (3.107)$$

calling it

$$L^\epsilon(x) \equiv L(\varphi^\epsilon(x), \partial \varphi^\epsilon(x)). \quad (3.108)$$
The associated action differs from the original one by

$$\delta_{\epsilon}A = \int dx \left\{ \left[ \frac{\partial \mathcal{L}^{(x)}}{\partial \epsilon(x)} - \partial_a \left( \frac{\partial \mathcal{L}}{\partial \partial_a \epsilon(x)} \right) \right] \delta \epsilon(x) + \partial_a \left[ \frac{\partial \mathcal{L}^{(x)}}{\partial \partial_a \epsilon(x)} \delta \epsilon(x) \right] \right\}. \quad (3.109)$$

For classical fields $\varphi(x) = \varphi_c(x)$ satisfying the Euler-Lagrange equation (2.40), the extremality of the action implies the vanishing of the first term, and thus the Euler-Lagrange-like equation

$$\frac{\partial \mathcal{L}^{(x)}}{\partial \epsilon(x)} - \partial_a \frac{\partial \mathcal{L}^{(x)}}{\partial \partial_a \epsilon(x)} = 0. \quad (3.110)$$

By assumption, the action changes by a pure surface term under the $x$-independent transformation (3.106), implying that

$$\frac{\partial \mathcal{L}}{\partial \epsilon} = \partial_a \Lambda^a. \quad (3.111)$$

Inserting this into (3.110) we find that

$$j^a = \frac{\partial \mathcal{L}^{(x)}}{\partial \partial_a \epsilon(x)} - \Lambda^a \quad (3.112)$$

has no four-divergence. This coincides with the previous Noether current density (3.101), as can be seen by differentiating (3.108) with respect to $\partial_a \epsilon(x)$:

$$\frac{\partial \mathcal{L}^{(x)}}{\partial \partial_a \epsilon(x)} = \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \Delta(\varphi, \partial \varphi). \quad (3.113)$$

### 3.5.3 Local Symmetries

In Chapter 2 we observed that charged particles and fields coupled to electromagnetism possess a more general symmetry. They are invariant under local gauge transformations (2.103). The scalar Lagrangian (2.135), for example, is invariant under the gauge transformations (2.103) and (2.136), and the Dirac Lagrange density (2.150) under (2.103) and (2.151). These are all of the form (3.98), but with a parameter $\epsilon$ which depends on spacetime. Thus the action is invariant under local substantial variations of the type (3.106), which were introduced in the last section only as an auxiliary tool for an alternative derivation of the Noether current density (2.135). For a locally invariant Lagrangian density, the Noether expression (3.112) vanishes identically. This does not mean, however, that the system does not possess a conserved current, as we have seen in Eqs. (2.139) and (2.155). Only Noether’s derivation breaks down. Let us discuss this phenomenon in more detail for the Lagrangian density (2.150).

If we restrict the gauge transformations (2.151) to $x$-spacetime-independent gauge transformations

$$\psi(x) \rightarrow e^{ieA/c} \psi(x), \quad (3.114)$$
we can easily derive a conserved Noether current density of the type (3.101) for the Dirac field. The result is the known Dirac current density (2.153). It is the source of the electromagnetic field, with a minimal coupling between them. A similar structure exists for many internal symmetries giving rise to nonabelian versions of electromagnetism, which govern strong and weak interactions. What happens to Noether’s derivation of conservation laws in such theories.

As observed above, the formula (3.112) for the current density which was so useful in the globally invariant theory would yield a Noether current density

$$j_a = \frac{\delta L}{\partial \partial_a \Lambda}$$

which vanishes identically, due to local gauge invariance, Thus it would not provide us with a current density. We may, however, subject only the Dirac field to a local gauge transformation at fixed gauge fields. Then we obtain the conserved current

$$j_a = \frac{\partial L}{\partial \partial_a \Lambda} \bigg|_{\Lambda^*}.$$  

(3.116)

Since the complete change under local gauge transformations $\delta^* \mathcal{L}$ vanishes identically, we can alternatively vary only the gauge fields and keep the particle orbit fixed

$$j_a = -\frac{\partial L}{\partial \partial_a \Lambda} \bigg|_{\psi}.$$  

(3.117)

This is done most simply by forming the functional derivative with respect to the gauge field and omitting the contribution of $\psi^* \mathcal{L}$, i.e., by applying it only to the Lagrangian of the charge particles $\mathcal{L} \equiv \mathcal{L} - \psi^* \mathcal{L}$:

$$j^a = -\frac{\partial \mathcal{L}}{\partial \partial_a \Lambda} = -\frac{\partial \mathcal{L}}{\partial A_a}.$$  

(3.118)

As a check we apply the rule (3.118) to Dirac complex Klein-Gordon fields with the actions (2.140) and (2.27), and re-obtain the conserved current densities (2.153) and (2.138) (the extra factor $c$ is a convention). From the Schrödinger action (2.50) we derive the conserved charge current density

$$j(x, t) \equiv e \frac{i}{2m} \overline{\psi}^*(x, t) \nabla \psi(x, t) - \frac{e^2}{c} A \psi^*(x, t) \psi(x, t),$$  

(3.119)

to be compared with the particle current density (2.70) which satisfied the conservation law (2.71) together with the charge density $\rho(x, t) \equiv e \psi^*(x, t) \psi(x, t)$.

An important consequence of local gauge invariance can be found for the gauge field itself. If we form the variation of the pure gauge field action

$$\delta_s \mathcal{A} = \int d^4x \ \mathrm{tr} \left( \delta_s A_a \frac{\delta A^e_m}{\partial A_a} \right).$$  

(3.120)
and insert for $\delta_s^x A$ an infinitesimal pure gauge field configuration
\[ \delta_s^x A_a = - \partial_a \Lambda(x) \] (3.121)
the right-hand side must vanish for all $\Lambda(x)$. After a partial integration this implies
the local conservation law $\partial_a j^a(x) = 0$ for the Noether current
\[ j^a(x) = - \frac{\delta A}{\delta A_a}. \] (3.122)
Recalling the explicit form of the action in Eqs. (16.20) and (2.83), we find
\[ j^a(x) = - \partial_b F^{ab}. \] (3.123)
The Maxwell equation (2.86) can therefore be written as
\[ j^a(x) = - e_j^a(x), \] (3.124)
where we have emphasized the fact that the current $j^a$ contains only the fields of the
charge particles by a superscript $e$. In the form (3.124), the Maxwell equation implies
the vanishing of the total current density consisting of the sum of the conserved
current (3.117) of the charges and the Noether current (3.122) of the electromagnetic
field:
\[ j^a(x) = e_j^a(x) + j^a(x) = 0. \] (3.125)
This unconventional way of phrasing the Maxwell equation (2.86) will be useful for
understanding later the Einstein field equation (17.149) by analogy.

At this place we make an important observation. In contrast to the conservation
laws derived for matter fields, which are valid only if the matter fields obey the Euler-
Lagrange equations, the current conservation law for the Noether current (3.122) of
the gauge fields
\[ \partial_a j^a(x) = - \partial_a \partial_b F^{ab} = 0 \] (3.126)
is valid for all field configurations. The right-hand side vanishes identically since the
vector potential $A^a$ as an observable field in any fixed gauge satisfies the Schwarz
integrability condition (2.88).

## 3.6 Canonical Energy-Momentum Tensor

As an important example for the field-theoretic version of the Noether theorem
consider a Lagrangian density that does not depend explicitly on the spacetime
coordinates $x$:
\[ \mathcal{L}(x) = \mathcal{L}(\varphi(x), \partial \varphi(x)). \] (3.127)
We then perform a translation of the coordinates along an arbitrary direction \( b = 0, 1, 2, 3 \) of spacetime

\[
x'^a = x^a - \epsilon^a,
\]

under which field \( \varphi(x) \) transforms as

\[
\varphi'(x') = \varphi(x),
\]

so that

\[
\mathcal{L}'(x') = \mathcal{L}(x).
\]

If \( \epsilon^a \) is infinitesimally small, the field changes by

\[
\delta_s \varphi(x) = \varphi'(x) - \varphi(x) = \epsilon^b \partial_b \varphi(x),
\]

and the Lagrangian density by

\[
\delta_s \mathcal{L} \equiv \mathcal{L}(\varphi'(x), \partial \varphi'(x)) - \mathcal{L}(\varphi(x), \partial \varphi(x))
\]

\[= \frac{\partial \mathcal{L}}{\partial \varphi(x)} \delta_s \varphi(x) + \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_a \delta_s \varphi(x),
\]

which is a pure divergence term

\[
\delta_s \mathcal{L}(x) = \epsilon^b \partial_b \mathcal{L}(x).
\]

Hence the requirement (3.99) is satisfied and \( \delta_s \varphi(x) \) is a symmetry transformation, with a function \( \Lambda \) which happens to coincide with the Lagrangian density

\[
\Lambda = \mathcal{L}.
\]

We can now define four four-vectors of current densities \( j^a_b \), one for each component of \( \epsilon^b \). For the spacetime translation symmetry, they are denoted by \( \Theta^a_b \):

\[
\Theta^a_b = \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_b \varphi - \delta^a_b \mathcal{L}.
\]

Since \( \epsilon^b \) is a vector, this \( 4 \times 4 \) object is a tensor field, the so-called energy-momentum tensor of the scalar field \( \varphi(x) \). According to Noether’s theorem, this has no divergence in the index \( a \) [compare (3.102)]:

\[
\partial_a \Theta^a_b(x) = 0.
\]

The four conserved charges \( Q_b \) associated with these current densities [see the definition (3.103)]

\[
P_b = \int d^3 x \Theta^0_b(x),
\]
3.6 Canonical Energy-Momentum Tensor

are the components of the total four-momentum of the system.

The alternative derivation of this conservation law follows Subsection 3.1.4 by introducing the local variations

\[ \delta_s^x \varphi(x) = \epsilon^b(x) \partial_b \varphi(x) \quad (3.138) \]

under which the Lagrangian density changes by

\[ \delta_s^x \mathcal{L}(x) = \epsilon^b(x) \partial_b \mathcal{L}(x). \quad (3.139) \]

Applying the chain rule of differentiation we obtain, on the other hand,

\[ \delta_s^x \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi(x)} \epsilon^b(x) \partial_b \varphi(x) + \frac{\partial \mathcal{L}}{\partial \partial_a \varphi(x)} \left\{ [\partial_a \epsilon^b(x)] \partial_b \varphi + \epsilon^b \partial_a \partial_b \varphi(x) \right\}, \quad (3.140) \]

which shows that

\[ \frac{\partial \mathcal{L}}{\partial \partial_a \epsilon^b(x)} = \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_b \varphi. \quad (3.141) \]

Forming for each \( b \) the combination (3.101), we obtain again the conserved energy-momentum tensor (3.135).

Note that by analogy with (3.26), we can write (3.141) as

\[ \frac{\partial \mathcal{L}}{\partial \partial_a \epsilon^b(x)} = \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_b \varphi. \quad (3.142) \]

Note further that the component \( \Theta^0_0 \) of the canonical energy momentum tensor

\[ \Theta^0_0 = \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi} \partial_0 \varphi - \mathcal{L} \quad (3.143) \]

coincides with Hamiltonian density (2.61) derived in the canonical formalism by a Legendre transformation of the Lagrangian density.

3.6.1 Electromagnetism

As an important physical application of the field-theoretic Noether theorem, consider the free electromagnetic field with the action

\[ \mathcal{L} = -\frac{1}{4} F_{\text{cd}} F^{\text{cd}}, \quad (3.144) \]

where \( F_{\text{cd}} \) are the field strength \( F_{\text{cd}} \equiv \partial_c A_d - \partial_d A_c \). Under a translation of the spacetime coordinates from \( x^a \) to \( x^a - \epsilon^a \), the vector potential undergoes a similar change as the scalar field in (3.129):

\[ A'^a(x') = A^a(x). \quad (3.145) \]

For infinitesimal translations, this can be written as
\[ \delta_s A^c(x) \equiv A^c(x') - A^c(x) \]
\[ = A^c(x' + \epsilon) - A^c(x) \]
\[ = \epsilon^b \partial_b A^c(x). \quad (3.146) \]

Under this, the field tensor changes as follows

\[ \delta_s F^{cd} = \epsilon^b \partial_b F^{cd}, \quad (3.147) \]

and we verify that the Lagrangian density (3.144) is a total four divergence:

\[ \delta_s \mathcal{L} = -\epsilon^b \frac{1}{2} \left( \partial_b F_{cd} F^{cd} + F_{cd} \partial_b F^{cd} \right) = \epsilon^b \partial_b \mathcal{L}. \quad (3.148) \]

Thus, the spacetime translations (3.146) are symmetry transformations, and Eq. (3.100) yield the four Noether current densities, one for each \( \epsilon^b \):

\[ \Theta^a_b = \frac{1}{c} \left[ \frac{\partial \mathcal{L}}{\partial \partial_a A^c} \partial_b A^c - \delta^a_b \mathcal{L} \right]. \quad (3.149) \]

The factor \( 1/c \) is introduced to give the Noether current the dimension of the energy-momentum tensors introduced in Section 1.13, which are momentum densities. Here we have found the canonical energy-momentum tensor of the electromagnetic field, which satisfy the local conservation laws

\[ \partial_a \Theta^a_b (x) = 0. \quad (3.150) \]

Inserting the derivatives \( \partial \mathcal{L} / \partial \partial_a A^c = -F^a_c \), we obtain

\[ \Theta^a_b = \frac{1}{c} \left[ -F^a_c \partial_b A^c + \frac{1}{4} \delta^a_b F^{cd} F_{cd} \right]. \quad (3.151) \]

### 3.6.2 Dirac Field

We now turn to the Dirac field whose transformation law under spacetime translations

\[ x'^a = x^a - \epsilon^a \quad (3.152) \]

is

\[ \psi'(x') = \psi(x). \quad (3.153) \]

Since the Lagrangian density in (2.140) does not depend explicitly on \( x \) we calculate, as in (3.130):

\[ \mathcal{L}' (x') = \mathcal{L} (x). \quad (3.154) \]
The infinitesimal variations
\[ \delta_s \psi(x) = \epsilon^a \partial_a \psi(x). \]
(3.155)
produce the pure derivative term
\[ \delta_s D \mathcal{L} (x) = \epsilon^a \partial_a D \mathcal{L} (x), \]
(3.156)
and the combination (3.101) yields the Noether current densities
\[ \Theta^a_b = \frac{\partial D \mathcal{L}}{\partial \partial_a \psi^c} \partial_b \psi^c + cc - \delta^a_b D \mathcal{L}, \]
(3.157)
which satisfies local conservation laws
\[ \partial_a \Theta^a_b (x) = 0. \]
(3.158)
From (2.140) we see that
\[ \frac{\partial D \mathcal{L}}{\partial \partial_a \psi^c} = \frac{1}{2} \bar{\psi} \gamma^a \]
(3.159)
so that we obtain the canonical energy-momentum tensor of the Dirac field:
\[ \Theta^a_b = \frac{1}{2} \bar{\psi} \gamma^a \partial_b \psi^c + cc - \delta^a_b D \mathcal{L}. \]
(3.160)

3.7 Angular Momentum

Let us now turn to angular momentum in field theory. Consider first the case of a scalar field \( \phi(x) \). Under a rotation of the coordinates,
\[ x^{i'} = R^i_j x^j \]
(3.161)
the field does not change, if considered at the same space point, i.e.,
\[ \phi'(x^i) = \phi(x^i). \]
(3.162)
The infinitesimal substantial variation is:
\[ \delta_s \phi(x) = \phi'(x) - \phi(x). \]
(3.163)
For infinitesimal rotations (3.48),
\[ \delta_s x^i = -\phi_k \epsilon_{kij} x^j = -\omega_{ij} x^j, \]
(3.164)
we see that
\[ \delta_s \phi(x) = \phi'(x^0, x^{i'}, -\delta x^i) - \phi(x) = \partial_i \phi(x) x^i \omega_{ij}. \]
(3.165)
For a rotationally Lorentz-invariant Lagrangian density which has no explicit $x$-dependence:

$$\mathcal{L}(x) = \mathcal{L}(\varphi(x), \partial \varphi(x)),$$

(3.166)

the substantial variation is

$$\delta_s \mathcal{L}(x) = \mathcal{L}(\varphi'(x), \partial \varphi'(x)) - \mathcal{L}(\varphi(x), \partial \varphi(x))$$

$$= \frac{\partial \mathcal{L}}{\partial \varphi} \delta_s \varphi(x) + \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_a \delta_s \varphi(x).$$

(3.167)

Inserting (3.165), this becomes

$$\delta_s \mathcal{L} = \left[ \frac{\partial \mathcal{L}}{\partial \varphi} \partial_i \varphi x^j + \frac{\partial \mathcal{L}}{\partial \varphi} \partial_a \partial_i \varphi \delta_a \varphi \right] \omega_{ij}$$

$$= \left[ (\partial_i \mathcal{L}) x^j + \frac{\partial \mathcal{L}}{\partial \partial_i \varphi} \partial_j \varphi \right] \omega_{ij}.$$  

(3.168)

Since we are dealing with a rotation-invariant local Lagrangian density $\mathcal{L}(x)$ by assumption, the derivative $\partial \mathcal{L}/\partial \partial_a \varphi$ is a vector proportional to $\partial_a \varphi$. Hence the second term in the brackets is symmetric and vanishes upon contraction with the antisymmetric $\omega_{ij}$. This allows us to express $\delta_s \mathcal{L}$ as a pure derivative term

$$\delta_s \mathcal{L} = \partial_i \left( \mathcal{L} x^i \omega_{ij} \right).$$

(3.169)

Calculating $\delta_s \mathcal{L}$ once more using the Euler-Lagrange equations gives

$$\delta_s \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi} \partial_i \varphi x^j + \frac{\partial \mathcal{L}}{\partial \partial_a \partial_i \varphi} \partial_a \partial_i \varphi$$

$$= \partial_a \left( \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_i \varphi x^j \right) \omega_{ij}.$$  

(3.170)

Thus we find the Noether current densities (3.101):

$$L^{ij,a} = \left( \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} \partial_i \varphi x^j - \delta^a_i \mathcal{L} x^j \right) - (i \leftrightarrow j),$$

(3.171)

which have no four-divergence

$$\partial_a L^{ij,a} = 0.$$  

(3.172)

The current densities can be expressed in terms of the canonical energy-momentum tensor (3.135) as

$$L^{ij,a} = x^i \Theta^{ja} - x^j \Theta^{ia}.$$  

(3.173)

The associated Noether charges

$$L^{ij} = \int d^3 x L^{ij,a}$$

(3.174)

are the time-independent components of the total angular momentum of the field system.
3.8 Four-Dimensional Angular Momentum

Consider now pure Lorentz transformations (1.27). An infinitesimal boost to a rapidity $\zeta^i$ is described by a coordinate change [recall (1.34)]

$$x'^a = \Lambda^a{}_b x^b = x^a - \delta^a{}_0 \zeta^i x^i - \delta^a{}_i \zeta^i x^0.$$  \hfill (3.175)

This can be written as

$$\delta x^a = \omega^a{}_b x^b,$$  \hfill (3.176)

where for passive boosts

$$\omega_{ij} = 0, \quad \omega_{0i} = -\omega_{i0} = \zeta^i.$$  \hfill (3.177)

With the help of the tensor $\omega^a{}_b$, the boosts can be treated on the same footing with the passive rotations (1.36), for which (3.176) holds with

$$\omega_{ij} = \epsilon_{ijk} \varphi^k, \quad \omega_{0i} = \omega_{i0} = 0.$$  \hfill (3.178)

For both types of transformations, the substantial variations of the field are

$$\delta_s \varphi(x) = \varphi'(x'^a - \delta x^a) - \varphi(x) = -\partial^a \varphi(x) x^b \omega^a{}_b.$$  \hfill (3.179)

For a Lorentz-invariant Lagrangian density, the substantial variation can be shown, as in (3.169), to be a total derivative:

$$\delta_s \varphi = -\partial^a (L x^b) \omega^a{}_b,$$  \hfill (3.180)

and we obtain the Noether current densities

$$\mathcal{L}^{ab,c} = -\left( \frac{\partial \mathcal{L}}{\partial \varphi} x^b - \delta^{ac} \mathcal{L} x^b \right) - (a \leftrightarrow b).$$  \hfill (3.181)

The right-hand side can be expressed in terms of the canonical energy-momentum tensor (3.135), yielding

$$\mathcal{L}^{ab,c} = -\left( \frac{\partial \mathcal{L}}{\partial \varphi} x^b - \delta^{ac} \mathcal{L} x^b \right) - (a \leftrightarrow b) = x^a \Theta^{bc} - x^b \Theta^{ac}.$$  \hfill (3.182)

According to Noether’s theorem (3.102), these current densities have no four-divergence:

$$\partial_c \mathcal{L}^{ab,c} = 0.$$  \hfill (3.183)
The charges associated with these current densities:

\[ L_{ab} = \int d^3 x \, L_{ab}^0 \]  

are independent of time. For the particular form (3.177) of \( \omega_{ab} \), we recover the time independent components \( L_{ij} \) of angular momentum.

The time-independence of \( L_{i0} \) is the relativistic version of the center-of-mass theorem (3.66). Indeed, since

\[ L_{i0} = \int d^3 x \left( x^i \Theta^{00} - x^0 \Theta^{i0} \right), \]  

we can then define the relativistic center of mass

\[ x_{\text{CM}}^i = \frac{\int d^3 x \, x^i \Theta^{00}}{\int d^3 x \, \Theta^{00}} \]  

and the average velocity

\[ v_{\text{CM}}^i = c \frac{\int d^3 x \, \Theta^{i0}}{\int d^3 x \, \Theta^{00}} = c \frac{P^i}{P^0}. \]  

Since \( \int d^3 x \, \Theta^{i0} = P^i \) is the constant momentum of the system, also \( v_{\text{CM}}^i \) is a constant. Thus, the constancy of \( L_{i0} \) implies the center of mass to move with the constant velocity

\[ x_{\text{CM}}(t) = x_{\text{CM},0} + v_{\text{CM},0} t \]  

with \( x_{\text{CM},0} = L_{i0} / P^0 \).

The Noether charges \( L_{ab} \) are the four-dimensional angular momenta of the system.

It is important to point out that the vanishing divergence of \( L_{ab,c} \) makes \( \Theta^{ba} \) symmetric:

\[ \partial_c L_{ab,c} = \partial_c (x^a \Theta^{bc} - x^b \Theta^{ac}) = \Theta^{ba} - \Theta^{ba} = 0. \]  

Thus, field theories which are invariant under spacetime translations and Lorentz transformations must have a symmetric canonical energy-momentum tensor.

\[ \Theta^{ab} = \Theta^{ba} \]  

### 3.9 Spin Current

If the field \( \varphi(x) \) is no longer a scalar but has several spatial components, then the derivation of the four-dimensional angular momentum becomes slightly more involved.


3.9 Spin Current

3.9.1 Electromagnetic Fields

Consider first the case of electromagnetism where the relevant field is the four-vector potential \( A^a(x) \). When going to a new coordinate frame

\[
x'^a = \Lambda^a_b x^b
\]

(3.191)

the vector field at the same point remains unchanged in absolute spacetime. But since the components \( A^a \) refer to two different basic vectors in the different frames, they must be transformed simultaneously with \( x^a \). Since \( A^a(x) \) is a vector, it transforms as follows:

\[
A'^a(x') = \Lambda^a_b A^b(x).
\]

(3.192)

For an infinitesimal transformation

\[
\delta_s x^a = \omega^a_b x^b
\]

(3.193)

this implies the substantial variation

\[
\delta_s A^a(x) = A'^a(x) - A^a(x) = A^a(x - \delta x) - A^a(x)
\]

\[
= \omega^b_b A^b(x) - \omega^c_b x^b \partial_c A^a.
\]

(3.194)

The first term is a spin transformation, the other an orbital transformation. The orbital transformation can also be written in terms of the generators \( \hat{L}_{ab} \) of the Lorentz group defined in (3.82) as

\[
\delta_{\text{orb}} s A^a(x) = -i\omega^{bc} \hat{L}_{bc} A^a(x).
\]

(3.195)

The spin transformation of the vector field is conveniently rewritten with the help of the \( 4 \times 4 \) generators \( \hat{L}_{ab} \) in Eq. (1.51). Adding the two together, we form the operator of total four-dimensional angular momentum

\[
\hat{J}_{ab} \equiv 1 \times \hat{L}_{ab} + L_{ab} \times 1,
\]

(3.196)

and can write the transformation (3.194) as

\[
\delta_{s \text{orb}} A^a(x) = -i\omega^{ab} \hat{J}_{ab} A^a(x).
\]

(3.197)

If the Lagrangian density involves only scalar combinations of four-vectors \( A^a \) and if it has no explicit \( x \)-dependence, it changes under Lorentz transformations like a scalar field:

\[
\mathcal{L}'(x') \equiv \mathcal{L}(A'(x'), \partial'^a A'(x')) = \mathcal{L}(A(x), \partial A(x)) \equiv \mathcal{L}(x).
\]

(3.198)

Infinitesimally, this amounts to

\[
\delta_s \mathcal{L} = - (\partial_a \mathcal{L} x^b) \omega^a_b.
\]

(3.199)
With the Lorentz transformations being symmetry transformations in the Noether sense, we calculate as in (3.171) the current of total four-dimensional angular momentum:

\[
J^{ab,c} = \frac{1}{c} \left[ \frac{\partial L}{\partial A^a} A_b - \left( \frac{\partial L}{\partial \partial_c A^d} \partial^a A^d x^b - \delta^{ac} \mathcal{L} x^b \right) - (a \leftrightarrow b) \right].
\]  

(3.200)

The prefactor \(1/c\) is chosen to give these Noether currents of the electromagnetic field the conventional physical dimension. In fact, the last two terms have the same form as the current \(L^{ab,c}\) of the four-dimensional angular momentum of the scalar field. Here they are the corresponding quantities for the vector potential \(A^a(x)\):

\[
L^{ab,c} = -\frac{1}{c} \left( \frac{\partial L}{\partial A^d} \partial^a A^d x^b - \delta^{ac} \mathcal{L} x^b \right) + (a \leftrightarrow b).
\]  

(3.201)

Note that this current has the form

\[
L^{ab,c} = \frac{1}{c} \left\{ -i \frac{\partial L}{\partial \partial_c A^d} \hat{L}^{ab} A^d + \left[ \delta^{ac} \mathcal{L} x^b - (a \leftrightarrow b) \right] \right\},
\]  

(3.202)

where \(\hat{L}^{ab}\) are the differential operators of four-dimensional angular momentum (1.103) satisfying the commutation rules (1.71) and (1.72).

Just as the scalar case (3.182), the currents (3.201) can be expressed in terms of the canonical energy-momentum tensor as

\[
L^{ab,c} = x^a \Theta^{bc} - x^b \Theta^{ac}.
\]  

(3.203)

The first term in (3.200),

\[
\Sigma^{ab,c} = \frac{1}{c} \left[ \frac{\partial L}{\partial \partial_c A^b} A^b - (a \leftrightarrow b) \right],
\]  

(3.204)

is referred to as the **spin current**. It can be written in terms of the \(4 \times 4\)-generators (1.51) of the Lorentz group as

\[
\Sigma^{ab,c} = -\frac{i}{c} \frac{\partial L}{\partial \partial_c A^d} (L^{ab})_{d\sigma} A^\sigma.
\]  

(3.205)

The two currents together

\[
J^{ab,c}(x) \equiv L^{ab,c}(x) + \Sigma^{ab,c}(x)
\]  

(3.206)

are conserved, \(\partial_c J^{ab,c}(x) = 0\). Individually, the terms are not conserved.

The total angular momentum is given by the charge

\[
J^{ab} = \int d^3x \ J^{ab,0}(x).
\]  

(3.207)
It is a constant of motion. Using the conservation law of the energy-momentum tensor we find, just as in (3.189), that the orbital angular momentum satisfies

$$\partial_c L^{ab,c}(x) = - \left[ \Theta^{ab}(x) - \Theta^{ba}(x) \right].$$  \hspace{1cm} (3.208)

From this we find the divergence of the spin current

$$\partial_c \Sigma^{ab,c}(x) = - \left[ \Theta^{ab}(x) - \Theta^{ba}(x) \right].$$  \hspace{1cm} (3.209)

For the charges associated with orbital and spin currents

$$L^{ab}(t) \equiv \int d^3 x L^{ab,0}(x), \quad \Sigma^{ab}(t) \equiv \int d^3 x \Sigma^{ab,0}(x),$$

this implies the following time dependence:

$$\dot{L}^{ab}(t) = - \int d^3 x \left[ \Theta^{ab}(x) - \Theta^{ba}(x) \right],$$  \hspace{1cm} (3.210)

$$\dot{\Sigma}^{ab}(t) = \int d^3 x \left[ \Theta^{ab}(x) - \Theta^{ba}(x) \right].$$  \hspace{1cm} (3.211)

Fields with spin have always have a non-symmetric energy momentum tensor. Then the current $J^{ab,c}$ becomes, now back in natural units,

$$J^{ab,c} = \left( \frac{\partial \delta^x_s \mathcal{L}}{\partial \omega^{ab}(x)} - \delta^{ac} \mathcal{L} x^b \right) - (a \leftrightarrow b)$$ \hspace{1cm} (3.212)

By the chain rule of differentiation, the derivative with respect to $\partial, \omega^{ab}(x)$ can come only from field derivatives. For a scalar field

$$\frac{\partial \delta^x_s \mathcal{L}}{\partial \omega^{ab}(x)} = \frac{\partial \mathcal{L}}{\partial \varphi} \left( \frac{\partial \delta^x_s \varphi}{\partial \omega^{ab}(x)} \right),$$  \hspace{1cm} (3.213)

and for a vector field

$$\frac{\partial \delta^x_s \mathcal{L}}{\partial \omega^{ab}(x)} = \frac{\partial \mathcal{L}}{\partial A^d} \left( \frac{\partial \delta^x_s A^d}{\partial \omega^{ab}} \right).$$  \hspace{1cm} (3.214)

The alternative rule of calculating angular momenta is to introduce spacetime-dependent transformations

$$\delta^x x = \omega^a_b(x)x^b$$

under which the scalar fields transform as

$$\delta^x \varphi = -\partial^x \omega^c_b(x)x^b$$  \hspace{1cm} (3.215)

and the Lagrangian density as

$$\delta^x \mathcal{L} = -\partial^x \mathcal{L} \omega^c_b(x)x^b = -\partial^x (x^b \mathcal{L}) \omega^c_b(x).$$  \hspace{1cm} (3.216)

By separating spin and orbital transformations of $\delta^x_s A^d$ we find the two contributions $\sigma^{ab,c}$ and $L^{ab,c}$ to the current $J^{ab,c}$ of the total angular momentum, the latter receiving a contribution from the second term in (3.212).
3.9.2 Dirac Field

We now turn to the Dirac field. Under a Lorentz transformation (3.191), this transforms according to the law
\[ \psi(x') \xrightarrow{\Lambda} \psi_\Lambda'(x) = D(\Lambda)\psi(x), \]
where \( D(\Lambda) \) are the 4 \( \times \) 4 spinor representation matrices of the Lorentz group. Their matrix elements can most easily be specified for infinitesimal transformations. For an infinitesimal Lorentz transformation
\[ \Lambda_a^b = \delta_a^b + \omega_a^b, \]
under which the coordinates are changed by
\[ \delta s x^a = \omega_a^b x^b, \]
the spin transforms under the representation matrix
\[ D(\delta_a^b + \omega_a^b) = \left( 1 - i \frac{1}{2} \omega_a^b \sigma_a^b \right) \psi(x), \]
where \( \sigma_a^b \) are the 4 \( \times \) 4 matrices acting on the spinor space defined in Eq. (1.222). We have shown in (1.220) that the spin matrices \( \Sigma_a^b \equiv \sigma_a^b/2 \) satisfy the same commutation rules (1.71) and (1.72) as the previous orbital and spin-1 generators \( \hat{L}_{aba} \) and \( L_{ab} \) of Lorentz transformations.

The field has the substantial variation [compare (3.194)]:
\[
\delta_s \psi(x) = \psi'(x) - \psi(x) = D(\delta_a^b + \omega_a^b)\psi(x - \delta x) - \psi(x) = -i \frac{1}{2} \omega_a^b \sigma_a^b \psi(x) - \omega_a^b x^b \partial_a \psi(x) = -i \frac{1}{2} \omega_a^b \left( S_a^b + \hat{L}_{ab} \right) \psi(x) \equiv -i \frac{1}{2} \omega_a^b \hat{J}_{ab} \psi(x), \]
the last line showing the separation into spin and orbital transformation for a Dirac particle.

Since the Dirac Lagrangian is Lorentz-invariant, it changes under Lorentz transformations like a scalar field:
\[ \mathcal{L}'(x') = \mathcal{L}(x). \]
Infinitesimally, this amounts to
\[ \delta_s \mathcal{L} = - (\partial_a \mathcal{L} x^b) \omega_a^b. \]

With the Lorentz transformations being symmetry transformations in the Noether sense, we calculate the current of total four-dimensional angular momentum extending the formulas (3.182) and (3.200) for scalar field and vector potential. The result is
\[
J^{ab,c} = \left( -i \frac{\partial \mathcal{L}}{\partial \psi^a} \sigma^{ab} \psi - i \frac{\partial \mathcal{L}}{\partial \psi^b} \hat{L}_{ab} \psi + cc \right) + \left[ \delta^{ac} \mathcal{L} x^b - (a \leftrightarrow b) \right]. \]
As before in (3.201) and (3.182), the orbital part of (3.225) can be expressed in terms of the canonical energy-momentum tensor as

\[ L_{ab,c} = x^a \Theta^{bc} - x^b \Theta^{ac}. \]  

(3.226)

The first term in (3.225) is the spin current

\[ \Sigma^{ab,c} = \frac{1}{2} \left( -i \frac{\partial L}{\partial \partial_c \psi} \sigma^{ab} \psi + cc \right). \]  

(3.227)

Inserting (3.159), this becomes explicitly

\[ \Sigma^{ab,c} = -\frac{i}{2} \bar{\psi} \gamma^c \sigma^{ab} \psi = \frac{1}{2} \bar{\psi} \gamma^{[a} \gamma^{b} \gamma^c] \psi = \frac{1}{2} \epsilon^{abcd} \bar{\psi} \gamma^d \psi. \]  

(3.228)

The spin density is completely antisymmetric in the three indices\(^1\).

Due to the presence of spin, the energy-momentum tensor is not symmetric.

## 3.10 Symmetric Energy-Momentum Tensor

Since the presence of spin is the cause for the asymmetry of the canonical energy-momentum tensor, it is suggestive that by an appropriate use of the spin current it should be possible to construct a new modified momentum tensor

\[ T^{ab} = \Theta^{ab} + \Delta \Theta^{ba} \]  

(3.229)

which is symmetric, while still having the fundamental property of \( \Theta^{ab} \), that the integral \( P^a = \int d^3x T^{a0} \) is the total energy-momentum vector of the system. This is the case if \( \Delta \Theta^{ab} \) is a three-divergence of a spatial vector. Such a construction was found by Belinfante in 1939. He introduced the tensor \([4]\)

\[ T^{ab} = \Theta^{ab} - \frac{1}{2} \partial_c (\Sigma^{ab,c} - \Sigma^{bc,a} + \Sigma^{ca,b}), \]  

(3.230)

whose symmetry is manifest, due to (3.209) and the symmetry of the last two terms under \( a \leftrightarrow b \). Moreover, by the components

\[ T^{a0} = \Theta^{a0} - \frac{1}{2} \partial_c (\Sigma^{a0,c} - \Sigma^{0c,a} + \Sigma^{ca,0}) = x^a T^{bc} - x^b T^{ac} \]  

(3.231)

which gives the same total angular momentum as the canonical expression (3.206):

\[ J^{ab} = \int d^3x J^{ab,0}. \]  

(3.232)

\(^1\)This property is important for being able to construct a consistent quantum mechanics in a space with torsion. See Ref. [2].
Indeed, the zeroth component of (3.231) is
\[ x^a \Theta^{b0} - x^b \Theta^{a0} - \frac{1}{2} \left[ \partial_k (\Sigma^{0,k} - \Sigma^{0,k,a} + \Sigma^{k,0}) x^b - (a \leftrightarrow b) \right]. \] (3.233)

Integrating the second term over \( d^3x \) and performing a partial integration gives, for \( a = 0, b = i \):
\[ -\frac{1}{2} \int d^3x \left[ x^0 \partial_k (\Sigma^{0,k} - \Sigma^{0,k,i} + \Sigma^{k,i,0}) - x^i \partial_k (\Sigma^{i,0} - \Sigma^{0,i} + \Sigma^{k,0}) \right] = \int d^3x \Sigma^{0,i,0}, \] (3.234)
and for \( a = i, b = j \)
\[ -\frac{1}{2} \int d^3x \left[ x^i \partial_k (\Sigma^{j,0,k} - \Sigma^{0,k,j} + \Sigma^{k,j,0}) - (i \leftrightarrow j) \right] = \int d^3x \Sigma^{ij,0}. \] (3.235)
The right-hand sides are the contributions of the spin to the total angular momentum.

For the electromagnetic field, the spin current (3.204) reads explicitly
\[ \Sigma^{ab,c} = \frac{-1}{c} \left[ F^{ca} A_b - (a \leftrightarrow b) \right]. \] (3.236)

From this we calculate the Belinfante correction
\[ \Delta \Theta^{ab} = \frac{1}{2c} \left[ \partial_c (F^{ca} A_b - F^{cb} A^a) - \partial_c (F^{ab} A^c - F^{ac} A^b) + \partial_c (F^{bc} A^a - F^{ba} A^c) \right] \]
\[ = \frac{1}{c} \partial_c (F^{bc} A^a). \] (3.237)

Adding this to the canonical energy-momentum tensor (3.151)
\[ \Theta^{ab} = \frac{1}{c} \left[ -F^{bc} \partial_a A^c + \frac{1}{4} g^{ab} F^{cd} F_{cd} \right], \] (3.238)
we find the symmetric energy-momentum tensor
\[ T^{ab} = \frac{1}{c} \left[ -F^{bc} F_{ac} + \frac{1}{4} g^{ab} F^{cd} F_{cd} + (\partial_c F^{bc}) A^a \right]. \] (3.239)
The last term vanishes for a free Maxwell field which satisfies \( \partial_c F^{ab} = 0 \) [recall (2.86)], and can be dropped. In this case \( T^{ab} \) agrees with the previously constructed symmetric energy-momentum tensor (1.272) of the electromagnetic field. The symmetry of \( T^{ab} \) can easily be verified using once more the Maxwell equation \( \partial_c F^{ab} = 0 \).

We have seen in (1.269) that the component \( c T^{00}(x) \) agrees with the well-known energy density \( \mathcal{E}(x) = \left( \mathbf{E}^2 + \mathbf{B}^2 \right) / 2 \) of the electromagnetic field, and that the components \( c^2 T^{ab}(x) \) are equal to the *Poynting vector* of energy current density \( \mathbf{S}(x) = c \mathbf{E} \times \mathbf{B} \), so that the energy conservation law \( c^2 \partial_a T^{0a}(0) \) can be written as \( \partial_t \mathcal{E}(x) + \nabla \cdot \mathbf{S}(x) = 0 \).
In the presence of an external current, where the Lagrangian density is (2.83), the canonical energy-momentum tensor becomes

$$\Theta^{ab} = \frac{1}{c} \left[ -F^b_c \partial^a A^c + \frac{1}{4} g^{ab} F^{cd} F_{cd} + \frac{1}{c} g^{ab} j^c A^c \right],$$  \hspace{1cm} (3.240)

generalizing (3.238).

The spin current is again given by Eq. (3.236), leading to the Belinfante energy-momentum tensor

$$T^{ab} = \Theta^{ab} + \frac{1}{c} \partial_c (F^{bc} A^a)$$

$$= \frac{1}{c} \left[ -F^b_c F^{ac} + \frac{1}{4} g^{ab} F^{cd} F_{cd} + \frac{1}{c} g^{ab} j^c A^c - \frac{1}{c} j^b A^a \right].$$  \hspace{1cm} (3.241)

The last term prevents $T^{ab}$ from being symmetric, unless the current vanishes. Due to the external current, the conservation law $\partial_b T^{ab} = 0$ is modified to

$$\partial_b T^{ab} = \frac{1}{c} A_c(x) \partial^a j^c(x).$$  \hspace{1cm} (3.242)

### 3.11 Internal Symmetries

In quantum field theory, an important role in classifying various actions is played by internal symmetries. They do not involve any change in the spacetime coordinate of the fields, i.e., they have the form

$$\phi'(x) = e^{-i \alpha_r G_r} \phi(x)$$  \hspace{1cm} (3.243)

where $G_r$ are the generators of some Lie group, and $\alpha_r$ the associated transformation parameters. If the field has $N$ components, the generators $G_r$ are $N \times N$-matrices. They satisfy commutation rules of the general form [recall (1.65)]

$$[G_r, G_s] = i f_{rst} G_t, \hspace{1cm} (r, s, t = 1, \ldots, \text{rank}),$$  \hspace{1cm} (3.244)

where $f_{rst}$ are the structure constants of the Lie algebra.

The infinitesimal symmetry transformations are substantial variations of the form

$$\delta_s \varphi = -i \alpha_r G_r \varphi,$$  \hspace{1cm} (3.245)

The associated conserved current densities read

$$j_r^a = -i \frac{\partial \mathcal{L}}{\partial \partial_a \varphi} G_r \varphi.$$  \hspace{1cm} (3.246)

These can also be written as

$$j_r^a = -i \pi G_r \varphi,$$  \hspace{1cm} (3.247)
where $\pi(x) \equiv \partial L(x)/\partial \partial_a \varphi(x)$ is the canonical momentum of the field $\varphi(x)$ [compare (2.60)].

The most important example is that of a complex field $\phi$ and a generator $G = 1$, where the symmetry transformation (3.243) is simply a multiplication by a constant phase factor. One also speaks of U(1)-symmetry. Other important examples are those of a triplet or an octet of fields $\phi$ with $G_r$ being the generators of an SU(2) or SU(3) representation. The U(1)-symmetry leads to charge conservation in electromagnetic interactions, the other two are responsible for isospin SU(2) and SU(3) invariance in strong interactions. The latter symmetries are, however, not exact.

### 3.11.1 U(1)-Symmetry and Charge Conservation

Consider a Lagrangian density $L(x) = L(\varphi(x), \partial \varphi(x), x)$ that is invariant under U(1)-transformations

$$\delta_s \varphi(x) = -i \alpha \varphi(x) \quad (3.248)$$

i.e., $\delta_s L = 0$. By the chain rule of differentiation we find, using the Euler-Lagrange equation (2.40)

$$\delta_s L = \left( \frac{\partial L}{\partial \varphi} - \frac{d}{dt} \frac{\partial L}{\partial \partial_a \varphi} \right) \delta_s \varphi + \left[ \frac{\partial L}{\partial \partial_a \varphi} \delta_s \varphi \right] = 0. \quad (3.249)$$

Inserting (3.248), we find that

$$j_\mu = -\frac{\partial L}{\partial \partial_\mu \varphi} \varphi$$

is a conserved current.

For a free relativistic complex scalar field with a Lagrangian density

$$L(x) = \partial_\mu \varphi^* \partial_\mu \varphi - m^2 \varphi^* \varphi \quad (3.251)$$

we have to add the contributions of real and imaginary parts of the field $\phi$ in formula (3.250), and obtain the conserved current density

$$j_\mu = -i \varphi^* \overleftarrow{\partial}_\mu \varphi$$

where the symbol $\varphi^* \overleftarrow{\partial}_\mu \varphi$ denotes the left-minus-right derivative:

$$\varphi^* \overleftarrow{\partial} \varphi \equiv \varphi^* \partial_\mu \varphi - (\partial_\mu \varphi^*) \varphi. \quad (3.253)$$

For a free Dirac field, the current density (3.250) takes the form

$$j^\mu(x) = \bar{\psi}(x) \gamma^\mu \psi(x). \quad (3.254)$$
3.12 Generating the Symmetry Transformations on Quantum Fields

As in quantum mechanical systems, the charges associated with the conserved currents obtained in the previous section can be used to generate the transformations of the fields from which they were derived. One merely has to invoke the canonical field commutation rules.

For the currents (3.246), the charges are
\[ Q_r = -i \int d^3x \frac{\partial \mathcal{L}}{\partial a_r} G_r \varphi \]  
(3.257)
and can be written as
\[ Q_r = -i \int d^3x \pi G_r \varphi, \]  
(3.258)
where \( \pi(x) \equiv \frac{\partial \mathcal{L}(x)}{\partial a_r} \varphi(x) \) is the canonical momentum of the field \( \varphi(x) \). After quantization, these fields satisfy the canonical commutation rules:
\[ [\hat{\pi}(x, t), \hat{\varphi}(x', t)] = -i\delta^{(3)}(x - x'), \]  
\[ [\hat{\varphi}(x, t), \hat{\varphi}(x', t)] = 0, \]  
(3.259)
\[ [\hat{\pi}(x, t), \hat{\pi}(x', t)] = 0. \]

From this we derive directly the commutation rule between the quantized version of the charges (3.258) and the field operator \( \hat{\varphi}(x) \):
\[ [\hat{Q}_r, \hat{\varphi}(x)] = -\alpha_r G_r \varphi(x). \]  
(3.260)

We also find that the commutation rules among the quantized charges \( \hat{Q}_r \) are the same as those of the generators \( G_r \) in (3.244):
\[ [\hat{Q}_r, \hat{Q}_s] = f_{rst} \hat{Q}_t, \quad (r, s, t = 1, \ldots, \text{rank}). \]  
(3.261)
Hence the operators $\hat{Q}_r$ form a representation of the generators of symmetry group in the many-particle Hilbert space generated by the quantized fields $\hat{\varphi}(x)$ (Fock space).

As an example, we may derive in this way the commutation rules of the conserved charges associated with the Lorentz generators (3.226):

$$J^{ab} \equiv \int d^3x J^{ab,0}(x).$$  \(3.262\)

They are obviously the same as those of the $4 \times 4$-matrices (1.51), and those of the quantum mechanical generators (1.103):

$$[\hat{j}^{ab}, \hat{j}^{ac}] = -ig^{aa} \hat{j}^{bc}. \tag{3.263}$$

The generators $J^{ab} \equiv \int d^3x J^{ab,0}(x)$, are sums $J^{ab} = L^{ab}(t) + \Sigma^{ab}(t)$ of charges (3.210) associated with orbital and spin rotations. According to (3.211), these individual charges are time dependent, only their sum being conserved. Nevertheless, they both generate Lorentz transformations: $L^{ab}(t)$ on the spacetime argument of the fields, and $\Sigma^{ab}(t)$ on the spin indices. As a consequence, they both satisfy the commutation relations (3.263):

$$[\hat{L}^{ab}, \hat{L}^{ac}] = -ig^{aa} \hat{L}^{bc}, \quad [\hat{\Sigma}^{ab}, \hat{\Sigma}^{ac}] = -ig^{aa} \hat{\Sigma}^{bc}. \tag{3.264}$$

It is important to realize that the commutation relations (3.260) and (3.261) remain valid also in the presence of symmetry-breaking terms as long as these do not contribute to the canonical momentum of the theory. Such terms are called soft symmetry-breaking terms. The charges are no longer conserved, so that we must attach a time argument to the commutation relations (3.260) and (3.261). All times in these relations must be the same, in order to invoke the equal-time canonical commutation rules.

The commutators (3.261) have played an important role in developing a theory of strong interactions, where they first appeared in the form of a charge algebra of the broken symmetry $SU(3) \times SU(3)$ of weak and electromagnetic charges. This symmetry will be discussed in more detail in Chapter 10.

### 3.13 Energy-Momentum Tensor of Relativistic Massive Point Particle

If we want to study energy and momentum of charged relativistic point particles in an electromagnetic field it is useful to consider the action (3.68) with (3.70) as an integral over a Lagrangian density:

$$A = \int d^4x \mathcal{L}(x), \quad \text{with} \quad \mathcal{L}(x) = \int_{\tau_a}^{\tau_b} d\tau L(\dot{x}^a(\tau)) \delta^{(4)}(x - x(\tau)). \tag{3.265}$$

This allows us to derive for point particles local conservation laws in the same way as for fields. Instead of doing this, however, we shall take advantage of the
3.13 Energy-Momentum Tensor of Relativistic Massive Point Particle

previously derived global conservation laws and convert them into local ones by inserting appropriate δ-functions with the help of the trivial identity

\[ \int d^4x \delta^{(4)}(x - x(\tau)) = 1. \]  

(3.266)

Consider for example the conservation law (3.72) for the momentum (3.73). With the help of (3.266) this becomes

\[ 0 = -\int d^4x \int_{-\infty}^{\infty} d\tau \left[ \frac{d}{d\tau} p^c(\tau) \right] \delta^{(4)}(x - x(\tau)). \]  

(3.267)

Note that in this expression the boundaries of the four-volume contain the information on initial and final times. We then perform a partial integration in \( \tau \), and rewrite (3.267) as

\[ 0 = -\int d^4x \int_{-\infty}^{\infty} d\tau \frac{d}{d\tau} [p^c(\tau) \delta^{(4)}(x - x(\tau))] + \int d^4x \int_{-\infty}^{\infty} d\tau p^c(\tau) \partial_\tau \delta^{(4)}(x - x(\tau)). \]  

(3.268)

The first term vanishes if the orbits come from and disappear into infinity. The second term can be rewritten as

\[ 0 = -\int d^4x \partial^b \left[ \int_{-\infty}^{\infty} d\tau p^c(\tau) \dot{x}^b(\tau) \delta^{(4)}(x - x(\tau)) \right]. \]  

(3.269)

This shows that

\[ \Theta^{cb}(x) \equiv m \int_{-\infty}^{\infty} d\tau \dot{x}^c(\tau) \dot{x}^b(\tau) \delta^{(4)}(x - x(\tau)) \]  

(3.270)

satisfies the local conservation law

\[ \partial^b \Theta^{cb}(x) = 0. \]  

(3.271)

This is the conservation law for the energy-momentum tensor of a massive point particle.

The total momenta are obtained from the spatial integrals over \( \Theta^{c0} \):

\[ P^a(t) \equiv \int d^3x \Theta^{c0}(x). \]  

(3.272)

For point particles, they coincide with the canonical momenta \( p^a(t) \). If the Lagrangian depends only on the velocity \( \dot{x}^a \) and not on the position \( x^a(t) \), the momenta \( p^a(t) \) are constants of motion: \( p^a(t) \equiv p^a \).

The Lorentz invariant quantity

\[ M^2 = P^2 = g_{ab}P^aP^b \]  

(3.273)

is called the total mass of the system. For a single particle it coincides with the mass of the particle.
Subjecting the orbits $x^a(\tau)$ to Lorentz transformations according to the rules of the last section we find the currents of total angular momentum

$$L^{ab,c} \equiv x^a \Theta^{bc} - x^b \Theta^{ac},$$  \hspace{1cm} (3.274)

to satisfy the conservation law:

$$\partial_c L^{ab,c} = 0.$$  \hspace{1cm} (3.275)

A spatial integral over the zeroth component of the current $L^{ab,c}$ yields the conserved charges:

$$L^{ab}(t) \equiv \int d^3 x \, L^{ab,0}(x) = x^a p^b(t) - x^b p^a(t).$$  \hspace{1cm} (3.276)

### 3.14 Energy-Momentum Tensor of Massive Charged Particle in Electromagnetic Field

Let us also consider an important combination of a charged point particle and an electromagnetic field Lagrangian

$$A = -mc \int_{\tau_a}^{\tau_b} d\tau \sqrt{g_{ab} \dot{x}^a(\tau) \dot{x}^b(\tau)} - \frac{1}{4} \int d^4 x F_{ab} F^{ab} - \frac{e}{c} \int_{\tau_a}^{\tau_b} d\tau \dot{x}^a(\tau) A_a(x(\tau)).$$  \hspace{1cm} (3.277)

By varying the action in the particle orbits, we obtain the Lorentz equation of motion

$$\frac{dp^a}{d\tau} = \frac{e}{c} F^{a}_{\ b} \dot{x}^b(\tau).$$  \hspace{1cm} (3.278)

By varying the action in the vector potential, we find the Maxwell-Lorentz equation

$$-\partial_b F^{ab} = \frac{e}{c} \dot{x}^b(\tau).$$  \hspace{1cm} (3.279)

The action (3.277) is invariant under translations of the particle orbits and the electromagnetic fields. The first term is obviously invariant, since it depends only on the derivatives of the orbital variables $x^a(\tau)$. The second term changes under translations by a pure divergence [recall (3.133)]. Also the interaction term changes by a pure divergence, which is seen as follows: Since the substantial variation changes $x^b(\tau) \to x^b(\tau) - \epsilon^b$, under which $\dot{x}^a(\tau)$ is invariant,

$$\dot{x}^a(\tau) \to \dot{x}^a(\tau),$$  \hspace{1cm} (3.280)

and $A_a(x^b)$ changes as follows:

$$A_a(x^b) \to A'_a(x^b) = A_a(x^b + \epsilon^b) = A_a(x^b) + \epsilon^b \partial_a A_a(x^b).$$  \hspace{1cm} (3.281)

Altogether we obtain

$$\delta_a \mathcal{L} = \epsilon^b \partial_b \mathcal{L}.$$  \hspace{1cm} (3.282)
We now calculate the same variation once more invoking the equations of motion. This gives

$$\delta_s A = \int d\tau \frac{d}{d\tau} L^m_{\nu} \delta_s x^a + \int d^4x \frac{\partial L_{em}}{\partial \dot{x}^a} \delta_s A^a. \quad (3.283)$$

The first term can be treated as in (3.268)–(3.269) after which it acquires the form

$$- \int_{\tau}^{\tau_b} d\tau \frac{d}{d\tau} \left( p^a + \frac{e}{c} A_a \right) = - \int d^4x \int_{-\infty}^{\infty} d\tau \frac{d}{d\tau} \left[ \left( p^a + \frac{e}{c} A_a \right) \delta^{(4)}(x-x(\tau)) \right] \quad (3.284)$$

and thus, after dropping boundary terms,

$$- \int_{\tau}^{\tau_b} d\tau \frac{d}{d\tau} \left( p^a + \frac{e}{c} A_a \right) = - \int d^4x \int_{-\infty}^{\infty} d\tau \left( p^a + \frac{e}{c} A_a \right) \frac{dx}{dt} \delta^{(4)}(x-x(\tau)). \quad (3.285)$$

The electromagnetic part is the same as before, since the interaction contains no derivative of the gauge field. In this way we find the canonical energy-momentum tensor

$$\Theta^{ab}(x) = \int d\tau \left( p^a + \frac{e}{c} A^a \right) \dot{x}^b(\tau) \delta^{(4)}(x-x(\tau)) - F^b c \partial^a A^c + \frac{1}{4} g^{ab} F_{cd} F_{cd}. \quad (3.286)$$

Let us check its conservation by calculating the divergence:

$$\partial_b \Theta^{ab}(x) = \int d\tau \left( p^a + \frac{e}{c} A^a \right) \dot{x}^b(\tau) \partial_b \delta^{(4)}(x-x(\tau))$$

$$- \partial_b F^b c \partial^a A^c - F^b c \partial_b \delta^{(4)}(x-x(\tau)) \quad (3.287)$$

The first term is, up to a boundary term, equal to

$$- \int d\tau \left( p^a + \frac{e}{c} A^a \right) \frac{dx}{d\tau} \delta^{(4)}(x-x(\tau)) = \int d\tau \left[ \frac{dx}{d\tau} \left( p^a + \frac{e}{c} A^a \right) \right] \delta^{(4)}(x-x(\tau)). \quad (3.288)$$

Using the Lorentz equation of motion (3.278), this becomes

$$\frac{e}{c} \int_{-\infty}^{\infty} d\tau \left( F^a_b \dot{x}^b(\tau) + \frac{d}{d\tau} A^a(\tau) \right) \delta^{(4)}(x-x(\tau)). \quad (3.289)$$

Inserting the Maxwell equation

$$\partial_b F^{ab} = - \epsilon \int d\tau (dx^a/d\tau) \delta^{(4)}(x-x(\tau)), \quad (3.290)$$

the second term in Eq. (3.287) can be rewritten as

$$- \frac{e}{c} \int_{-\infty}^{\infty} d\tau \frac{dx}{d\tau} \delta^{(4)}(x-x(\tau)), \quad (3.291)$$
which is the same as

\[ -\frac{e}{c} \int d\tau \left( \frac{dx_a}{d\tau} F^{ac} + \frac{dx_c}{d\tau} \partial^c A^a \right) \delta^{(4)}(x - x(\tau)), \]

(3.292)

thus canceling (3.289). The third term in (3.287) is, finally, equal to

\[ -F^b_c \partial^a F^c_b + \frac{1}{4} \partial^a (F^{cd} F_{cd}), \]

(3.293)

due to the antisymmetry of $F^{bc}$. By rewriting the homogeneous Maxwell equation, the Bianchi identity (2.88), in the form

\[ \partial_c F^{ab} + \partial_a F^{bc} + \partial_b F^{ca} = 0, \]

(3.294)

and contracting it with $F^{ab}$, we see that the term (3.293) vanishes identically.

It is easy to construct from (3.286) Belinfante’s symmetric energy momentum tensor. We merely observe that the spin density is entirely due to the vector potential, and hence the same as before [see (3.236)]

\[ \Sigma^{ab,c} = - \left[ F^{ca} A^b - (a \leftrightarrow b) \right]. \]

(3.295)

Hence the additional piece to be added to the canonical energy momentum tensor is again [see (3.237)]

\[ \Delta \Theta^{ab} = \partial_c (F^{ab} A^c) = \frac{1}{2} (\partial_c F^{bc} A^a + F^{bc} \partial_c A^a). \]

(3.296)

The last term in this expression serves to symmetrize the electromagnetic part of the canonical energy-momentum tensor and brings it to the Belinfante form:

\[ T^{em}_{ab} = -F^b_c F^{ac} + \frac{1}{4} g^{ab} F^{cd} F_{cd}. \]

(3.297)

The second-last term in (3.296), which in the absence of charges vanished, is needed to symmetrize the matter part of $\Theta^{ab}$. Indeed, using once more Maxwell’s equation, it becomes

\[ -\frac{e}{c} \int d\tau \dot{x}^b(\tau) A^a \delta^{(4)}(x - x(\tau)), \]

(3.298)

thus canceling the corresponding term in (3.286). In this way we find that the total energy-momentum tensor of charged particles plus electromagnetic fields is simply the term of the two symmetric energy-momentum tensor.

\[ T^{ab} = \frac{m}{\dot{T}}^{ab} + \frac{e}{\dot{T}}^{em}_{ab} \]

(3.299)

\[ = m \int_{-\infty}^{\infty} d\tau \dot{x}^a \dot{x}^b \delta^{(4)}(x - x(\tau)) - F^b_c F^{ac} + \frac{1}{4} g^{ab} F^{cd} F_{cd}. \]
For completeness, let us cross check also its conservation. Forming the divergence \( \partial_b T^{ab} \), the first term gives now only

\[
\frac{e}{c} \int d\tau \dot{x}^b(\tau) F^a_b(x(\tau)),
\]

in contrast to (3.289), which is canceled by the divergence in the second term

\[
-\partial_b F^b_c F^{ac} = -\frac{e}{c} \int d\tau \dot{x}^c(\tau) F^{ac}(x(\tau)),
\]

in contrast to (3.292).

Notes and References

For more details on classical electromagnetic fields see

The individual citations refer to:

See also
E. Bessel-Hagen, Math. Ann. 84, 258 (1926);
L. Rosenfeld, Me. Acad. Roy. Belg. 18, 2 (1938);
F.J. Belinfante, Physica 6, 887 (1939).
[4] The Belinfante energy-momentum tensor is discussed in detail in
There are painters who transform the sun to a yellow spot, but there are others who transform a yellow spot into the sun.

Pablo Picasso (1881 - 1973)

4

Multivalued Gauge Transformations in Magnetostatics

For the upcoming development of a theory of gravitation with torsion it will be important to realize that it is possible to find a way to transform physical laws in Euclidean space into spaces with curvature and torsion. This can be done by a geometric generalization of a well-known field-theoretic technique developed by Dirac to introduce magnetic monopoles into electrodynamics. So far, no magnetic monopoles have been discovered in nature, but the mathematics used by Dirac will suggest us how to proceed in the geometric situation.

4.1 Vector Potential of Current Distribution

Let us begin by recalling the standard description of magnetism in terms of vector potentials. Since there are no magnetic monopoles in nature, a magnetic field $B(x)$ satisfies the identity $\nabla \cdot B(x) = 0$, implying that only two of the three field components of $B(x)$ are independent. To account for this, one usually expresses a magnetic field $B(x)$ in terms of a vector potential $A(x)$, setting $B(x) = \nabla \times A(x)$. Then Ampère’s law, which relates the magnetic field to the electric current density $j(x)$ by $\nabla \times B = j(x)$, becomes a second-order differential equation for the vector potential $A(x)$ in terms of an electric current

$$\nabla \times [\nabla \times A](x) = j(x).$$

(4.1)

In this chapter we are using natural units with $c = 1$ to save recurring factors of $c$.

The vector potential $A(x)$ is a gauge field. Given $A(x)$, any locally gauge-transformed field

$$A(x) \rightarrow A'(x) = A(x) + \nabla \Lambda(x)$$

(4.2)

yields the same magnetic field $B(x)$. This reduces the number of physical degrees of freedom in the gauge field $A(x)$ to two, just as those in $B(x)$. In order for
this to hold, the transformation function must be single-valued, i.e., it must have commuting derivatives
\[(\partial_i \partial_j - \partial_j \partial_i)\Lambda(\mathbf{x}) = 0.\] (4.3)
The equation for absence of magnetic monopoles \(\nabla \cdot \mathbf{B} = 0\) is ensured if the vector potential has commuting derivatives
\[(\partial_i \partial_j - \partial_j \partial_i)\mathbf{A}(\mathbf{x}) = 0.\] (4.4)
This integrability property makes \(\nabla \cdot \mathbf{B} = 0\) a *Bianchi identity* in this gauge field representation of the magnetic field [recall the generic definition after Eq. (2.87)].

In order to solve (4.1), we remove the gauge ambiguity by choosing a particular gauge, for instance the *transverse gauge* \(\nabla \cdot \mathbf{A}(\mathbf{x}) = 0\) in which \(\nabla \times [\nabla \times \mathbf{A}(\mathbf{x})] = -\nabla^2 \mathbf{A}(\mathbf{x})\), and obtain
\[\mathbf{A}(\mathbf{x}) = \frac{1}{4\pi} \int d^3 x' \frac{\mathbf{j}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}.\] (4.5)
The associated magnetic field is
\[\mathbf{B}(\mathbf{x}) = \frac{1}{4\pi} \int d^3 x' \frac{\mathbf{j}(\mathbf{x}') \times \mathbf{R}'}{R'^3}, \quad \mathbf{R}' \equiv \mathbf{x}' - \mathbf{x}.\] (4.6)
This standard representation of magnetic fields is not the only possible one. There exists another one in terms of a scalar potential \(\Lambda(\mathbf{x})\), which must, however, be multivalued to account for the two physical degrees of freedom in the magnetic field.

### 4.2 Multivalued Gradient Representation of Magnetic Field

Consider an infinitesimally thin closed wire carrying an electric current \(I\) along the line \(L\). It corresponds to a current density
\[\mathbf{j}(\mathbf{x}) = I \delta(\mathbf{x}; L),\] (4.7)
where \(\delta(\mathbf{x}; L)\) is the \(\delta\)-function on the closed line \(L\):
\[\delta(\mathbf{x}; L) = \int_L d\mathbf{x}' \delta^{(3)}(\mathbf{x} - \mathbf{x}').\] (4.8)
For a closed line \(L\), this function has zero divergence:
\[\nabla \cdot \delta(\mathbf{x}; L) = 0.\] (4.9)
This follows from the property of the \(\delta\)-function on an arbitrary open line \(L_{x_1}^{x_2}\) connecting the points \(\mathbf{x}_1\) and \(\mathbf{x}_2\) defined by
\[\delta(\mathbf{x}; L_{x_1}^{x_2}) = \int_{x_1}^{x_2} d\mathbf{x}' \delta^{(3)}(\mathbf{x} - \mathbf{x}').\] (4.10)
Multivalued Gauge Transformations in Magnetostatics

B(x) = \nabla \Omega(x)

Figure 4.1 Infinitesimally thin closed current loop L. The magnetic field B(x) at the point x is proportional to the solid angle \Omega(x) under which the loop is seen from x. In any single-valued definition of \Omega(x), there is some surface S across which \Omega(x) jumps by 4\pi. In the multivalued definition, this surface is absent.

which satisfies

$$\nabla \cdot \delta(x; L_{x_1}^{x_2}) = \delta(x_1) - \delta(x_2).$$

(4.11)

For closed loops, the right-hand side of (4.11) vanishes.

As an example, take a line L_{x_1}^{x_2} which runs along the positive z-axis from z_1 to z_2, so that

$$\delta(x; L_{x_1}^{x_2}) = \int_{z_1}^{z_2} dz' \delta(x) \delta(y) \delta(z - z') = \delta(x) \delta(y) [\Theta(z - z_1) - \Theta(z - z_2)],$$

(4.12)

and

$$\nabla \cdot \delta(x; L_{x_1}^{x_2}) = \delta(x) \delta(y) [\delta(z - z_1) - \delta(z - z_2)] = \delta(x_1) - \delta(x_2).$$

(4.13)

From Eq. (4.5) we obtain the associated vector potential

$$A(x) = \frac{I}{4\pi} \int_L dx' \frac{1}{|x - x'|},$$

(4.14)

yielding the magnetic field

$$B(x) = \frac{I}{4\pi} \int_L \frac{dx' \times R'}{R'^3}, \quad R' \equiv x' - x.$$

(4.15)

The same result will now be derived from a multivalued scalar field. Let \Omega(x; S) be the solid angle under which the current loop L is seen from the point x (see
Fig. 4.1). If $S$ denotes an arbitrary smooth surface enclosed by the loop $L$, and $dS'$ a surface element, then $\Omega(x; S)$ can be calculated from the surface integral

$$\Omega(x; S) = \int_S \frac{dS' \cdot R'}{R'^3}. \quad (4.16)$$

The argument $S$ in $\Omega(x; S)$ emphasizes that the definition depends on the choice of the surface $S$. The range of $\Omega(x; S)$ is from $-2\pi$ to $2\pi$, as can most easily be seen if $L$ lies in the $xy$-plane and $S$ is chosen to lie in the same place. Then we find for $\Omega(x; S)$ the value $2\pi$ for $x$ just below $S$, and $-2\pi$ just above. We form the vector field

$$B(x; S) = \frac{I}{4\pi} \nabla \Omega(x; S), \quad (4.17)$$

which is equal to

$$B(x; S) = \frac{I}{4\pi} \int_S dS' \nabla \frac{R'_k}{R'^3} = -\frac{I}{4\pi} \int_S dS' \nabla \frac{R'_k}{R'^3}. \quad (4.18)$$

This can be rearranged to

$$B_i(x; S) = -\frac{I}{4\pi} \left[ \int_S (dS'_k \partial'_i \frac{R'_k}{R'^3} - dS'_i \partial'_k \frac{R'_k}{R'^3}) + \int_S dS'_i \partial'_k \frac{R'_k}{R'^3} \right]. \quad (4.19)$$

With the help of Stokes’ theorem

$$\int_S (dS_k \partial_k - dS_i \partial_i) f(x) = \epsilon_{kil} \int_L dx_l f(x), \quad (4.20)$$

and the relation $\partial'_k (R'_k/R'^3) = 4\pi \delta^{(3)}(x - x')$, this becomes

$$B(x; S) = -I \left[ \frac{1}{4\pi} \int_L \frac{dx' \times R'}{R'^3} + \int_S dS' \delta^{(3)}(x - x') \right]. \quad (4.21)$$

The first term is recognized to be precisely the magnetic field (4.15) of the current $I$. The second term is the singular magnetic field of an infinitely thin magnetic dipole layer lying on the arbitrarily chosen surface $S$ enclosed by $L$.

The second term is a consequence of the fact that the solid angle $\Omega(x; S)$ was defined by the surface integral (4.16). If $x$ crosses the surface $S$, the solid angle jumps by $4\pi$.

It is useful to re-express Eq. (4.18) in a slightly different way. By analogy with (4.22) we define a $\delta$-function on a surface as

$$\delta(x; S) = \int_S dS' \delta^{(3)}(x - x'), \quad (4.22)$$

and observe that Stokes’ theorem (4.20) can be written as an identity for $\delta$-functions:

$$\nabla \times \delta(x; S) = \delta(x; L), \quad (4.23)$$
where \( L \) is the boundary of the surface \( S \). This equation proves once more the zero divergence (4.9).

Using the \( \delta \)-function on a surface \( S \), we can rewrite (4.16) as

\[
\Omega(x; S) = \int d^3 x' \delta(x'; S) \cdot \frac{R'_{x}}{R''}, \tag{4.24}
\]

and (4.18) as

\[
B(x; S) = -\frac{I}{4\pi} \int d^3 x' \delta(x'; S) \nabla \cdot \frac{R'_{x}}{R''}, \tag{4.25}
\]

and (4.19), after an integration by parts, as

\[
B_i(x; S) = -\frac{I}{4\pi} \left\{ \int d^3 x' \left[ \partial'_i \delta_k(x'; S) - \partial'_k \delta_i(x'; S) \right] \frac{R'_{k}}{R''} - \int d^3 x' \delta_i(x'; S) \nabla' \cdot \frac{R'_{x}}{R''} \right\}. \tag{4.26}
\]

The divergence at the end yields a \( \delta(3) \)-function, and we obtain

\[
B_i(x; S) = -\frac{I}{4\pi} \left[ \int d^3 x' \left[ \nabla \delta(x; S) \right] \times \frac{R'}{R''} + \int d^3 x' \delta(x'; S) \delta^{(3)}(x - x') \right]. \tag{4.27}
\]

Using (4.23) and (4.22), this is once more equal to (4.21).

Stokes theorem written in the form (4.23) displays an important property. If we move the surface \( S \) to \( S' \) with the same boundary, the \( \delta \)-function \( \delta(x; S) \) changes by

\[
\delta(x; S) \rightarrow \delta(x; S') = \delta(x; S) + \nabla \delta(x; V), \tag{4.28}
\]

where

\[
\delta(x; V) \equiv \int d^3 x' \delta^{(3)}(x - x'), \tag{4.29}
\]

and \( V \) is the volume over which the surface has swept. Under this transformation, the curl on the left-hand side of (4.23) is invariant. Comparing (4.28) with (4.2) we identify (4.28) as a novel type of gauge transformation [1, 2]. The magnetic field in the first term of (4.27) is invariant under this, the second is not. It is then obvious how to find a gauge-invariant magnetic field: we simply subtract the singular \( S \)-dependent term and form

\[
B(x) = \frac{I}{4\pi} \left[ \nabla \Omega(x; S) + 4\pi \delta(x; S) \right]. \tag{4.30}
\]

This field is independent of the choice of \( S \) and coincides with the magnetic field (4.15) derived in the usual gauge theory. To verify this explicitly we calculate the change of the solid angle (4.16) under a change of \( S \). For this we rewrite (4.24) as

\[
\Omega(x; S) = -\int d^3 x' \nabla' \frac{1}{R} \cdot \delta(x'; S) = -\frac{4\pi}{\nabla^2} \nabla \cdot \delta(x; S). \tag{4.31}
\]
Performing the vortex gauge transformation (4.28), the solid angle changes by
\[ \Delta \Omega(x; S) = -\frac{4\pi}{\nabla^2} \nabla \cdot \nabla \delta(x; V) = -4\pi \delta(x; L), \] (4.32)
so that (4.30) is indeed invariant.

Hence the description of the magnetic field as a gradient of field \( \Omega(x; S) \) is completely equivalent to the usual gauge field description in terms of the vector potential \( A(x) \). Both are gauge theories, but of a completely different type.

The gauge freedom (4.28) can be used to move the surface \( S \) into a standard configuration. One possibility is to choose \( S \) so that the third component of \( \delta(x; S) \) vanishes. This is called the axial gauge. If \( \delta(x; S) \) does not have this property, we can always shift \( S \) by a volume \( V \) determined by the equation
\[ \delta(V) = -\int_{-\infty}^{\infty} \delta_z(x; S), \] (4.33)
and the transformation (4.28) will produce a \( \delta(x; S) \) in the axial gauge \( \delta_3(x; S) = 0 \).

A general differential relation between \( \delta \)-functions on volumes and surfaces related to (4.33) is
\[ \nabla \delta(x; V) = -\delta(x; S). \] (4.34)

There exists another possibility of defining a solid angle \( \Omega(x; L) \) which is independent of the shape of the surface \( S \) and depends only on the boundary line \( L \) of \( S \). This is done by analytic continuation of \( \Omega(x; S) \) through the surface \( S \). This removes the jump and produces a \textit{multivalued function} \( \Omega(x; L) \) ranging from \(-\infty\) to \(\infty\). At each point in space, there are infinitely many Riemann sheets whose branch line is \( L \). The values of \( \Omega(x; L) \) on the sheets differ by integer multiples of \( 4\pi \). From this multivalued function, the magnetic field (4.15) can be obtained as a simple gradient:
\[ B(x) = \frac{I}{4\pi} \nabla \Omega(x; L). \] (4.35)

Ampère’s law (4.1) implies that the multivalued solid angle \( \Omega(x; L) \) satisfies the equation
\[ (\partial_i \partial_j - \partial_j \partial_i) \Omega(x; L) = 4\pi \epsilon_{ijk} \delta_k(x; L). \] (4.36)

Thus, as a consequence of its multivaluedness, \( \Omega(x; L) \) violates the Schwarz integrability condition. This makes it an unusual mathematical object to deal with. It is, however, perfectly suited to describe the magnetic field of an electric current along \( L \).

Let us see explicitly how Eq. (4.36) is fulfilled by \( \Omega(x; L) \), let us go to two dimensions where the loop corresponds to two points (in which the loop intersects a plane). For simplicity, we move one of them to infinity, and place the other at the
coordinate origin. The role of the solid angle \( \Omega(\mathbf{x}; L) \) is now played by the azimuthal angle \( \varphi(\mathbf{x}) \) of the point \( \mathbf{x} \) with respect to the origin:

\[
\varphi(\mathbf{x}) = \arctan \frac{x^2}{x^1}.
\]  

(4.37)

The function \( \arctan(x^2/x^1) \) is usually made unique by cutting the \( \mathbf{x} \)-plane from the origin along some line \( C \) to infinity, preferably along a straight line to \( \mathbf{x} = (-\infty, 0) \), and assuming \( \varphi(\mathbf{x}) \) to jump from \( \pi \) to \(-\pi \) when crossing the cut, as shown in Fig. 4.2(a). The cut corresponds to the magnetic dipole surface \( S \) in the integral

![Figure 4.2 Single- and multi-valued definitions of \( \arctan \varphi \).](image)

(4.16). In contrast to this, we shall take \( \varphi(\mathbf{x}) \) to be the multivalued analytic continuation of this function. Then the derivative \( \partial_i \) yields

\[
\partial_i \varphi(\mathbf{x}) = -\epsilon_{ij} \frac{x_j}{(x^1)^2 + (x^2)^2}.
\]  

(4.38)

This is in contrast to the derivative \( \partial_i \varphi(\mathbf{x}) \) of the single-valued definition of \( \partial_i \varphi(\mathbf{x}) \) which would contain an extra \( \delta \)-function \( \epsilon_{ij}\delta_j(C; \mathbf{x}) \) across the cut \( C \), corresponding to the second term in (4.21). When integrating the curl of the derivative (4.38) across the surface \( s \) of a small circle \( c \) around the origin, we obtain by Stokes’ theorem

\[
\int_s d^2x (\partial_i \partial_j - \partial_j \partial_i) \varphi(\mathbf{x}) = \int_c dx_i \partial_i \varphi(\mathbf{x}),
\]  

(4.39)

which is equal to \( 2\pi \) for the multivalued definition of \( \varphi(\mathbf{x}) \) shown in Fig. 4.2(b) and the book cover. This result implies the violation of the integrability condition as in (4.48):

\[
(\partial_i \partial_2 - \partial_2 \partial_i) \varphi(\mathbf{x}) = 2\pi \delta^{(2)}(\mathbf{x}),
\]  

(4.40)
4.2 Multivalued Gradient Representation of Magnetic Field

whose three-dimensional generalization is (4.36). In the single-valued definition of \( \varphi(x) \) with the jump by \( 2\pi \) across the cut \( C \), the right-hand side of (4.39) would vanish, since the contribution from the jump would cancel the integral along \( c \), so that \( \varphi(x) \) would satisfy the integrability condition (4.36).

On the basis of Eq. (4.40) we may construct a Green function for solving the corresponding differential equation with an arbitrary source, which is a superposition of infinitesimally thin line-like currents piercing the two-dimensional space at the points \( x_n \):

\[
j(x) = \sum_n I_n \delta^{(2)}(x - x_n),
\]

(4.41)

where \( I_n \) are currents. We may then easily solve the differential equation

\[
(\partial_1 \partial_2 - \partial_2 \partial_1) f(x) = j(x),
\]

(4.42)

with the help of the Green function

\[
G(x, x') = \frac{1}{2\pi} \varphi(x - x')
\]

(4.43)

which satisfies

\[
(\partial_1 \partial_2 - \partial_2 \partial_1) G(x - x') = \delta^{(2)}(x - x').
\]

(4.44)

The solution of (4.42) is obviously

\[
f(x) = \int d^2x' G(x, x') j(x).
\]

(4.45)

The gradient of \( f(x) \) yields the magnetic field of an arbitrary set of line-like currents vertical to the plane under consideration.

It is interesting to realize that the Green function (4.43) is the imaginary part of the complex function \( (1/2\pi) \log(z - z') \) with \( z = x_1 + ix_2 \), whose real part \( (1/2\pi) \log |z - z'| \) is the Green function \( G_{\Delta}(x - x') \) of the two-dimensional Poisson equation:

\[
(\partial_1^2 + \partial_2^2) G_{\Delta}(x - x') = \delta^{(2)}(x - x').
\]

(4.46)

It is important to point out that the superposition of line-like currents cannot be smeared out into a continuous distribution. The integral (4.45) yields the superposition of multivalued functions

\[
f(x) = \frac{1}{2\pi} \sum_n I_n \arctan \frac{x_2 - x_{2n}}{x_1 - x_{1n}},
\]

(4.47)

which is properly defined only if one can clearly continue it analytically into all Riemann sheets branching off from the endpoints of the cut at the origin. If we
were to replace the sum by an integral, this possibility would be lost. Thus it is, strictly speaking, impossible to represent arbitrary continuous magnetic fields as gradients of superpositions of scalar potentials $\Omega(\mathbf{x}; L)$. This, however, is not a severe disadvantage of this representation since arbitrary currents can be approximated by a superposition of line-like currents with any desired accuracy, and the same will be true for the associated magnetic fields.

The arbitrariness of the shape of the jumping surface is the origin of a further interesting gauge structure which has interesting physical consequences discussed in Subsection 4.6.

### 4.3 Generating Magnetic Fields by Multivalued Gauge Transformations

After this first exercise in multivalued functions, we now turn to another example in magnetism which will lead directly to our intended geometric application. We observed before that the local gauge transformation (4.2) produces the same magnetic field $\mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x})$ only, as long as the function $\Lambda(\mathbf{x})$ satisfies the Schwarz integrability criterion (4.36)

$$
(\partial_i \partial_j - \partial_j \partial_i) \Lambda(\mathbf{x}) = 0.
$$

Any function $\Lambda(\mathbf{x})$ violating this condition would change the magnetic field by

$$
\Delta B_k(\mathbf{x}) = \epsilon_{kij} (\partial_i \partial_j - \partial_j \partial_i) \Lambda(\mathbf{x}),
$$

thus being no proper gauge function. The gradient of $\Lambda(\mathbf{x})$

$$
\mathbf{A}(\mathbf{x}) = \nabla \Lambda(\mathbf{x})
$$

would be a nontrivial vector potential.

By analogy with the multivalued coordinate transformations violating the integrability conditions of Schwarz as in (4.36), the function $\Lambda(\mathbf{x})$ will be called nonholonomic gauge function.

Having just learned how to deal with multivalued functions we may change our attitude towards gauge transformations and decide to generate all magnetic fields approximately in a field-free space by such improper gauge transformations $\Lambda(\mathbf{x})$.

By choosing for instance

$$
\Lambda(\mathbf{x}) = \frac{\Phi}{4\pi} \Omega(\mathbf{x}),
$$

we see from (4.36) that this generates a field

$$
B_k(\mathbf{x}) = \epsilon_{kij} (\partial_i \partial_j - \partial_j \partial_i) \Lambda(\mathbf{x}) = \Phi \delta_k(\mathbf{x}; L).
$$

This is a magnetic field of total flux $\Phi$ inside an infinitesimal tube. By a superposition of such infinitesimally thin flux tubes analogous to (4.45) we can obviously generate a discrete approximation to any desired magnetic field in a field-free space.
4.4 Magnetic Monopoles

Multivalued fields have also been used to describe magnetic monopoles [4, 5, 6]. A monopole charge density $\rho_m(x)$ is the source of a magnetic field $B(x)$ as defined by the equation

$$\nabla \cdot B(x) = \rho_m(x).$$  \hfill (4.53)

If $B(x)$ is expressed in terms of a vector potential $A(x)$ as $B(x) = \nabla \times A(x)$, equation (4.53) implies the noncommutativity of derivatives in front of the vector potential $A(x)$:

$$\frac{1}{2} \epsilon_{ijk} \partial_i \partial_j A_k(x) = \rho_m(x).$$  \hfill (4.54)

Thus $A(x)$ must be multivalued. Dirac in his famous theory of monopoles [7, 8, 9] made the field single-valued by attaching to the world line of the particle a jumping world surface, whose intersection with a coordinate plane at a fixed time forms the Dirac string, along which the magnetic field of the monopole is imported from infinity. This world surface can be made physically irrelevant by quantizing it appropriately with respect to the charge. Its shape in space is just as irrelevant as that of the jumping surface $S$ in Fig. 4.1. The invariance under shape deformations constitute once more a second gauge structure of the type mentioned earlier and discussed in Refs. [10, 4, 11, 12, 2].

Once we allow ourselves to work with multivalued fields, we may easily go one step further and express also $A(x)$ as a gradient of a scalar field as in (4.50). Then the condition becomes

$$\epsilon_{ijk} \partial_i \partial_j \partial_k \Lambda(x) = \rho_m(x).$$  \hfill (4.55)

Let us construct the field of a magnetic monopole of charge $g$ at a point $x_0$, which satisfies (4.53) with $\rho(x) = g\delta^{(3)}(x - x_0)$. Physically, this can be done only by setting up an infinitely thin solenoid (Dirac string) along an arbitrary line $L^{x_0}$ which imports the flux from somewhere at infinity to the point $x_0$ where the flux emerges. The superscript 0 indicates that the line ends at $x_0$. Inside this solenoid, the magnetic field is infinite, equal to

$$B_{\text{inside}}(x; L) = g\delta(x; L^{x_0}),$$  \hfill (4.56)

where $\delta(x; L^{x_0})$ is a modification of (4.10) in which the integral runs along the line $L^{x_0}$ to $x_0$:

$$\delta(x; L^{x_0}) = \int_{L^{x_0}}^{x} d^3x' \delta^{(3)}(x - x').$$  \hfill (4.57)

The divergence of this function is concentrated at the endpoint $x_0$ of the solenoid:

$$\nabla \cdot \delta(x; L^{x_0}) = -\delta^{(3)}(x - x_0).$$  \hfill (4.58)
Similarly we may define a $\delta$-function along a line $L_{x_0}$ which starts at $x_0$ and runs to somewhere at infinity:

$$\delta(x; L_{x_0}) = \int_{x_0} d^3 x' \delta^{(3)}(x - x'),$$

which satisfies

$$\nabla \cdot \delta(x; L_{x_0}) = \delta^{(3)}(x - x_0).$$

This describes a thin solenoid (Dirac string) which exports the magnetic flux from $x_0$ to infinity, corresponding to an antimonopole at $x_0$.

As an example, take a line $L_{x_0}$ which carries the flux from positive infinity to the origin along the $z$-axis. If $\hat{z}$ denotes the unit vector along the $z$-axis, then

$$\delta(x; L_{x_0}) = \hat{z} \int_{-\infty}^{\infty} dz' \delta(x) \delta(y) \delta(z - z') = \hat{z} \delta(x) \delta(y) \Theta(-z),$$

where $S$ is the surface over which $L_{x_0}$ has swept on its way to $L'_{x_0}$. Under this gauge transformation, the relation (4.58) is obviously invariant. We shall call this monopole gauge invariance. The flux (4.56) inside the solenoid is therefore a monopole gauge field.

Note that with respect to the previous gauge transformations (4.28) which shifted a surface, the gradient is exchanged by a curl, and the opposite exchange relates the invariants, which was a boundary line found from a curl in Eq. (4.23), and is here the starting point of the line $L_{x_0}$ found from the divergence in Eq. (4.58).

It is straightforward to construct the associated ordinary gauge field $A(x)$ of the monopole. Consider first the $L_{x_0}$-dependent field

$$A(x; L_{x_0}) = \frac{g}{4\pi} \int d^3 x' \frac{\nabla' \times \delta(x'; L_{x_0})}{R'} = -\frac{g}{4\pi} \int d^3 x' \delta(x'; L_{x_0}) \times \frac{R'}{R'^3}. $$

The curl of the first expression is

$$\nabla \times A(x; L_{x_0}) = \frac{g}{4\pi} \int d^3 x' \frac{\nabla' \times \delta(x'; L_{x_0})}{R'}$$

H. Kleinert, GRAVITY WITH TORSION
and consists of two terms

\[
\frac{g}{4\pi} \int d^3x' \frac{\nabla' \cdot \delta(x'; L^{x_0})}{R'} - \frac{g}{4\pi} \int d^3x' \frac{\nabla'^2 \delta(x'; L^{x_0})}{R'}.
\] (4.66)

After an integration by parts, and using (4.58), the first term is $L^{x_0}$-independent and reads

\[
\frac{g}{4\pi} \int d^3x' \delta^{(3)}(x - x_0) \nabla' \frac{1}{R} = \frac{g}{4\pi} \frac{x - x_0}{|x - x_0|^3}.
\] (4.67)

The second term becomes, after two integration by parts,

\[
g \delta(x'; L^{x_0}).
\] (4.68)

The first term is the desired magnetic field of the monopole. Its divergence is $\delta^{(3)}(x - x_0)$, which we wanted to archive. The second term is the monopole gauge field, the magnetic field inside the solenoid. The total divergence of this field is, of course, zero.

By analogy with (4.30) we now subtract the latter term and find the $L^{x_0}$-independent magnetic field of the monopole

\[
B(x) = \nabla \times A(x; L^{x_0}) - g \delta(x; L^{x_0}),
\] (4.69)

which depends only on $x_0$ and satisfies $\nabla \cdot B(x) = g \delta^{(3)}(x - x_0)$.

Let us calculate the vector potential explicitly for the monopole where the solenoid comes in along $L^{x_0}$. Inserting (4.61) into the right-hand side of (4.64), we obtain

\[
A^{(g)}(x; L^{x_0}) = -\frac{g}{4\pi} \int_0^\infty dz' \frac{\hat{z} \times x}{\sqrt{x'^2 + y^2 + (z' - z)^2}}
\]

\[
= -\frac{g}{4\pi} \frac{\hat{z} \times x}{R(R - z)} = \frac{g}{4\pi} \frac{y, -x, 0}{R(R - z)}.
\] (4.70)

Alternatively, if $L^{x_0}$ runs to $-\infty$, so that $\delta(x; L^{x_0})$ is equal to $-\hat{z} \Theta(-z) \delta(x) \delta(y)$, we obtain

\[
A^{(g)}(x; L^{x_0}) = -\frac{g}{4\pi} \int_0^\infty dz' \frac{\hat{z} \times x}{\sqrt{x'^2 + y^2 + (z' - z)^2}^{3/2}}
\]

\[
= \frac{g}{4\pi} \frac{\hat{z} \times x}{R(R + z)} = -\frac{g}{4\pi} \frac{y, -x, 0}{R(R + z)}.
\] (4.71)

The vector potential has only azimuthal components. If we parametrize $(x, y, z)$ in terms of spherical coordinates as $r \sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta$, these are

\[
A^{(g)}_{\varphi}(x; L^{x_0}) = \frac{g \sin \theta}{4\pi R(1 + \cos \theta)} \quad \text{or} \quad A^{(g)}_{\varphi}(x; L^{x_0}) = -\frac{g \sin \theta}{4\pi R(1 - \cos \theta)},
\] (4.72)
respectively.

The shape of the line \( L^{x_0} \) (or \( L_{x_0} \)) can be brought to a standard form, which corresponds to fixing a gauge of the field \( \delta(x; L^{x_0}) \) or \( \delta(x; L_{x_0}) \). For example, we may always choose \( L^{x_0} \) to run along the positive \( z \)-axis.

An interesting observation is the following: If the gauge function \( \Lambda(x) \) is considered as a nonholonomic displacement in some fictitious crystal dimension, then the magnetic field of a current loop which gives rise to noncommuting derivatives
\[
(\partial_j \partial_i - \partial_i \partial_j) \Lambda(x) \neq 0
\]
is the analog of a dislocation [compare (14.3)], and thus implies torsion in the crystal. A magnetic monopole, on the other hand, arises from noncommuting derivatives
\[
(\partial_i \partial_j - \partial_j \partial_i) \partial_k \Lambda(x) \neq 0
\]
in Eq. (4.55).

### 4.5 Minimal Magnetic Coupling of Particles from Multivalued Gauge Transformations

Multivalued gauge transformations are the perfect tool to minimally couple electromagnetism to any type of matter. Consider for instance a free nonrelativistic point particle with a Lagrangian
\[
L = \frac{M}{2} \dot{x}^2. \tag{4.73}
\]
The equations of motion are invariant under a gauge transformation
\[
L \to L' = L + \nabla \Lambda(x) \dot{x}, \tag{4.74}
\]

since this changes the action \( \mathcal{A} = \int_a^b dt L \) merely by a surface term:
\[
\mathcal{A}' \to \mathcal{A} = \mathcal{A} + \Lambda(x_b) - \Lambda(x_a). \tag{4.75}
\]
The invariance is absent if we take \( \Lambda(x) \) to be a multivalued gauge function. In this case, a nontrivial vector potential \( A(x) = \nabla \Lambda(x) \) (working in natural units with \( e = 1 \)) is created in the field-free space, and the nonholonomically gauge-transformed Lagrangian corresponding to (4.74),
\[
L' = \frac{M}{2} \dot{x}^2 + A(x) \dot{x}, \tag{4.76}
\]
describes correctly the dynamics of a free particle in an external magnetic field.

The coupling derived by multivalued gauge transformations is automatically invariant under additional ordinary single-valued gauge transformations of the vector potential
\[
A(x) \to A'(x) = A(x) + \nabla \Lambda(x), \tag{4.77}
\]
since these add to the Lagrangian (4.76) once more the same pure derivative term which changes the action by an irrelevant surface term as in (4.75).
4.5 Minimal Magnetic Coupling of Particles from Multivalued Gauge Transformations

The same procedure leads in quantum mechanics to the minimal coupling of the Schrödinger field $\psi(x)$. The action is $A = \int dt d^3x \, L$ with a Lagrange density (in natural units with $\hbar = 1$)

$$L = \psi^*(x) \left( i\partial_t + \frac{1}{2M} \nabla^2 \right) \psi(x). \quad (4.78)$$

The physics described by a Schrödinger wave function $\psi(x)$ is invariant under arbitrary local phase changes

$$\psi(x, t) \rightarrow \psi'(x, t) = e^{i\Lambda(x)} \psi(x, t), \quad (4.79)$$
called local U(1) transformations. This implies that the Lagrange density (4.78) may equally well be replaced by the gauge-transformed one

$$L = \psi^*(x, t) \left( i\partial_t + \frac{1}{2M} D^2 \right) \psi(x, t), \quad (4.80)$$

where $-iD \equiv -i\nabla - \nabla\Lambda(x)$ is the operator of physical momentum.

We may now go over to nonzero magnetic fields by admitting gauge transformations with multivalued $\Lambda(x)$ whose gradient is a nontrivial vector potential $A(x)$ as in (4.50). Then $-iD$ turns into the covariant momentum operator

$$\hat{P} = -iD = -i\nabla - A(x), \quad (4.81)$$

and the Lagrange density (4.80) describes correctly the magnetic coupling in quantum mechanics.

As in the classical case, the coupling derived by multivalued gauge transformations is automatically invariant under ordinary single-valued gauge transformations under which the vector potential $A(x)$ changes as in (4.77), whereas the Schrödinger wave function undergoes a local U(1)-transformation (4.79). This invariance is a direct consequence of the simple transformation behavior of $D\psi(x, t)$ under gauge transformations (4.77) and (4.79) which is

$$D\psi(x, t) \rightarrow D\psi'(x, t) = e^{i\Lambda(x)}D\psi(x, t). \quad (4.82)$$

Thus $D\psi(x, t)$ transforms just like $\psi(x, t)$ itself, and for this reason, $D$ is called gauge-covariant derivative. The generation of magnetic fields by a multivalued gauge transformation is the simplest example for the power of the nonholonomic mapping principle.

After this discussion it is quite suggestive to introduce the same mathematics into differential geometry, where the role of gauge transformations is played by reparametrizations of the space coordinates.

4.6 Equivalence of Multivalued Scalar and
Single-Valued Vector Potential Representation

In the previous sections we have given examples for the use of multivalued fields in describing magnetic phenomena. The multivalued gauge transformations by which we created line-like nonzero field configurations were shown to be the natural origin of the minimal couplings to the classical actions as well as to the Schrödinger equation. It is interesting to establish the complete equivalence of the multivalued scalar theory with the usual vector potential theory of magnetism. This is done by a proper treatment of the the degrees of freedom of the jumping surfaces $S$. For this purpose we set up an action formalism for calculating the magnetic energy of a current loop in the gradient representation of the magnetic field. In Euclidean field theory, the action is provided by the field energy

$$ H = \frac{1}{2} \int d^3 x \mathbf{B}^2(\mathbf{x}). \quad (4.83) $$

Inserting the gradient representation (4.35) of the magnetic field, we can write this as

$$ H = \frac{I^2}{2(4\pi)^2} \int d^3 x [\nabla \Omega(\mathbf{x})]^2. \quad (4.84) $$

This holds for the multivalued solid angle $\Omega(\mathbf{x})$ which is independent of $S$. In order to perform field theoretic calculations, we must go over to the single-valued representation (4.30) of the magnetic field for which the energy is

$$ H = \frac{I^2}{2(4\pi)^2} \int d^3 x [\nabla \Omega(\mathbf{x}; S) + 4\pi \delta(\mathbf{x}; S)]^2. \quad (4.85) $$

The $\delta$-function removes the unphysical field energy on the artificial magnetic dipole layer on $S$.

The Hamiltonian is extremized by the scalar field (4.24). Moreover, due to infinite field strength on the surface, all field configurations $\Omega(\mathbf{x}; S')$ with a jumping surface $S'$ different from $S$ will have an infinite energy. Thus we may omit the argument $S$ in $\Omega(\mathbf{x}; S)$ and admit an arbitrary field $\Omega(\mathbf{x})$ to the Hamiltonian (4.85). Only the field (4.24) will give a finite contribution.

Let us calculate the magnetic field energy of the current loop from the energy (4.85). For this we rewrite the energy (4.85) in terms of an independent auxiliary vector field $\mathbf{B}(\mathbf{x})$ as

$$ H = \int d^3 x \left\{ -\frac{1}{2} \mathbf{B}^2(\mathbf{x}) - \frac{1}{4\pi} \mathbf{B}(\mathbf{x}) \cdot [\nabla \Omega(\mathbf{x}) + 4\pi I \delta(\mathbf{x}; S)] \right\}. \quad (4.86) $$

A partial integration brings the second term to

$$ \int d^3 x \frac{1}{4\pi} \nabla \cdot \mathbf{B}(\mathbf{x}) \Omega(\mathbf{x}). $$
Extremizing this in $\Omega(x)$ yields the equation
\[ \nabla \cdot \mathbf{B}(x) = 0, \] (4.87)
implying that the field lines of $\mathbf{B}(x)$ form closed loops. This equation may be enforced identically (as a Bianchi identity) by expressing $\mathbf{B}(x)$ as a curl of an auxiliary vector potential $\mathbf{A}(x)$, setting
\[ \mathbf{B}(x) \equiv \nabla \times \mathbf{A}(x). \] (4.88)
This ansatz brings the energy (4.86) to the form
\[ H = \int d^3 x \left\{ -\frac{1}{2} \left[ \nabla \times \mathbf{A}(x) \right]^2 - I \left[ \nabla \times \mathbf{A}(x) \right] \cdot \delta(x; S) \right\}. \] (4.89)
A partial integration of the second term leads to
\[ H = \int d^3 x \left\{ -\frac{1}{2} \left[ \nabla \times \mathbf{A}(x) \right]^2 - I \mathbf{A}(x) \cdot \left[ \nabla \times \delta(x; S) \right] \right\}, \] (4.90)
and we identify the factor of $\mathbf{A}(x)$ in the linear term as the auxiliary current
\[ \mathbf{j}(x) \equiv I \nabla \times \mathbf{A}(x; S) = I \delta(x; L). \] (4.91)
In the last step we have used Stokes’ law (4.23). According to Eq. (4.9), this current is conserved for loops $L$.

The representation (4.90) of the energy is called the dually transformed version of the original energy (4.85).

By extremizing the energy (4.89), we obtain Ampère’s law (4.1). Thus the auxiliary quantities $\mathbf{B}(x)$, $\mathbf{A}(x)$, and $\mathbf{j}(x)$ coincide with the usual magnetic quantities with the same name. If we insert the explicit solution (4.5) of Ampère’s law into the energy, we obtain the Biot-Savart energy for an arbitrary current distribution
\[ H = \frac{1}{8\pi} \int d^3 x \int d^3 x' \mathbf{j}(x) \frac{1}{|x-x'|} \mathbf{j}(x'). \] (4.92)
If we insert here two current filaments running parallel in thin wires, the energy (4.92) decreases with increasing distance suggesting, for a moment, that the force between them is repulsive. The experimental force, however, is attractive. The sign change is due to the fact that when increasing the distance of the wires we must perform work against the inductive forces in order to maintain the constant currents. This work is not calculated above and turns out to be exactly twice the energy gain implied by (4.92). The energy responsible for discussing the forces of external current distributions is the free magnetic energy
\[ F = \frac{1}{2} \int d^3 x \mathbf{B}^2(x) - \int d^3 x \mathbf{j}(x) \cdot \mathbf{A}(x). \] (4.93)
Extremizing this in \( \mathbf{A}(\mathbf{x}) \) yields the vector potential (4.5), and reinserting this into (4.93) we find, indeed, that the free Biot-Savart energy is the opposite of (4.92):

\[
F|_{\text{ext}} = -\frac{1}{8\pi} \int d^3x \, d^3x' \, j(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{x}'|} j(\mathbf{x}').
\]  

(4.94)

As a consequence, parallel wires with fixed currents attract each other rather than repel.

Note that the energy (4.89) is invariant under two mutually dual gauge transformations, the usual magnetic one in (4.2), by which the vector potential receives a gradient of an arbitrary scalar field, and the gauge transformation (4.28), by which the irrelevant surface \( S \) is moved to another configuration \( S' \).

Thus we have proved the complete equivalence of the gradient representation of the magnetic field to the usual gauge field representation. In the gradient representation, there exists a new type of gauge invariance which expresses the physical irrelevance of the jumping surface appearing when using single-valued solid angles.

The energy (4.90) describes magnetism in terms of a double gauge theory [13], in which both the gauge of \( \mathbf{A}(\mathbf{x}) \) and the shape of \( S \) can be changed arbitrarily. By setting up a grand-canonical partition function of many fluctuating surfaces it is possible to describe a large family of phase transitions mediated by the proliferation of line-like defects. Examples are vortex lines in the superfluid-normal transition in helium, to be discussed in the next chapter, and dislocation and disclination lines in the melting transition of crystals, to be discussed later [10, 4, 11, 12, 2].

4.7 Multivalued Field Theory of Magnetic Monopoles and Electric Currents

Let us now go through the analogous discussion for a gas of monopoles at \( \mathbf{x}_n \) with strings \( L^{x_n} \) importing their fluxed from infinity, and electric currents along closed \( L_{n'} \). The free energy of fixed currents is given by the energy of the magnetic field (4.69) coupled to the currents as in the action Eq. (4.93):

\[
F = \int d^3x \left\{ \frac{1}{2} \left[ \mathbf{\nabla} \times \mathbf{A} - g \sum_n \delta(\mathbf{x}; L^{x_n}) \right]^2 - I \mathbf{A}(\mathbf{x}) \cdot \sum_{n'} \delta(\mathbf{x}, L_{n'}) \right\}.
\]  

(4.95)

Extremizing this in \( \mathbf{A}(\mathbf{x}) \) we obtain

\[
\mathbf{A}(\mathbf{x}) = -\frac{1}{\nabla^2} \left[ g \sum_n \mathbf{\nabla} \times \delta(\mathbf{x}; L^{x_n}) + I \sum_{n'} \delta(\mathbf{x}, L_{n'}) \right].
\]  

(4.96)

Reinserting this into (4.95) yields three terms. First, there is an interaction between the current lines

\[
H_{II} = -\frac{I^2}{2} \int d^3x \sum_{n,n'} \delta(\mathbf{x}; L_n) \frac{1}{\nabla^2} \delta(\mathbf{x}; L_{n'}) = -\frac{I^2}{2} \sum_{n,n'} \int_{L_n} d\mathbf{x}_n \int_{L_{n'}} d\mathbf{x}_{n'} \frac{1}{|\mathbf{x}_n - \mathbf{x}_{n'}|}.
\]  

(4.97)
4.7 Multivalued Field Theory of Magnetic Monopoles and Electric Currents

which corresponds to (4.94). Second, there is an interaction between monopole strings

\[
\frac{g^2}{2} \int d^3x \left\{ \left[ \sum_n \delta(x; L^x_n) \right]^2 + \left[ \sum_n \nabla \times \delta(x; L^x_n) \right] \frac{1}{\nabla^2} \left[ \sum_n \nabla \times \delta(x; L^x_n) \right] \right\},
\]

(4.98)

which can be brought to the form

\[
H_{gg} = \frac{g^2}{2} \int d^3x \left[ \sum_n \nabla \cdot \delta(x; L^x_n) \right]^2 = \frac{g^2}{2} \int d^3x \left[ \sum_n \delta(x - x_n) \right]^2
= \frac{g^2}{8\pi} \sum_{n,n'} \left| x_n - x_{n'} \right|.
\]

(4.99)

Finally, there is an interaction between the monopoles and the currents

\[
H_{gl} = -g I \int d^3x \sum_{n,n'} \nabla \times \delta(x; L^x_n) \frac{1}{\nabla^2} \delta(x; L^x_{n'}). \tag{4.100}
\]

An integration by parts brings this to the form

\[
H_{gl} = -g I \int d^3x \sum_{n,n'} \delta(x; L^x_n) \frac{1}{\nabla^2} \nabla \times \delta(x; L^x_{n'})
= -g I \int d^3x \sum_{n,n'} \delta(x; L^x_n) \frac{1}{\nabla^2} \nabla \times [\nabla \times \delta(x; S_{n'})], \tag{4.101}
\]

which is equal to

\[
H_{gl} = H'_{gl} + \Delta H_{gl}, \tag{4.102}
\]

with

\[
H'_{gl} = -g I \int d^3x \sum_{n,n'} \delta(x; L^x_n) \nabla \frac{1}{\nabla^2} \left[ \nabla \cdot \delta(x; S_{n'}) \right], \tag{4.103}
\]

and

\[
\Delta H_{gl} = g I \int d^3x \sum_{n,n'} \delta(x; L^x_n) \delta(x; S_{n'}). \tag{4.104}
\]

Each integral in the sum yields an integer number which counts how often the lines \( L_n \) pierce the surface \( S_{n'} \), so that

\[
\Delta H_{gl} = g I k, \quad k = \text{integer}. \tag{4.105}
\]

Recalling (4.31), the interaction (4.103) can be rewritten as

\[
H_{gl} = -\frac{g I}{4\pi} \int d^3x \sum_{n,n'} \delta(x; L^x_n) \nabla \Omega(x; S_{n'}). \tag{4.106}
\]
An integration by parts and the relation (4.58) brings this to the form

\[ H_{gI} = \frac{gI}{4\pi} \sum_{n,n'} \Omega(x_n; S_{n'}). \]  

(4.107)

It is proportional to the sum of the solid angles \( \Omega(x_n; S_{n'}) \) under which the current loops \( L_n \) are seen from the monopoles at \( x_n \). The result does not depend on the surfaces \( S_n \), only on the boundary lines \( L_n \) along which the currents flow.

The total interaction is obviously invariant under shape deformations of \( S \), except for the term (4.105). This term, however, is physically irrelevant provided we subject the charges \( Q \) in the currents to quantization rule

\[ Qg = 2\pi k, \quad k = \text{integer}. \]  

(4.108)

This rule was first found by Dirac [7].

The quantization rule is a consequence of quantum theory. This is governed by amplitudes which can be calculated from the classical action by means of a functional integral

\[ \text{Amplitude} = \sum_{\text{field configurations}} e^{iA/\hbar}, \]  

(4.109)

where \( A \) is the full four-dimensional action of the system, which for static currents and monopoles is simply

\[ A = - \int dt H = -Qg. \]  

(4.110)

This shows that \( \Delta H \) in (4.105) does indeed not contribute to (4.109) if the quantization condition (4.108) is fulfilled since it does not change the amplitude \( e^{iA/\hbar} \).

At this place it should be mentioned that the Dirac quantization condition guarantees the invisibility of the Dirac string only for electric charges of integer spin. For electrons and all particles of half-integer spin, the wave function is double valued since it returns to its original value only after rotating it by \( 4\pi \). For these particles the electric charge must be twice as big as for integer spins, and satisfy the Schwinger quantization condition [8]

\[ Qg = 4\pi k, \quad k = \text{integer}. \]  

(4.111)

Notes and References


5

Multivalued Fields in Superfluids and Superconductors

Multivalued fields play an important role in understanding a great variety of phase transitions. In this chapter we shall discuss two simple but important examples.

5.1 Superfluid Transition

The simplest phase transitions which can be explained by multivalued field theory is the so-called λ-transition of superfluid helium. The name has its origin in the shape of the peak in the specific heat observed at a critical temperature $T_c \approx 2.18$ K shown in Fig. 5.1.

![Figure 5.1 Specific heat of superfluid $^4$He. For very small $T$, it shows the typical power behavior $\propto T^3$ characteristic for massless excitations in three dimensions in the Debye theory of specific heat. Here these excitations are phonons of the second sound. The peak is caused by the proliferation of vortex loops at the superfluid-normal transition.](image)

For temperatures $T$ below $T_c$, the fluid shows no friction and possesses only massless excitations. These are the quanta of the second sound, called phonons. They
5.1 Superfluid Transition

Figure 5.2 Energies of the elementary excitations in superfluid $^4$He measured by neutron scattering showing the roton minimum near $k \approx 2 \pi \AA$ (data are taken from Ref. [1]).

cause the typical temperature behavior of the specific heat which in $D$ dimensions starts out like

$$C \sim T^D. \tag{5.1}$$

This was first explained in 1912 by Debye in his theory of specific heat [2], in which he generalized Planck’s theory of black-body radiation to solid bodies.

As the temperature rises, another type of excitations appears in the superfluid. These are the famous rotons whose existence was deduced in 1947 by Landau from the thermodynamic properties of the superfluid [3, 4]. Rotons are the excitations of wavenumber near $2 \pi \AA$ where the phonon dispersion curve has a minimum. The full shape of this curve can be measured by neutron scattering and is displayed in Fig. 5.2.

As long as $T$ stays sufficiently far below $T_c$, the thermodynamic properties of the superfluid are dominated by phonons and rotons. If the temperature approaches $T_c$, the rotons join side by side and form large surfaces, a shown in Fig. 5.3. The adjacent

Figure 5.3 Near $T_c$, more and more rotons join side by side to form surfaces whose boundary appears as a large vortex loop. The adjacent roton boundaries cancel each other.
boundaries of the rotons cancel each other, so that the memory of the surfaces is lost, their shape becomes irrelevant, and only the boundaries of the surfaces are physical objects, observable as *vortex loops*. At $T_c$, the vortex loops become infinitely long and proliferate. The large activation energies for creating single rotons are overcome by the high configurational entropy of the long vortex loops.

The inside of a vortex line consists of normal fluid since the large rotation velocity destroys superfluidity. For this reason, the proliferation of the vortex loops fills the system with normal fluid, and the fluid looses its superfluid properties. The existence of such a mechanism for a phase transition was realized more than fifty years ago by Onsager in 1949 [5]. It was re-emphasized by Feynman in 1955 [6], and turned into a proper disorder field theory in the 1980’s by the author [7]. The same idea was advanced in 1952 by Shockley [8] who proposed a proliferation of defect lines in solids to be responsible for the melting transition. His work instigated the author to develop a detailed theory of melting in the textbook [9].

The disorder field theory of superfluids and superconductors will be derived in Subsection 5.1.10, the melting theory in Chapter 10.

### 5.1.1 Configuration Entropy

There exists a simple estimate for the temperature of a phase transition based on the proliferation of line-like excitations. A long line of length $l$ with an energy per length $\epsilon$ is suppressed strongly by a Boltzmann factor $e^{-\epsilon l/T}$. This suppression is, however, counteracted by configurational entropy. Suppose the line can bend easily on a length scale $\xi$ which is of the order of the coherence length of the system. Hence it can occur in approximately $(2D)^{l/\xi}$ possible configurations, where $D$ is the space dimension [10]. A rough approximation to the partition function of a single loop of arbitrary length is given by the integral

$$Z_1 \approx \oint \frac{dl}{l} (2D)^{l/\xi} e^{-\epsilon l/T}. \tag{5.2}$$

The factor $1/l$ in the integrand accounts for the cyclic invariance of the loop. By exponentiating this one-loop expression we obtain the partition function of a grand-canonical ensemble of loops of arbitrary length $l$, whose free energy is therefore $F = -Z_1/\beta$.

The integral (5.2) converges only below a critical temperature

$$T < T_c = \epsilon \xi / \log(2D). \tag{5.3}$$

Above $T_c$, the integral diverges and the ensemble undergoes a phase transition in which the loops proliferate and become infinitely long. This process will be called *condensation* of loops. A Monte-Carlo simulation of this process is shown in Fig. 5.4.

From Eq. (5.3) we can immediately deduce a relation between the critical temperature and the roton energy in superfluids. The size of a roton will be roughly
5.1 Superfluid Transition

Figure 5.4 Vortex loops in XY-model with periodic boundary condition for different values of $\beta = 1/k_BT$. Close to $T_c \equiv 3$, the loops proliferate, with some becoming infinitely long (from Ref. [11]). The plots show the views of left and right eye. To perceive the loops three-dimensionally, place a sheet of paper vertically between the pictures and point the eyes parallel until you see only one picture.

$\pi \xi$. Its energy is therefore $E_{\text{roton}} \approx \pi \xi \epsilon$. Inserting this into Eq. (5.3) we estimate the critical temperature of a line ensemble as

$$T_c = c_{\text{lines}} E_{\text{rot}}.$$  \hspace{1cm} (5.4)

It is proportional to the roton energy with a proportionality constant in $D = 3$ dimensions

$$c_{\text{lines}} \approx 1/\pi \log 6 \approx 1/5.6.$$  \hspace{1cm} (5.5)

This prediction was recently verified experimentally [12].
5.1.2 Origin of Massless Excitations

The massless excitations in superfluid helium are a consequence of a spontaneous breakdown of a continuous symmetry of the Hamiltonian. Such massless excitations are called Nambu-Goldstone modes. These arise as follows. Superfluid \(^4\)He is described by a complex order field \(\phi(x)\) which is the wave function of the condensate.

Near the transition and for smooth spatial variations, the energy density is given by the Hamiltonian of Landau, Ginzburg, and Pitaevskii \([13]\)

\[
H[\phi] = \frac{1}{2} \int d^3x \left\{ |\nabla \phi|^2 + \tau |\phi|^2 + \frac{\lambda}{2} |\phi|^4 \right\},
\]

where \(\tau\) is proportional to the relative temperature distance from the critical temperature

\[
\tau \equiv \mu_0 (T/T_c - 1).
\]

Below the critical temperature where \(\tau < 0\), the ground state lies at

\[
\phi(x) = \phi_0 = \sqrt{-\frac{\tau}{\lambda}} e^{i\alpha}.
\]

This field value is called the order parameter of the superfluid.

The ground state is not unique but infinitely degenerate. Only its absolute value of \(|\phi_0|\) is fixed, the phase \(\alpha\) is arbitrary. For this reason, the entropy does not go to zero at zero temperature. The degeneracy in \(\alpha\) is due to the fact that the Hamiltonian density (5.6) is invariant under constant U(1) phase transformations

\[
\phi(x) \rightarrow e^{i\alpha} \phi(x).
\]

The Nambu-Goldstone theorem states that such a degenerate ground state possesses massless excitations, unless there is another massless excitation which prevents this by mixing with the Goldstone excitation. In the field theory with Hamiltonian (5.6), the appearance of massless excitations is easily understood by decomposing the order field \(\phi(x)\) into size and phase variables,

\[
\phi(x) = \rho(x) e^{i\theta(x)},
\]

and rewriting (5.6) as

\[
H[\rho, \theta] = \frac{1}{2} \int d^3x \left[ (\nabla \rho)^2 + \rho^2 (\nabla \theta)^2 + \tau \rho^2 + \frac{\lambda}{2} \rho^4 \right].
\]

If \(\tau\) is negative, the size of the order field is frozen at the minimum (5.8), implying \(\rho_0 = \sqrt{-\tau/\lambda}\), and the Hamiltonian (5.6) can be approximated by its so-called hydrodynamic limit, also called London limit (see Section 7.2)

\[
H^{\text{hy}}[\theta] = \rho_0^2 \int d^3x (\nabla \theta)^2.
\]
5.1 Superfluid Transition

We have omitted a constant \textit{condensation energy}

\[ H^\text{hy}_c = - \int d^3x \frac{\tau^2}{2\lambda}. \]  

(5.13)

The Hamiltonian density (5.12) shows that the energy of a plane-wave excitation of the phase grows with the square of the wave vector $k$, and goes to zero for $k \to 0$. These are the massless Nambu-Goldstone modes. By rewriting (5.12) as

\[ H^\text{hy}[\theta] = \frac{\rho_s}{2M} \int d^3 x (\nabla \theta)^2; \]  

(5.14)

we obtain the usual hydrodynamic kinetic energy, and identify

\[ \rho_s = M \rho_0^2 \]  

(5.15)

as the \textit{superfluid density}.

Apart from the constant field $\phi(x) = \phi_0$, there exist nontrivial field configurations which extremize the Hamiltonian (5.6). They represent vortex lines which play a crucial role in many phenomena encountered in superfluid helium. Some relevant properties of these lines are discussed in Appendix 5A. Here we only note that at the center of each line, the size $\rho(x)$ of order field vanishes. The question arises as to what happens to these solutions in the hydrodynamic limit where $\rho(x)$ is constant everywhere? The alert reader may have noticed that in going from (5.6) to (5.11) we have made an important error which for the discussion of the Nambu-Goldstone mechanism was irrelevant but becomes important for the understanding of the $\lambda$-transition. We have used the chain rule of differentiation to express

\[ \nabla \phi(x) = \{i[\nabla \theta(x)]\rho + \nabla \rho(x)\} e^{i\theta(x)}. \]  

(5.16)

However, this rule cannot be applied here. Since $\theta(x)$ is the phase of the complex order field $\phi(x)$, it is a \textit{multivalued field}. At every point $x$ it is possible to add an arbitrary integer-multiple of $2\pi$ without changing $e^{i\theta(x)}$.

The correct chain rule is

\[ \nabla \phi(x) = \{i [\nabla \theta(x) - 2\pi \delta(x; S)]\} \rho(x) + \nabla \rho(x)\} e^{i\theta(x)} \]  

(5.17)

where $\delta(x; S)$ is the $\delta$-functions on the surface $S$ defined in (4.22) across which $\theta(x)$ jumps by $2\pi$. With this, we may approximate in the London limit:

\[ |\nabla \phi(x)|^2 \xrightarrow{\text{London limit}} |\phi|^2 [\nabla \theta(x) - \theta^\gamma(x)]^2 \]  

(5.18)

where we have introduced the field

\[ \theta^\gamma(x) \equiv 2\pi \delta(x, S). \]  

(5.19)

With this notation, the correct version of (5.11) reads

\[ H[\rho, \theta] = \frac{1}{2} \int d^3 x \left[ (\nabla \rho)^2 + \rho^2 (\nabla \theta - \theta^\gamma)^2 + \tau \rho^2 + \frac{\lambda}{2} \rho^4 \right], \]  

(5.20)
where
\[ \Theta^\prime(x) \equiv 2\pi \delta(x, S). \]  
(5.21)

In the London limit, the gradient energy density (5.14) must be corrected accordingly, so that the hydrodynamic limit of the Ginzburg-Landau Hamiltonian containing phonons and vortex lines reads, from now on in natural units with \( \rho_s/M = 1 \),

\[ H_{\psi}^{hy}[\Theta] = \frac{1}{2} \int d^3x \left( \nabla \Theta - \Theta^\prime \right)^2. \]  
(5.22)

This Hamiltonian density is obviously gauge invariant under deformations of the surface, under which \( \Theta^\prime(x) \) and \( \theta(x) \) change by

\[ \Theta^\prime(x) \rightarrow \Theta^\prime(x) + \nabla \Lambda_\delta^\prime(x), \quad \theta(x) \rightarrow \theta(x) + \Lambda_\delta^\prime(x), \]  
(5.23)

with the gauge functions

\[ \Lambda_\delta^\prime(x) = 2\pi \delta(x; V). \]  
(5.24)

Thus we encounter again the gauge transformations (4.28) and (4.29) of the gradient representation of magnetic fields. In the present context, the field \( \Theta^\prime(x) \) is called vortex gauge field.

In the sequel we shall see that all physical properties of the complex field theory can be found in the theory of the multivalued field \( \theta(x) \) with the vortex gauge-invariant Hamiltonian (5.22). Care has to be taken that all observable quantities are vortex gauge-invariant.

### 5.1.3 Vortex Density

As in the magnetic discussion in Section 4.2, the physical content of the vortex gauge field \( \Theta^\prime(x) \) appears when forming its curl. By Stokes’ theorem (4.23) we find the vortex density

\[ \nabla \times \Theta^\prime(x) \equiv j^\psi(x) = 2\pi \delta(x; L). \]  
(5.25)

As a consequence of Eq. (4.9), the vortex density satisfies the conservation law

\[ \nabla \cdot j^\psi(x) = 0, \]  
(5.26)

implying that vortex lines are closed.

The conservation law is a trivial consequence of \( j^\psi \) being the curl of \( \Theta^\prime \). It is therefore a Bianchi identity associated with the vortex gauge field structure.

The expression (5.22) is in general not the complete energy of a vortex configuration. It is possible to add a gradient energy in the vortex gauge field, which introduces an extra core energy to the vortex line. The extended Hamiltonian of the hydrodynamic limit of the Ginzburg-Landau Hamiltonian containing phonons and vortex lines with an extra core energy reads

\[ H_{vc}^{hy} = \int d^3x \left[ \frac{1}{2} (\nabla \Theta - \Theta^\prime)^2 + \frac{\epsilon_c}{2} (\nabla \times \Theta')^2 \right]. \]  
(5.27)
5.1 Superfluid Transition

The extra core energy does not destroy the invariance under vortex gauge transformations (5.23).

The core energy term is proportional to the square of a $\delta$-function which is highly singular. The singularity is a consequence of the hydrodynamic limit in which the field $\rho(x)$ in (5.11) is completely frozen at the minimum of (5.20). Moreover, the coherence length of the $\rho$-field is zero, and this is the origin of the above $\delta$-functions. With this in mind we may regularize the $\delta$-functions in the core energy physically by smearing them out over the actual small coherence length $\xi$ of the superfluid, which is of the order of a few rA. Whatever the size of $\xi$, the regularized last term yields an energy proportional to the total length of the vortex lines.

5.1.4 Partition Function

The partition function of the Nambu-Goldstone modes and all fluctuating vortex lines may be written as a functional integral

$$Z_{\text{hy}}^{\text{vc}} = \sum_{\{S\}} \int_{-\pi}^{\pi} D\theta e^{-\beta H_{\text{hy}}^{\text{vc}}},$$

(5.28)

where $\beta$ is the inverse temperature $\beta \equiv 1/T$ in natural units where the Boltzmann constant $k_B$ is equal to unity. The measure $\int_{-\pi}^{\pi} D\theta$ is defined by discretizing the space into a fine simple cubic lattice of spacing $a$ and integrating at each lattice point over all $\theta \in (-\pi, \pi)$, and taking the continuum limit $a \to 0$. The sum over all surface configurations $\sum_{\{S\}}$ is defined on the lattice by setting at each lattice point $x$

$$\theta_i^v(x; S) \equiv 2\pi n_i(x),$$

(5.29)

where $n_i(x)$ is an integer-valued version of the vortex gauge field $\theta_i^v(x; S)$, and by summing over all integer numbers $n_i(x)$:

$$\sum_{\{S\}} \equiv \sum_{\{n_i(x)\}}.$$

(5.30)

The partition function (5.28) is the continuum limit of the lattice partition function

$$Z_V = \sum_{\{n_i(x)\}} \left[ \prod_{x} \int_{-\pi}^{\pi} d\theta(x) \right] e^{-\beta H_V},$$

(5.31)

where $H_V$ is the lattice version of the Hamiltonian (5.27):

$$H_V = \frac{1}{2} \sum_x [\nabla \theta(x) - 2\pi n(x)]^2 + \frac{1}{2} \epsilon_c |\nabla \times n(x)|^2.$$  

(5.32)

Here the symbol $\nabla_i$ denotes lattice derivative which act on an arbitrary function $f(x)$ as

$$\nabla_i f(x) \equiv a^{-1}[f(x + a\epsilon_i) - f(x)],$$

(5.33)
where $\mathbf{e}_i$ are the unit vectors to the nearest neighbors in the plane, and $a$ is the lattice spacing. There exists also a conjugate lattice derivative

$$\nabla_i f(\mathbf{x}) \equiv a^{-1}[f(\mathbf{x}) - f(\mathbf{x} - a\mathbf{e}_i)],$$  

which arises in the lattice version of partial integration

$$\sum_{\mathbf{x}} f(\mathbf{x}) \nabla_i g(\mathbf{x}) = -\sum_{\mathbf{x}} [\nabla_i f(\mathbf{x})] g(\mathbf{x}),$$  

which holds for functions $f(\mathbf{x})$, $g(\mathbf{x})$ vanishing on the surface, or satisfying periodic boundary conditions. In Fourier space, the eigenvalues of $\nabla_i$, $\nabla_i$ are

$$K_i = \frac{e^{ik_a} - 1}{a},$$

$$K_i = \frac{1 - e^{-ika}}{a},$$

respectively, where $k_i$ are the wave numbers in the $i$-direction.

The lattice version of the Laplacian $\nabla^2$ is the **lattice Laplacian** $\nabla\nabla$. Its eigenvalues are in $D$ dimensions [15]

$$\mathbf{K}K = 2 \sum_{i=1}^{D} (1 - \cos k_i a),$$

where $k_i \in (-\pi/a, \pi/a)$ are the wave numbers in the Brillouin zone of the lattice [15]. In the continuum limit $a \to 0$, both lattice derivatives reduce to the ordinary derivative $\partial_i$, and $\mathbf{K}K$ goes over into $k^2$.

In the Hamiltonian (5.32), the lattice spacing $a$ has been set equal to unity, for simplicity.

In the partition function (5.31), the integer-valued vortex gauge fields $n_i(\mathbf{x})$ are summed without restriction. Alternatively, we may fix a gauge of $n_i(\mathbf{x})$ by some functional $\Phi[n]$, and obtain [7]

$$Z_V = \sum_{\{n_i(\mathbf{x})\}} \Phi[n] \left[ \prod_{\mathbf{x}} \int_{-\infty}^{\infty} d\theta(\mathbf{x}) \right] e^{-\beta H_V}.$$  

On the lattice we can always enforce the axial gauge [16]

$$n_3(\mathbf{x}) = 0.$$  

Note that in contrast to continuum gauge fields it is impossible to choose the Lorentz gauge $\nabla \cdot n(\mathbf{x}) = 0$.

In the formulation (5.37), the gauge freedom has been absorbed into the $\theta(\mathbf{x})$-field which now runs, for each $\mathbf{x}$, from $\theta = -\infty$ to $\infty$ rather than from $-\pi$ to $\pi$ in (5.31). This has the advantage that the integrals over all $\theta(\mathbf{x})$ can be done yielding

$$Z_V = \text{Det}^{1/2}[-\nabla\nabla^{-1}] \sum_{\{n_i(\mathbf{x})\}} \Phi[n] e^{-\beta H'_V},$$  

with

$$\beta H'_V = \sum_{\mathbf{x}} \left[ \frac{4\pi^2}{2} \left\{ n^2(\mathbf{x}) - [\nabla \cdot n(\mathbf{x})] \frac{1}{\nabla \nabla} [\nabla \cdot n(\mathbf{x})] \right\} + \frac{1}{2} \epsilon_c [\nabla \times n(\mathbf{x})]^2 \right].$$
5.1 Superfluid Transition

In calculating partition functions we shall always ignore trivial overall factors. If we introduce lattice curls of the integer-valued jump fields (5.29):

\[ \mathbf{l}(\mathbf{x}) \equiv \nabla \times \mathbf{n}(\mathbf{x}), \]

(5.41)

we can rewrite the Hamiltonian (5.40) as

\[ \beta H'_{\mathbf{V}} = \sum_{\mathbf{x}} \left[ \frac{4\pi^2}{2} \mathbf{l}(\mathbf{x}) \cdot \frac{1}{-\nabla \nabla} \mathbf{l}(\mathbf{x}) + \frac{\epsilon c \mathbf{l}^2(\mathbf{x})}{2} \right]. \]

(5.42)

Being lattice curls, the fields \( \mathbf{l}(\mathbf{x}) \) satisfy \( \nabla \cdot \mathbf{l}(\mathbf{x}) = 0 \). They are, of course, integer-valued versions of the vortex density \( j^v(\mathbf{x})/2\pi \) defined in Eq. (5.25). The energy (5.42) is the interaction energy between the vortex loops.

The inverse lattice Laplacian \(-\nabla \nabla^{-1}\) in (5.31) and (5.42) is the lattice version of the inverse Laplacian \(-\nabla^2\) whose local matrix elements \( \langle \mathbf{x}_2 | -\nabla^2 | \mathbf{x}_1 \rangle \) yield the Coulomb potential of the coordinate difference \( \mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1 \):

\[ V_0(\mathbf{x}) \equiv \int \frac{d^3k}{a^3(2\pi)^3} \frac{e^{-i\mathbf{k} \cdot \mathbf{x}}}{4\pi r} = \frac{1}{4\pi r}, \quad r \equiv |\mathbf{x}|. \]

(5.43)

The corresponding matrix elements on the lattice \( \langle \mathbf{x}_2 | -\nabla \nabla^{-1} | \mathbf{x}_1 \rangle \) are given by

\[ v_0(\mathbf{x}) = \int_{BZ} \frac{d^3k}{a^3(2\pi)^3} \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{4\pi r} = \frac{1}{a} \left[ \prod_{i=1}^3 \int_{-\pi/a}^{\pi/a} \frac{d^3(ak_i)e^{i\mathbf{k}_i \cdot \mathbf{x}_i}}{(2\pi)^3} \right] \frac{1}{2 \sum_{i=1}^3 (1 - \cos ak_i)}, \]

(5.44)

where the subscript BZ of the momentum integral indicates the restriction to the Brillouin zone.

The lattice Coulomb potential (5.44) is the zero-mass limit of the lattice Yukawa potential

\[ v_m(\mathbf{x}) = \frac{1}{a} \left[ \prod_{i=1}^3 \int_{-\pi/a}^{\pi/a} \frac{d^3(ak_i)e^{i\mathbf{k}_i \cdot \mathbf{x}_i}}{(2\pi)^3} \right] \frac{1}{2 \sum_{i=1}^3 (1 - \cos ak_i) + m^2a^2}, \]

(5.45)

whose continuum limit is the ordinary Yukawa potential

\[ V_m(r) = \int \frac{d^3k}{(2\pi)^3} \frac{e^{-i\mathbf{k} \cdot \mathbf{x}}}{\mathbf{k}^2 + m^2} \frac{1}{4\pi r} = \frac{e^{-mr}}{4\pi r}, \quad r \equiv |\mathbf{x}|. \]

(5.46)

In terms of the lattice Coulomb potential, we can write the partition function (5.39) for zero extra core energy as

\[ Z_V = \text{Det}^{1/2} \left[ \tilde{v}_0 \right] \sum_{\mathbf{l}, \mathbf{l'}=0} e^{-(4\pi^2\beta \alpha/2)\sum_{\mathbf{x}, \mathbf{x'}} \mathbf{l}(\mathbf{x}) \cdot \mathbf{l}(\mathbf{x}) \mathbf{l}(\mathbf{x'})}, \]

(5.47)

where \( \tilde{v}_0 \) abbreviates the operator \(-\nabla \nabla^{-1}\).
The momentum integrals over the different lattice directions can be done separately by applying the Schwinger trick to express the denominator as an integral over an exponential

\[ \frac{1}{a} = \int_0^\infty ds \, e^{-sa}, \]  

so that

\[ \frac{1}{2 \sum_{i=1}^3 (1 - \cos ak_i) + m^2 a^2} = \int_0^\infty ds \, e^{-(6+m^2a^2)s} \prod_{i=1}^3 e^{\cos k_ia} . \]  

Using now the integral representation of the modified Bessel functions of the first kind

\[ I_n(z) = \int_{-\pi}^{\pi} d\kappa e^{iz + \cos \kappa z}, \]  

we find

\[ v_m(x) = \frac{1}{a} \int_0^\infty ds \, e^{-(6+m^2a^2)s} I_{x_1/a}(2s) I_{x_2/a}(2s) I_{x_3/a}(2s). \]  

In contrast to the continuum version (5.46), the lattice potential \( v_m(x) \) is finite at the origin. The values of \( v_m(0) \) as a function of \( m^2 a^2 \) are plotted in Fig. 5.5. The Coulomb potential has the value \( v_0(0) \approx 0.2527/a \) [17].

**Figure 5.5** Lattice Yukawa potential at the origin and the associated Trace log. The plot shows the subtracted expression \( \text{Tr} \log(-\nabla \nabla + m^2)/N + \log(6/a^2 + m^2) \), where \( N \) is the number of sites on the lattice.

The functional determinant of the lattice Laplacian appearing in (5.39) and (5.42) as a prefactor can be calculated easily from the Yukawa potential. We simply use the relation

\[ \text{Det}^{-1/2}(-\nabla \nabla + m^2) = \text{Det}^{1/2}(\hat{v}_m) = e^{-\frac{1}{2} \text{Tr} \log(-\nabla \nabla + m^2)} = e^{-\frac{a^3}{2} \sum_x \log(-\nabla \nabla + m^2)|x|}, \]  

and calculate

\[ \text{Tr} \log(-\nabla \nabla + m^2) = \frac{a^3}{2} \int dm^2 \sum_x (x|(-\nabla \nabla + m^2)^{-1}|x) = \frac{Na^3}{2} \int dm^2 \, v_m(0), \]  

where \( N \) is the number of lattice sites and the constant is the limes \( m^2 \to 0 \) of \( \log m^2 \). Performing the integral over \( m^2 \) in (5.51), we obtain

\[ a^3 \int dm^2 \, v_m(0) = -\int_0^\infty \frac{ds}{s} e^{-(6+m^2a^2)s} I_0^2(2s). \]
5.1 Superfluid Transition

The divergence at \( s = 0 \) can be removed by subtracting a similar integral representation

\[
a^2 \int dm^2 (6 + m^2 a^2)^{-1} = - \int_0^\infty \frac{ds}{s} e^{-(6+m^2a^2)s}. \tag{5.55}
\]

Thus we obtain the finite result

\[
\frac{1}{N} \text{Tr} \log(-\nabla^2 + m^2) - \log(6/a^2 + m^2) = - \int_0^\infty \frac{ds}{s} e^{-(6+m^2a^2)s} \left[ t_0^3(2s) - 1 \right]. \tag{5.56}
\]

The \( m^2 \)-behavior of this expression is displayed in Fig. 5.5.

In the form (5.47) it is easy to perform a graphical expansion of the partition function adding terms of longer and longer loops each term carrying a Boltzmann factor \( e^{-\beta \text{const}/2} \). This expansion converges fast for low temperatures. As the temperature is raised, fluctuations create more and longer loops. If there is no extra core energy \( \epsilon_c \), the loops become infinitely long and dense at a critical value \( T_c = 1/\beta_c = 1/0.33 \equiv 3 \), where the sum in (5.39) diverges. At that point the system is filled with vortex loops. Since the inside of each vortex loop consists of normal fluid, this condensation of vortex loops makes the entire fluid normal. See Fig. 5.4 for visualizing this condensation process. The successive orders of the associated specific heat are plotted in Fig. 5.6.

Without the extra core energy, \( Z_V \) defines the famous Villain model [18], a discrete Gaussian approximation to the so-called XY-model whose Hamiltonian is

\[
H_{XY} = \sum_x \sum_{i=1}^3 \cos[\nabla_i \theta(x)]. \tag{5.57}
\]

Both the XY-model and the Villain model can be simulated with Monte Carlo techniques on a computer and displays a second-order phase transition for \( \beta_c \approx 0.33 \). The critical exponents of the two models coincide. The resulting specific heat of the Villain model is shown in Fig. 5.6. It has the typical \( \lambda \)-shape observed in \(^4\text{He} \) in Fig. 5.1.

By analogy with the lattice formulation, we fix the gauge in the continuum partition function (5.28), with the energy (5.22) or (5.27), by inserting a gauge-fixing functional \( \Phi[\theta^x] \). The axial gauge is fixed by the \( \delta \)-functional

\[
\Phi[\theta^x] = \delta[\theta^x_3]. \tag{5.58}
\]

Note that since the partition function (5.28) contains the sum over the vortex gauge fields \( \theta^x \), it describes superfluid \(^4\text{He} \) not only at zero temperature, where the Nambu-Goldstone modes were identified, but at any not too large temperature. In particular, the temperature regime around the superfluid phase transition is included.

The vortex gauge field extends the partition function of fluctuating Nambu-Goldstone modes in the same way as the size of the order field \( \psi \) does in a Landau description of the phase transition. In fact, it is easy to show that near the transition, the partition function (5.28) can be transformed into a \( |\psi|^4 \)-theory of the Landau type [7].
Figure 5.6 Specific heat of Villain model in three dimensions plotted against $\beta = 1/T$ in natural units. The $\lambda$-transition is seen as a sharp peak, with properties near $T_c$ similar to the experimental curve in Fig. 5.1. The solid curves stem from analytic expansion in powers of $T \equiv 1/\beta$ (low-temperature or weak-coupling expansion) and in power of $T^{-1} = \beta$ (high-temperature or strong-coupling expansion) (see Ref. [19]).

5.1.5 Continuum Derivation of Interaction Energy

Let us calculate the interaction energy (5.42) between vortex loops once more in the continuum formulation. Omitting the core energy, for simplicity, the partition function with a fixed vortex gauge is given by

$$Z_{hv}^{hy} = \sum_{\{s\}} \Phi[\theta^\nu] \int_{-\infty}^{\infty} D\theta e^{-\beta H_{hv}^{hy}},$$

(5.59)

where

$$H_{hv}^{hy} = \frac{1}{2} \int d^3x (\nabla \theta - \theta^\nu)^2$$

(5.60)

is the energy (5.27) without core energy. Let us expand this into two parts

$$H_{hv}^{hy} = \frac{1}{2} \int d^3x \left[ (\nabla \theta)^2 + 2\theta \nabla \theta^\nu + \theta^\nu \right] = H_{hv1}^{hy} + H_{hv2}^{hy},$$

(5.61)

where

$$H_{hv1}^{hy} = \frac{1}{2} \int d^3x \theta \left( \nabla + \frac{1}{-\nabla^2} \nabla \cdot \theta^\nu \right) \left( \nabla + \frac{1}{-\nabla^2} \nabla \cdot \theta^\nu \right)$$

(5.62)

and

$$H_{hv2}^{hy} = \frac{1}{2} \int d^3x \left( \theta^\nu \nabla \cdot \theta^\nu \frac{1}{-\nabla^2} \nabla \cdot \theta^\nu \right).$$

(5.63)

Inserting this into (5.59), we can perform the Gaussian integrals over $\theta(\mathbf{x})$ at each $\mathbf{x}$ using the generalization of the Gaussian formula

$$\int_{-\infty}^{\infty} \frac{d\theta}{2\pi} e^{-a(\theta-c)^2/2} = a^{-1/2}$$

(5.64)
to fields $\theta(x)$ and differential operators $\hat{O}$ in $x$-space

\[
\int_{-\infty}^{\infty} \mathcal{D}\theta e^{-\int d^3x \left[ \theta(x)-c(x) \right] \hat{O}[\theta(x)-c(x)]/2} = [\text{Det} \hat{O}]^{-1/2}.
\]

(5.65)

Applying this formula to (5.59), we obtain

\[
Z_{hv}^\text{hy} = [\text{Det}(-\nabla^2)]^{-1/2} \sum_{\{\Phi\}} \Phi[v] e^{-\beta H_v},
\]

(5.66)

where $H_v$ is the interaction energy of the vortex loops corresponding to (5.42):

\[
H_v = \frac{1}{2} \int d^3x (\nabla \times \theta)^2 - \frac{1}{\nabla^2} (\nabla \times \theta') = \frac{1}{2} \int d^3x j^v \cdot \frac{1}{-\nabla^2} j^v
\]

\[
= \frac{1}{8\pi} \int d^3x d^3x' \frac{j^v(x)}{|x-x'|} j^v(x').
\]

(5.67)

This has the same form as the magnetic Biot-Savart energy (4.92) for current loops, implying that parallel vortex lines repel each other [as currents would do if no work were required to keep them constant against the inductive forces, which reverses the sign. Recall the discussion of the free magnetic energy (4.94)]. On a lattice, the partition function (5.66) takes once more, the form (5.39).

The process of removing some variables from a partition function by integration will occur frequently in the sequel and will be referred to as integrating out. It will be used also in discussions of Hamiltonians without writing always down the associated partition function in which the integrals are actually performed.

5.1.6 Physical Jumping Surfaces

The invariance of the energy (5.27) under vortex gauge transformations guarantees the physical irrelevance of the jumping surfaces $S$. If we destroy this invariance, the surfaces become physical objects. This may be done by destroying the original U(1)-symmetry explicitly. This will give a mass to the Nambu-Goldstone modes. To lowest approximation, it adds to the energy (5.27) without core energy a mass term $m^2\theta(x)^2$:

\[
H_{vim}^\text{by} = \frac{1}{2} \int d^3x \left\{ [\nabla \theta(x)]^2 + m^2\theta(x)^2 \right\}.
\]

(5.68)

The mass term gives an energy to the surfaces $S$. To see this we write the energy as

\[
H_{vim}^\text{by} = \frac{1}{2} \int d^3x \left\{ [\nabla \theta]^2 + m^2\theta^2 - \theta \nabla \cdot \theta + \theta^2 \right\},
\]

(5.69)

and decompose this into two parts as in (5.61):

\[
H_{vim1}^\text{by} = \frac{1}{2} \int d^3x \left( \theta + \frac{1}{-\nabla^2 + m^2} \nabla \cdot \theta' \right) (-\nabla^2 + m^2) \left( \theta + \frac{1}{-\nabla^2 + m^2} \nabla \cdot \theta' \right).
\]

(5.70)
and
\[ H_{vm}^{by} = \frac{1}{2} \int d^3 x \left( \Theta^2 - \nabla \cdot \Theta^v - \frac{1}{-\nabla^2 + m^2} \nabla \cdot \Theta^v \right). \] (5.71)

The Gaussian integrals over \( \theta(x) \) can be done as before, the partition function (5.66) becomes
\[ Z_{vm}^{by} = \text{Det}^{-1/2} \left[ -\nabla^2 + m^2 \right] \sum_{\{S\}} \Phi[\Theta^v] e^{-\beta H_{vm}^{by}}. \] (5.72)

The energy \( H_{vm}^{by} \) in the exponent can be rewritten as
\[ H_{vm}^{by} = \frac{1}{2} \int d^3 x \left[ (\nabla \times \Theta^v)^2 - \nabla^2 + m^2 (\nabla \times \Theta^v) \right] \] (5.73)

The presence of the mass \( m \) changes the long-range Coulomb-like interaction \( 1/R \) in Eq. (5.67) into a finite-range Yukawa-like interaction \( e^{-mR}/R \).

The second term in (5.73),
\[ H_{Sm} = \frac{m^2}{2} \int d^3 x \left( \Theta^v \frac{1}{-\nabla^2 + m^2} \Theta^v \right) = \frac{1}{8\pi} \int d^3 x d^3 x' \left( \frac{e^{-m|x-x'|}}{|x-x'|} \Theta^v(x') \right). \] (5.75)

is of a completely new type. It describes a Yukawa-like interaction between the normal vectors of the surface elements, and gives rise to a field energy within a layer of thickness \( 1/m \) around the surfaces \( S \). As a consequence, the surfaces acquire tension. Their shape is no longer irrelevant, but for a given set of vortex loops \( L \) at their boundaries, the surface will span minimal surfaces. For \( m = 0 \), the tension disappears and the shape of the surface becomes again irrelevant, thus restoring vortex gauge invariance.

This mechanism of generating surface tension will be used in Chapter 8 to construct a simple model of quark confinement.

### 5.1.7 Canonical Representation of Superfluid

We can set up an alternative representation of the partition function of the superfluid in which the vortex loops are more directly described by their physical vortex density, instead of their jumping surfaces \( S \). This is possible by eliminating the Nambu-Goldstone modes in favor of a new gauge field. It is canonically conjugate to the angular field \( \theta \) and called generically stress gauge field [7]. In the particular case
of the superfluid under discussion it is a *gauge field of superflow*. Recall that the canonically conjugate momentum variable $p(t)$ in an ordinary path integral [20]

$$\int \mathcal{D}x \exp \left( -\frac{M}{2} \int_{t_a}^{t_b} dt \dot{x}^2 \right)$$

(5.76)

is introduced by a quadratic completion, rewriting (5.76) as

$$\int \mathcal{D}x \mathcal{D}p \exp \left[ \int_{t_a}^{t_b} dt \left( ip\dot{x} - \frac{p^2}{2M} \right) \right].$$

(5.77)

By analogy, we introduce a canonically conjugate vector field $b(x)$ to rewrite the partition function (5.59) as

$$Z^h_v = \int_{-\infty}^{\infty} \mathcal{D} b \sum_{\{ S \}} \Phi[\theta^v] \int_{-\infty}^{\infty} \mathcal{D} \theta e^{-\beta \bar{H}^h_v}$$

(5.78)

where [21]

$$\beta \bar{H}^h_v = \int d^3x \left\{ \frac{1}{2\beta} b^2(x) - i b \left[ \nabla \theta(x) - \theta^v(x) \right] \right\}.$$  

(5.79)

In principle, the gradient energy could contain higher powers of $\partial_i \theta$. Then the canonical representation (5.79) would contain more complicated functions of $b_i(x)$.

Note that if we go over to a Minkowski space formulation in which $x^0 = -ix^3$ plays the role of time, the integral

$$\int_{-\infty}^{\infty} \mathcal{D} b_0 e^{-ib_0(\mathbf{x})\partial_0\theta(\mathbf{x})}$$

(5.80)

creates, on a discretized time axis, a product of $\delta$-functions

$$\langle \theta_{n+1} | \theta_n \rangle \langle \theta_n | \theta_{n-1} \rangle \langle \theta_{n-1} | \theta_{n-2} \rangle$$

(5.81)

with

$$\langle \theta_n | \theta_{n-1} \rangle = \delta_n(\theta_n - \theta_{n-1}).$$

(5.82)

These can be interpreted as Dirac scalar products in the Hilbert space of the system. On this Hilbert space, there exists an operator $\hat{b}_0(\mathbf{x})$ whose zeroth component is given by

$$\hat{b}_0 = -i\partial_0$$

(5.83)

and satisfies the equal-time canonical communication rule

$$[\hat{b}_0(\mathbf{x}_\perp, x_0), \theta(\mathbf{x}'_\perp, x_0)] = -i\delta^{(2)}(\mathbf{x}_\perp - \mathbf{x}'_\perp),$$

(5.84)

where $\mathbf{x}_\perp = (x^1, x^2)$ denotes the spatial components of the vector $(x^0, x^1, x^2)$. The charge associated with $\hat{b}_0(\mathbf{x})$,

$$\dot{Q}(x_0) = \int d^2x \hat{b}_0(\mathbf{x}_\perp, x_0),$$

(5.85)
generates a constant shift in $\theta$:

$$e^{-i\alpha \hat{Q}(x_0)} \theta(x_\perp, x_0) e^{i\alpha \hat{Q}(x_0)} = \theta(x_\perp, x_0) + \alpha.$$  \hspace{1cm} (5.86)

Thus it multiplies the original field $e^{i\theta(x)}$ by a phase factor $e^{i\alpha}$. The charge $\hat{Q}(x_0)$ is the generator of the U(1)-symmetry transformation whose spontaneous breakdown is responsible for the Nambu-Goldstone nature of the fluctuations of $\theta(x)$. Since the original theory is invariant under the transformations $\phi \rightarrow e^{i\alpha} \phi$, the energy (5.79) does not depend on $\theta$ itself, but only on $\partial_\theta$.

In the partition function (5.78) we may use the formula

$$\int_{-\infty}^{\infty} \mathcal{D}\theta \ e^{i \int d^3 x f(x) \theta(x)} = \delta[f(x)],$$  \hspace{1cm} (5.87)

and obtain from $f(x) = \nabla \cdot b(x)$ the conservation law

$$\nabla \cdot b(x) = 0,$$  \hspace{1cm} (5.88)

implying that $\hat{Q}(x_0)$ is a time-independent charge and $e^{i\alpha \hat{Q}}$ is a symmetry transformation.

If the energy in (5.78) would depend on $\theta$ itself, then the charge $\hat{Q}(x_0)$ would no longer be time-independent. However, it would still generate the above U(1)-transformation.

In general, the conjugate variable to the phase angle of a complex field is the particle number (recall Subsection 3.5.3). This role is played here by $\hat{Q}(x_0)$ which counts the number of particles in the superfluid. Thus we may identify the vector field $b(x)$ as the particle current density of the superfluid condensate:

$$j_s(x) \equiv b(x),$$  \hspace{1cm} (5.89)

also called the supercurrent density of the superfluid.

After integrating out the $\theta$-fields in the partition function (5.78), we can also perform the sum over all surface configurations of the vortex gauge field $\theta^v(x)$. For this we employ the following useful formula applicable to any function $b(x)$ with $\nabla \cdot b(x) = 0$:

$$\sum_{\{S\}} e^{2\pi i \int d^3 x \delta(x; S)b(x)} = \sum_{\{L\}} \delta[b(x) - \delta(x; L)].$$  \hspace{1cm} (5.90)

This can easily be proved by going on a lattice where this formula reads [recall (5.30)]

$$\sum_{\{n_i\}} e^{2\pi i \sum_{x_i} f_i(x)} = \sum_{\{m_i\}} \prod_i \delta(f_i(x) - m_i(x)), $$  \hspace{1cm} (5.91)

and using for each $x, i$ the Poisson formula [20]

$$\sum_n e^{2\pi i nf} = \sum_m \delta(f - m).$$  \hspace{1cm} (5.92)
Then we obtain for (5.78) the following alternative representation

\[ Z^\text{hy} = \sum_{\{L\}} e^{-\int d^3x \, b^2/2\beta}, \]  

(5.93)

where \( b = \delta(x; L) \). On the lattice, this partition function becomes

\[ Z^\text{hy} = \sum_{b, \nabla \cdot b = 0} e^{-\sum_x b^2/2\beta}, \]  

(5.94)

where \( b(x) \) is now an integer-valued divergenceless field representing the closed lines of superflow.

The partition function (5.94) can be evaluated graphically adding terms of longer and longer loops each term carrying a Boltzmann factor \( e^{-\text{const}/2\beta} \). This expansion converges fast for high temperature. The specific heat following from the lowest approximations obtained in this way are plotted in Fig. 5.6. For very high temperature, there is no loop of superflow. As the temperature is lowered, fluctuations create more and longer loops. The system has become superfluid.

Note that the superflow partition function (5.94) look quite similar to the vortex loop partition function (5.47). Both contain the same type of sum over non-self-backtracking loops. The main difference is the long-range Coulomb interaction between the loop elements. Suppose we forget for a moment the non-local parts of the Coulomb interaction and approximate the vortex loop partition function (5.47) keeping only the self-energy part of the Coulomb interaction:

\[ Z_V^\text{app} = \text{Det}^{1/2}[\hat{\nabla}_m] \sum_{1, \nabla \cdot l = 0} e^{-(4\pi^2/\beta a)\nu_0(0) \sum_x \Phi(x)}. \]  

(5.95)

Apart from a constant overall factor, this approximation coincides with the superflow partition function (5.94). By comparing the prefactors of the energy we see that (5.95) has a second-order phase transition at

\[ 4\pi^2 a/\beta v_0(0) \approx T_c \approx 3. \]  

(5.96)

This is solved by \( \beta \approx 3/4\pi^2 a v_0(0) \approx 0.30 \), corresponding to a critical temperature

\[ T_c^\text{app} \approx \frac{4\pi^2 a v_0(0)}{3} \approx 3.3. \]  

(5.97)

This is only 10% larger than the accurate value \( T_c = 1/\beta_c \approx 3 \), so that we conclude that the nonlocal parts of the Coulomb interaction in (5.47) have little effect upon the transition temperature.


5.1.8 Yukawa Loop Gas

The above observation allows us to estimate the transition temperature in partition function closely related to (5.47)

\[
Z_Y^V = \text{Det}^{1/2}[\hat{v}_m] \sum_{1,\nabla l=0} e^{-\left(4\pi^2 a^2/2\right)\Sigma_{x,x'} l(x)v_m(x-x') l(x')},
\]

(5.98)

where \( v_m(x) \) is the lattice version of the Yukawa potential (5.46), and \( \hat{v}_m \) the associated operator \( (-\nabla^2 + m^2)^{-1} \).

Performing also here the local approximation of the type (5.95),

\[
Z_{Y \text{ app}}^V = \text{Det}^{1/2}[\hat{v}_0] \sum_{1,\nabla l=0} e^{-\left(4\pi^2 a^2/2\right)v_m(0)\Sigma_{x} l^2(x)},
\]

(5.99)

we estimate the critical value \( \beta_{m,c} \) of the Yukawa loop gas by the equation corresponding to (5.96):

\[
4\pi^2 a\beta_{m,c} v_m(0) \approx T_c \approx 3.
\]

(5.100)

Since the Yukawa potential becomes more and more local for increasing \( m \), the local approximation (5.99) becomes exact. Thus we conclude that the error in the estimating of the critical temperature \( T_{m,c} = 1/\beta_{m,c} \) from Eq. (5.100) drops from 10% at \( m = 0 \) to zero as \( m \) goes to infinity. We have plotted the resulting critical values of \( T_{m,c} = 1/\beta_{m,c} \) in Fig. 5.7.

![Figure 5.7](image)

**Figure 5.7** Critical temperature \( T_{m,c} = 1/\beta_{m,c} \) of a loop gas with Yukawa interactions between line elements, estimated by Eq. (5.100). The error is with 10% the largest at \( m = 0 \), and decreases to zero for increasing \( m \). The dashed curve is the analytic approximation (5.104).

We conclude that the Yukawa loop gas (5.98) has a second-order phase transition as the Villain- and the XY-models. The critical exponents of the Yukawa loop gas are all of the same as those of the Villain-model, and thus also of the XY-model. In the terminology of the theory of critical phenomena, the Yukawa loop gases lie for all \( m \) in the same universality class as the XY-model.

It is possible to find a simple analytic approximation for the critical temperature plotted in Fig. 5.7. For this we use the so-called *hopping expansion* [7] of the lattice
5.1 Superfluid Transition

Yukawa potential (5.1.8). It is found by expanding the modified Bessel function \( I_{n/a}(2s) \) in Eq. (5.51) in powers of \( s \) using the series representation

\[
I_n(2s) = \sum_{k=0}^{\infty} \frac{s^{2k}}{k!\Gamma(n+k+1)}.
\]

(5.101)

At the origin \( x = 0 \), the integral over \( s \) yields the expansion

\[
v_m(0) = \frac{1}{a} \sum_{n=0,2,4} \frac{H_n}{(m^2a^2+6)^{n+1}}, \quad H_0 = 1, H_2 = 6, \ldots.
\]

(5.102)

To lowest order, this implies the approximate ratio \( v_m(0)/v_0(0) \equiv 1/(m^2a^2/6 + 1) \). A somewhat more accurate fit to the ratio

\[
\frac{v_m(0)}{v_0(0)} \approx \frac{1}{\sigma m^2a^2/6 + 1}, \quad \text{with} \quad \sigma \approx 1.6.
\]

(5.103)

Together with (5.97) this leads to the analytic approximation

\[
T_{m,c} = \frac{1}{\beta_{m,c}} \approx \frac{4\pi^2av_m(0)}{3} \approx \frac{4\pi^2av_0(0)}{3} \frac{1}{\sigma m^2a^2/6 + 1}.
\]

(5.104)

A comparison with the numerical evaluation of (5.100) is shown in Fig. 5.7. The fit has only a 10% error for \( m = 0 \) and becomes accurate for large \( m \).

5.1.9 Gauge Field of Superflow

The current conservation law \( \nabla \cdot \mathbf{b}(x) = 0 \) can be ensured automatically as a Bianchi identity, if we represent \( \mathbf{b}(x) \) as a curl of a gauge field of superflow

\[
\mathbf{b}(x) = \nabla \times \mathbf{a}(x).
\]

(5.105)

With respect to the gauge transformations (5.107), the current conservation law \( \nabla \cdot \mathbf{b}(x) = 0 \) is a Bianchi identity.

The energy (5.79), with the core energy reinserted, goes over into what is called the *dual representation*:

\[
\beta H_{avc} = \int d^3x \left[ \frac{1}{2\beta} (\nabla \times \mathbf{a})^2 + i\mathbf{a} \cdot (\nabla \times \mathbf{v}) + \frac{\beta \varepsilon}{2} (\nabla \times \mathbf{v})^2 \right].
\]

(5.106)

The energy is now double-gauge invariant. Apart from the invariance under the vortex gauge transformation (5.23), there is now the additional invariance under the gauge transformations of superflow

\[
\mathbf{a}(x) \rightarrow \mathbf{a}(x) + \nabla \Lambda(x),
\]

(5.107)

with arbitrary functions \( \Lambda(x) \).
The energy (5.106) can be expressed in terms of the vortex density of Eq. (5.25) as
\[
\beta H_{\text{vec}}' = \int d^3x \left[ \frac{1}{2\beta} (\nabla \times a)^2 + i a \cdot j + \frac{\beta c}{2} j^2 \right].
\] (5.108)

In this expression, the freely deformable jumping surfaces have disappeared and the energy depends only on the vortex lines. For a fixed set of vortex lines along \(L\), the action (5.108) has a similar form as the free magnetic energy of a given current distribution in Eq. (5.108). The only difference is a factor \(i\). Around a vortex line, the field \(b(x) = \nabla \times a(x)\) looks precisely like a magnetic field \(B(x) = \nabla \times A(x)\) around a current line, except for the factor \(i\). Extremizing the energy in \(a\) and reinserting the extremum yields once more the Biot-Savart interaction energy of the form Eq. (5.67) which is of the form (4.92) [not (4.92) due to the factor \(i\)].

If we want to express the partition function (5.78) in terms of the gauge field of superflow \(a(x)\), we must fix its gauge. Here we may choose the transverse gauge:
\[
\Phi_T[a] = \delta[\nabla \cdot a],
\] (5.109)

and the partition function (5.78) becomes
\[
Z_{hy}^b = \int D a \Phi_T[a] \sum_{\{S\}} \Phi[\theta]\, e^{-\beta H_{\text{vec}}'}. \tag{5.110}
\]

In terms of the Hamiltonian (5.108), the partition function becomes a sum over vortex lines \(L\):
\[
Z_{hy}^b = \int D a \Phi_T[a] \sum_{\{L\}} \Phi_T[j^\nu] e^{-\beta H_{\text{vec}}'}. \tag{5.111}
\]

where
\[
\Phi_T[j^\nu] = \delta[\nabla \cdot j^\nu]. \tag{5.112}
\]

ensures the closure of the vortex lines.

Note that if the energies \(H_{hy}^b\) or \(H_{hy}^\nu\) in (5.27) and (5.60) contain an explicit \(\theta\)-dependent term, such as the mass term in the Hamiltonian (5.68), there exists no reformulation of the \(\theta\)-fluctuations in terms of a gauge field \(a\). For a mass term, the formula (5.87) turns into
\[
\int_{-\infty}^{\infty} D \theta \, e^{-\int d^3x \left[ \frac{\beta m^2 \theta^2(x)}{2} - i f(x)\theta(x) \right]} = \delta_m[f(x)], \tag{5.113}
\]

where \(\delta_m[f(x)]\) denotes the softened \(\delta\)-functional
\[
\delta_m[f(x)] \propto e^{-\int d^3x f(x)/2\beta m^2}. \tag{5.114}
\]

For \(f(x) = \nabla \cdot b(x)\) this implies that \(b(x)\) is no longer purely transverse, as in (5.88). Hence it no longer possesses a curl representation (5.105).
5.1 Superfluid Transition

5.1.10 Disorder Field Theory

Let us discuss the sum over all vortex line configurations in the partition function (5.111) separately. For this we define an $a$-dependent vortex partition function

$$Z^v[a] = \sum_{\{L\}} \delta[\nabla \cdot j^v] \exp \left[ -\int d^3x \left( \frac{\beta \epsilon_c}{2} j^v_2 - i a \cdot j^v \right) \right].$$

(5.115)

It is possible to express this with the help of an auxiliary fluctuating vortex gauge field $\tilde{\theta}^v(x)$ singular on surfaces $\tilde{S}$ as a sum over auxiliary surface configurations $\tilde{S}$ as follows

$$Z^v[a] = \sum_{\{\tilde{S}\}} \int D\tilde{j}^v \delta[\nabla \cdot j^v] \exp \left\{ -\int d^3x \left[ \frac{\beta \epsilon_c}{2} j^v_2 - i j^v \cdot (\tilde{\theta}^v + a) \right] \right\}.$$  

(5.116)

In this expression, $j^v$ is an ordinary field. The sum over all $\tilde{S}$ configuration ensures via formula (5.91) that $j^v$ is a superposition of $\delta$-functions on lines $\bar{L}$, so that (5.116) agrees with (5.115).

Next we introduce an auxiliary field $\tilde{\theta}$, and rewrite the $\delta$ functional of the divergence of $j^v$ as a functional Fourier integral, so that we obtain the identity

$$Z^v[a] = \sum_{\{\tilde{S}\}} \int D\tilde{j}^v \int D\tilde{\theta} \exp \left\{ -\int d^3x \left[ \frac{\beta \epsilon_c}{2} j^v_2 + i j^v \cdot (\nabla \tilde{\theta} - \tilde{\theta}^v - a) \right] \right\}.$$  

(5.117)

Now $j^v$ is a completely unrestricted ordinary field. It can therefore be integrated to yield

$$Z^v[a] = \sum_{\{\bar{L}\}} \int D\tilde{\theta} \exp \left[ -\frac{1}{2\beta \epsilon_c} \int d^3x (\nabla \tilde{\theta} - \tilde{\theta}^v - a)^2 \right].$$  

(5.118)

Remembering the derivation of the Hamiltonian (5.22) from the hydrodynamic limit of the Ginzburg-Landau $|\phi|^4$ theory (5.6), we may interprete (5.118) as the partition function of the hydrodynamic limit of another U(1)-invariant Ginzburg-Landau theory with partition function

$$\hat{Z}^v[a] = \int D\psi D\psi^* \exp \left\{ -\frac{1}{2\beta} \int d^3x \left[ (\nabla - ia) \psi^* \psi^2 + m^2 |\psi|^2 + \frac{g}{2} |\psi|^4 \right] \right\},$$  

(5.119)

where $\psi(x)$ is another complex field $\psi$ with a $|\psi|^4$ interaction. Inserting this into (5.111) we obtain the combined partition function

$$Z^v_{\text{hy}} = \int_{-\infty}^{\infty} Da \Phi_T[a] \hat{Z}^v[a] \exp \left\{ -\int d^3x \left[ \frac{1}{2\beta} (\nabla \times a)^2 \right] \right\},$$  

(5.120)

which defines the desired disorder field theory.
The representation of ensembles of lines in terms of a single disorder field is the Euclidean version of what is known as second quantization in the quantum mechanics of many-particle systems.

At high temperature, the mass term \( m^2 \) of the \( \psi \)-field is negative and the disorder field acquires a nonzero expectation value \( \psi_0 = \sqrt{-m^2/g} \). Setting, as in (5.10),

\[
\psi(x) = \tilde{\rho}(x) e^{i\tilde{\theta}(x)}
\]  

(5.121)

and freezing out the fluctuations of \( \rho \) leads directly to the partition function (5.118).

The disorder field theory possesses similar vortex lines as the original Ginzburg-Landau theory with Hamiltonian (5.6), or its hydrodynamic limit (5.27). But in contrast to it, the fluctuations of the disorder field are “frozen out” at high temperature, as we can see from the prefactor \( 1/\beta \) in the exponents of (5.118) and (5.119), and the partition function (5.119) reduces to (5.118) in the hydrodynamic limit. As before in (5.59) we may perform the functional integral over \( \tilde{\theta} \). Here this removes the longitudinal part of \( \tilde{\theta} - a \), and (5.118) becomes

\[
Z^v[a] = \exp \left[ -\frac{m_a^2}{2\beta} \int d^3x \left( \tilde{\theta} - a \right)_T^2 \right]
\]  

(5.122)

where

\[
m_a^2 = \frac{1}{\epsilon_c},
\]  

(5.123)

and \( v_T \) denotes transverse part of the vector field \( v \). This and the longitudinal part \( v_L \) are defined by

\[
v_{T_i} \equiv \left( \delta_{ij} - \nabla_i \nabla_j \right) v_j, \quad v_{L_i} \equiv \nabla_i v_j.
\]  

(5.124)

At high temperatures, where the disorder field \( \psi \) has no vortex lines \( \tilde{L} \) (while the order field \( \phi \) has many vortex lines \( L \)), the partition function (5.122) becomes

\[
Z^v[a] \approx \exp \left( -\frac{m_a^2}{2\beta} \int d^3x a_T^2 \right),
\]  

(5.125)

and the exponent gives a mass to the transverse part \( a_T \) of the gauge field of superflow. Recalling the gradient term \( (1/\beta)(\nabla \times a)^2 \) of the \( a \)-field in (5.120) we see the mass has the value \( m_a \).

Having obtained this result we go once more back to the expression (5.115) and realize that the same mass can also be obtained from \( Z^v[a] \) by simply ignoring the \( \delta \)-function nature of \( j^v(x) = 2\pi \delta(x, L) \) and integrating \( j^v(x) \) out using the Gaussian formula (5.65). With such an approximate treatment, the partition function (5.115) yields for the vortex density the simple correlation function

\[
\langle j_i^v(x) j_j^v(x') \rangle = \frac{1}{\epsilon_c} \left( \delta_{ij} - \frac{\nabla_i \nabla_j}{\nabla^2} \right) \delta^{(3)}(x - x').
\]  

(5.126)
5.1 Superfluid Transition

The reason why this simplification is applicable in the high-temperature phase is easy to understand. On a lattice, the sums over lines $L$ in (5.115) correspond to Gaussian sums of the type $\sum_{n_i=-\infty}^{\infty} e^{-\beta \epsilon_c n_i^4/2}$ at each $x$, $i$. At high temperatures where $\beta$ is small, the sum over $n_i$ can obviously be replaced by $1/\sqrt{\beta}$ times an integral over the quasi-continuous variable $n_i \equiv \sqrt{\beta n_i}$. In general, if lines or surfaces of volumes are prolific, the statistical mechanics of fields proportional to the corresponding $\delta$-functions $\delta(x; L)$, $\delta(x; S)$, $\delta(x; V)$ can be treated as if they were ordinary fields. The sums of the geometric configurations turn into functional integrals.

The same mass generation can, of course, be observed in the complex disorder field theory (5.119). At high temperature, the mass term $m_\psi^2$ of the $\psi$-field is negative and the disorder field acquires a nonzero expectation value $\psi_0 = \sqrt{-m_\psi^2/g}$. This produces again the mass term (5.122) with $m_a^2 = \psi_0^2$.

Let us now look at the low-temperature phase. There the $\delta$-function nature of the density $j^v(x) = 2\pi \Phi^v(x; L)$ cannot be ignored in the partition function (5.115). At low temperatures, vortex lines appear only as small loops. An infinitesimal loop gives a simple curl contribution [22]

$$Z^v[a] \sim \exp \left[ -\frac{1}{2\beta} \int d^3x (\nabla \times a)^2 \right],$$

(5.127)

whereas larger loops contribute

$$Z^v[a] \sim \exp \left[ -\frac{1}{2} \int d^3x (\nabla \times a) f(-i\nabla)(\nabla \times a) \right],$$

(5.128)

where $f(k)$ is some smooth function of $k$ starting out with a constant, the so-called stiffness of the $a$-field. Hence the contributions of small vortex loops change only the dispersion of the gauge fields of superflow. Infinitely long vortex lines in $\theta^v$ are necessary to produce a mass term. These appear when the temperature is raised above the critical point, in particular at high temperatures, where the correlation function of the vortex densities is approximately given by (5.126), and (5.115) leads directly to the mass term in (5.125).

With the help of the disorder partition function $Z^v[a]$, the partition function (5.28) can be replaced by the completely equivalent dual partition function

$$\tilde{Z}^v_{\psi} = \int_{-\infty}^{\infty} D\Phi T[a] \sum_{\{\theta\}} \Phi[\theta^v] \int_{-\infty}^{\infty} D\tilde{\theta} e^{-\beta \tilde{H}^v_{\psi}}$$

(5.129)

with the exponent

$$\beta \tilde{H}^v_{\psi} = \frac{1}{2\beta} \int d^3x \left[ (\nabla \times a)^2 + m_a^2 \left( \nabla \tilde{\theta} - \tilde{\theta}^v - a \right)^2 \right].$$

(5.130)

This energy is invariant under the following two gauge transformations. First, there is invariance under the gauge transformations of superflow (5.107), if it is accompanied by a compensating transformation of the angular field $\tilde{\theta}$:

$$a(x) \rightarrow a(x) + \nabla \Lambda(x), \quad \tilde{\theta}(x) \rightarrow \tilde{\theta}(x) + 2\pi \Lambda(x).$$

(5.131)
Second, there is gauge invariance under the joint vortex gauge transformations Eq. (5.23) with (5.24) and the phase transformation of the disorder field,
\[ \tilde{\theta}(x) \rightarrow \tilde{\theta}(x) + \nabla \tilde{\Lambda}_x(x), \]
with gauge functions
\[ \tilde{\Lambda}_x(x) = 2\pi \delta(x; \tilde{V}). \]
At high temperatures, the vortex lines in \( \tilde{\theta} \) are frozen out and the energy (5.130) shows again the mass term (5.122).

The mass term implies that at high temperatures, the gauge field of superflow possesses a finite range. At some critical temperature superfluidity has been destroyed. This is the disorder analog of the famous Meissner effect in superconductors [23], to be discussed in Section 5.2.1. Without the gauge field of superflow, the field \( \tilde{\theta} \) would be of long range, i.e., massless. The gauge field of superflow absorbs this massless mode and the system has only short-range excitations. More precisely, it can be shown that all correlation functions involving local gauge-invariant observable quantities must be of short range in the high-temperature phase.

Take, for instance, the local gauge-invariant current operator of the disorder field
\[ j_s \equiv \nabla \tilde{\theta} - a. \]
Choosing \( \tilde{\theta} \) to absorb the longitudinal part of \( a \), only the transverse part of \( a \) remains in (5.134), which becomes \( j_s^\ast = -a_T \). [24]. From the Hamiltonian (5.130) we immediately find the free correlation function of superflow:
\[ \langle j^\ast_{i}(x_1)j^\ast_{j}(x_2) \rangle \propto \int \frac{d^3k}{(2\pi)^3} \frac{\delta_{ij} - k_{ik}k_{jk}/m_a^2}{k^2 + m_a^2} e^{i k(x_1 - x_2)}, \]
which has no zero-mass pole.

### 5.2 Phase Transition in Superconductor

The specific heat of a superconductor looks quite different from that of helium as shown in Fig. 5.8 (compare Fig. 5.1 on p. 130). It starts out with a behavior typical for an activation process, which is governed by a Boltzmann factor \( c_s \propto e^{-\Delta(0)/k_B T} \), where \( k_B \) is the Boltzmann constant. The activation energy \( \Delta(0) \) shows the energy gap in the electron spectrum at \( T = 0 \). It is equal to the binding energy of the Cooper pairs formed from electrons of opposite momentum near the Fermi sphere. At the critical temperature \( T_c \), the specific heat drops down to the specific heat of a free electron gas
\[ c_n = \frac{2}{3} n_e^2 \mathcal{N}(0) T, \]
where [25]
\[ \mathcal{N}(0) = \frac{3n_e}{4\epsilon_F} = \frac{3mn_e}{2p_F^2} = \frac{3n_e}{2mv_F^2}. \]
5.2 Phase Transition in Superconductor

Figure 5.8 Specific heat of superconducting aluminum [N.E. Phillips, Phys. Rev. 114, 676 (1959)]. For very small $T$, it shows the typical power behavior $e^{-\Delta(0)/k_B T}$ instead of the power behavior in superfluid helium. At the critical temperature $T_c \approx \Delta(0)/1.76$, there is a jump down to the linear behavior characteristic for a free electron gas. The ratio $\Delta c/c_n = 1.43$ agrees well with the BCS theory [26]. A normal metal shows only the linear behavior labeled by $c_s$.

is the density of electrons of mass $m$ at the surface of the Fermi sphere of energy $\epsilon_F$ and momentum $p_F$, velocity $v_F = p_F/m$, and $n_e$ is the density of electrons of both spin directions. The Fermi velocity $v_F$ is typically of the order $10^8$ (cm/sec)(~c/300).

According to the theory of Bardeen, Cooper, and Schrieffer (BCS) [26], the jump is given by the universal law [to be derived in Eq. (7A.24)]

$$\frac{c_s - c_n}{c_n} = \frac{\Delta c}{c_n} = 3 \frac{8}{\pi \zeta(3)} \approx 1.4261 \ldots, \quad (5.138)$$

where $\zeta(3)$ is Riemann’s zeta function $\zeta(z) \equiv \sum_{n=1}^{\infty} n^{-z}$, with $\zeta(3) = 1.202057 \ldots$. This jump agrees perfectly with the experiment in Fig. 5.8.

In the BCS-theory there exists also a universal ratio between the gap $\Delta(0)$ and $T_c$:

$$\frac{\Delta(0)}{T_c} = \pi e^{-\gamma} \approx 1.76, \quad (5.139)$$

where $\gamma \approx 0.577 \ldots$ is the Euler-Mascheroni constant. This ratio is also observed in Fig. 5.8.
5.2.1 Ginzburg-Landau Theory

The BCS theory can be used to derive the Ginzburg-Landau Hamiltonian for the superconducting phase transition \[2, 28\]

\[
H_{HL}[\psi, \psi^*, A] = \frac{1}{2} \int d^3 x \left\{ \left| (\nabla - i q A) \psi \right|^2 + \tau |\psi|^2 + \frac{g}{2} |\psi|^4 + (\nabla \times A)^2 \right\} \tag{5.140}
\]

governing the neighborhood of the critical point. The parameter \(q\) is the charge of the \(\psi\)-field, and \(\tau\) may be identified with \(\tilde{T}/T_{c}^{MF} - 1\), the relative temperature distance from the critical point. It is positive in the normal state and negative in the superconducting state. The field \(\psi(x)\) is a so-called collective field describing the Cooper pairs of electrons of opposite momenta slightly above and below the Fermi sphere \[29\]. The Cooper pairs carry a charge twice the electron charge, \(q = 2e\), and are coupled in (5.140) minimally to the vector potential \(A(x)\). For simplicity, we have set the light velocity \(c\) equal to unity, \(c = 1\). The size of \(\psi\) is equal to the energy gap in the electron spectrum, which is caused by the binding of electrons to Cooper pairs.

Ginzburg and Landau \[27\] found the Hamiltonian (5.140) by a formal expansion of the energy in powers of the energy gap which they considered as an order parameter. They convinced themselves that for small \(\tau\) only the terms up to \(\psi^4\) would be important. To this truncated expansion they added a gradient term to allow for spatial variations of the order parameter, making it an order field denoted by \(\psi(x)\). There exists an elegant derivation of the Ginzburg-Landau Hamiltonian (5.140) from the BCS theory via functional integration. For details see Ref. \[29\] and a short summary in Appendix 7A.

In the critical regime, the Ginzburg-Landau provides us with a simple explanation of many features of superconductors. In most applications, one may neglect fluctuations of the Ginzburg-Landau order field \(\phi(x)\), which is why one speaks of mean-field results, an why one attaches the superscript to the critical temperature \(\tilde{T}_{c}^{MF}\) in the above definition of \(\tau\). Close to the transition, the properties of a superconductor are well described by the Ginzburg-Landau Hamiltonian [compare (5.119)].

The Ginzburg-Landau Hamiltonian (5.140) possesses a conserved supercurrent which is found by applying Noether’s rule (3.117) to (5.140). The current density is

\[
j(x, t) = \frac{1}{2i} \left( \psi^\dagger(x, t) [\nabla - iqA(x, t)] \psi(x, t) - \{ [\nabla - iqA(x, t)] \psi(x, t) \}^\dagger \psi \right) = \frac{i}{2} \psi^*(x, t) \overrightarrow{\nabla} \psi(x, t) - qA \psi^*(x, t) \psi(x, t). \tag{5.141}\]

This differs from the Schrödinger current density (5.141) by the use of natural units \(m = 1, c = 1\).

Let us now proceed as in (5.10) and (5.121) and decompose the field \(\psi\) as

\[
\psi(x) = \tilde{\rho}(x) e^{i\tilde{\theta}(x)}. \tag{5.142}\]
Inserting this into \((5.140)\) and remembering \((5.17)\), we find

\[
H_{\text{GL}}[\tilde{\rho}, \tilde{\theta}, \tilde{\theta}^v, A] = \int d^3x \left[ \frac{\tilde{\rho}^2}{2} (\nabla \tilde{\theta} - \tilde{\theta}^v - qA)^2 + \frac{1}{2} (\nabla \tilde{\rho})^2 + V(\tilde{\rho}) + \frac{1}{2} (\nabla \times A)^2 \right],
\]

where \(V(\tilde{\rho})\) is the potential of the \(\tilde{\rho}\)-field:

\[
V(\tilde{\rho}) = \frac{\tau}{2} \tilde{\rho}^2 + \frac{g}{4} \tilde{\rho}^4. \tag{5.144}
\]

In the low-temperature phase we go to the hydrodynamic limit by setting \(\tilde{\rho}(x)\) equal to its value \(\tilde{\rho}_0 = \sqrt{-\tau/g}\) at the minimum of the energy \((5.140)\). The resulting hydrodynamic or London energy of the superconductor is

\[
H_{\text{hy}}^{\text{SC}}[\tilde{\theta}, \tilde{\theta}^v, A] = \int d^3x \left[ m_A^2 (\nabla \tilde{\theta} - \tilde{\theta}^v - qA)^2 + \frac{1}{2} (\nabla \times A)^2 \right], \tag{5.145}
\]

where we have introduced the density of superfluid particles

\[
n_0 = \tilde{\rho}_0^2. \tag{5.146}
\]

At very low temperatures where vortices are absent, the first term in \((5.145)\) gives a mass

\[
m_A = \sqrt{n_0 q} \tag{5.147}
\]

to the transverse part of the vector field. This causes a finite penetration depth \(\lambda = 1/m_A\) of the magnetic field in a superconductor, thus explaining the famous Meissner effect of superconductivity.

This mechanism is imitated in the standard model of electromagnetic and weak interactions to give the vector mesons \(W^+, W^0, Z\) a finite mass, thereby explaining the strong suppression of weak with respect to electromagnetic interactions. There the Meissner effect is called Higgs effect.

In the same limit, the current density of superfluid particles becomes

\[
\mathbf{j}_s = n_0 (\nabla \tilde{\theta} - \tilde{\theta}^v - qA). \tag{5.148}
\]

The partition function reads

\[
Z_{\text{hy}}^{\text{SC}} = \int D\mathbf{A} \Phi_T[A] \sum_{\{s\}} \Phi[\tilde{\theta}^v] \int_{-\infty}^{\infty} D\tilde{\theta} e^{-\beta H_{\text{hy}}^{\text{SC}}[\tilde{\theta}, \tilde{\theta}^v, A]}. \tag{5.149}
\]

To distinguish this discussion from the previous one of superfluid helium we call the temperature of the superconductor \(T\) and its inverse \(\beta\).

The energy \((5.145)\) has the same form as the energy in the disorder representation \((5.130)\) of superfluid \(^4\text{He}\). The role of the gauge field of superflow is now played by
the vector potential $A$ of magnetism. The energy has the following two types of
gauge symmetries: the magnetic invariance
\[ A(x) \rightarrow A(x) + q^{-1} \nabla \Lambda(x), \quad \tilde{\theta}(x) \rightarrow \tilde{\theta}(x) + \Lambda(x), \quad (5.150) \]
and the vortex gauge invariance
\[ \tilde{\theta}'(x) \rightarrow \tilde{\theta}'(x) + \partial_i \tilde{\Lambda}_\delta(x), \quad \tilde{\theta}(x) \rightarrow \tilde{\theta}(x) + \tilde{\Lambda}(x), \quad (5.151) \]
with gauge functions
\[ \tilde{\Lambda}_\delta(x) = 2\pi \delta(x; \tilde{V}). \quad (5.152) \]

As in the description of superfluid $^4$He with the partition function (5.28), the
partition function (5.149) gives us the statistical behavior of the superconductor not
only at zero temperature, where the energy (5.145) was constructed, but at all not
too large temperatures. The fluctuating vortex gauge field $\tilde{\theta}'$ ensures the validity
through the phase transition, path, fluctuations.

### 5.2.2 Disorder Theory of Superconductor

We shall now derive the disorder representation of this partition function in which
the vortex lines of the superconductor play a central role in describing the phase transition [23].

At low temperatures, the vortices are frozen, and the $\tilde{\theta}$-fluctuations in the partition function can be integrated out. This reduces the energy to
\[ H^{by} \sim \frac{m_A^2}{2} \int d^3x A^2, \quad (5.153) \]
i.e., to a simple transverse mass term for the vector potential $A$. This is the famous
Meissner effect in a superconductor, which limits the range of a magnetic field to a
finite penetration depth $\lambda = 1/m_A$. The effect is completely analogous to the one
observed previously in the disorder description of the superfluid where the superfluid
acquired a finite range in the normal phase.

To derive the disorder theory of the partition function (5.149), we supplement
the energy by a core energy of the vortex lines
\[ H_c = \frac{\tilde{\epsilon}_c}{2} \int d^3x (\nabla \times \tilde{\theta}')^2. \quad (5.154) \]
As in the partition function (5.78), an auxiliary $\tilde{b}_i$ field can be introduced to bring
the exponent in (5.145) to the canonical form
\[ \tilde{\beta}H_{SC}^{by} = \int d^3x \left[ \frac{1}{2\beta m_A^2} \tilde{b}^2 + i\tilde{b} (\nabla \tilde{\theta} - \tilde{\theta}' - qA) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 + \frac{\beta \tilde{\epsilon}_c}{2} (\nabla \times \tilde{\theta}')^2 \right]. \quad (5.155) \]
By integrating out the $\tilde{\theta}$-fields in the associated partition function, we obtain the
conservation law
\[ \nabla \cdot \tilde{b} = 0, \quad (5.156) \]
5.2 Phase Transition in Superconductor

which is fulfilled by expressing \( \tilde{b} \) as a curl of the gauge field \( \tilde{a} \) of superflow in the superconductor

\[
\tilde{b} = \nabla \times \tilde{a}.
\]

This brings the energy to the form

\[
\tilde{\beta} H_{SC}^{hy} = \int d^3 x \left[ \frac{1}{2\beta m_A^2} (\nabla \times \tilde{a})^2 - i q \tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 - i \tilde{a} \cdot \tilde{j}^v + \frac{\tilde{\beta} c}{2} \tilde{j}^v \right],
\]

where

\[
\tilde{j}^v = \nabla \times \tilde{\theta}^v
\]

is the vortex density in the superconductor. At low temperatures where \( \tilde{\beta} \) is large and the vortex lines are frozen out, the two last terms in the Hamiltonian can be neglected. Integrating out the \( \tilde{a} \)-field we re-obtain the transverse mass term (5.153) of the Meissner effect. At high temperatures, on the other hand, the vortex lines are prolific and the vortex density \( \tilde{j}^v \) can be integrated out in the associated partition function like an ordinary field using the analog to the correlation function (5.126).

This produces the transverse mass term

\[
\frac{1}{2\beta m_A^2} \int d^3 x \ m^2_A \tilde{a}_T^2
\]

where

\[
m^2_A = q^2 m_A^2 / \epsilon_c.
\]

This can immediately be seen to destroy the Meissner effect in the superconductor at high temperature. Indeed, inserting the curl (5.157) into the energy (5.155), and using the result (5.160), we obtain at high \( \tilde{T} \):

\[
\tilde{\beta} H_{SC}^{hy} = \int d^3 x \left[ \frac{1}{2\beta m_A^2} [(\nabla \times \tilde{a})^2 + m^2_A \tilde{a}_T^2] - i \tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 \right].
\]

If we integrate out the massive \( \tilde{a} \)-field, the Hamiltonian of the vector potential becomes

\[
H_A = \frac{1}{2} \int d^3 x \ \nabla \times A \left( 1 + \frac{m^2_A}{-\nabla^2 + m^2_A} \right) \nabla \times A.
\]

Expanding the denominator in powers of \( -\nabla^2 \) we see that only gradient energies appear, but no mass term. Thus, the vector potential \( A \) maintains its long range and yields Coulomb-like forces at large distances. Only the dispersion is modified to a more complicated \( k \)-dependence of the energy.

In the low-temperature phase, on the other hand, the mass \( m_A \) is zero, and the \( m^2_A \)-term in (5.163) produces again the transverse mass Hamiltonian (5.153) which is responsible for the Meissner effect.

We can represent the fluctuating vortices in the superconductor by a disorder field theory in the same way as we did for the vortices in the superfluid, by repeating
the transformations from (5.115) to (5.118). The angular field variable of disorder will now be denoted by $d \theta$, the vortex lines in the disorder theory by $\theta^\nu$. The disorder action reads

$$\tilde{\beta} \tilde{H}_{\text{hy}}^{\text{dis}} = \int d^3 x \left[ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 - i q \tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 + \frac{m_a^2}{2 \beta m_A^2} (\nabla \theta - \theta^\nu - \tilde{a})^2 \right]. \quad (5.164)$$

Near the phase transition, this is equivalent to a disorder field energy

$$\tilde{\beta} \tilde{H}_{\text{hy}}^{\text{dis}} \sim \int d^2 x \left[ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 - i q \tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 + \frac{1}{2} (\nabla - i \tilde{a}) \phi^2 + \frac{\tau}{2} |\phi|^2 + \frac{g}{4} |\phi|^4 \right]. \quad (5.165)$$

The vector potential $A$ fluctuates harmonically in such a way that the associated magnetic field is on the average equal to $\tilde{a}/\tilde{\beta}$. Integrating $A$ out we obtain from (5.165)

$$\tilde{\beta} \tilde{H}_{\text{hy}}^{\text{dis}} \sim \int d^2 x \left[ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 + m_a^2 \tilde{a}_T^2 + \frac{1}{2} (\nabla - i \tilde{a}) \phi^2 + \frac{\tau}{2} |\phi|^2 + \frac{g}{4} |\phi|^4 \right]. \quad (5.166)$$

This Hamiltonian is invariant under the gauge transformations

$$\phi(x) \to e^{i \tilde{\Lambda}(x)} \phi(x), \quad \tilde{a}(x) \to \tilde{a}(x) + \nabla \tilde{\Lambda}(x). \quad (5.167)$$

The partition function is

$$Z_{\text{SC}}^{\text{dual}} = \int D\phi \int D\phi^* \int D\tilde{a} \Phi[\tilde{a}] e^{-\tilde{\beta} \tilde{H}_{\text{hy}}^{\text{dis}}}, \quad (5.168)$$

where $\Phi[\tilde{a}]$ is some gauge-fixing functional.

This partition function can be evaluated perturbatively as a power series in $g$. The terms of order $g^n$ consist of Feynman integrals which can be pictured by Feynman diagrams with $n + 1$ loops [30]. These loops are pictures of the topology of vortex loops in the superconductor.

The disorder field theory for the superconductor was for a long time the only formulation which has led to a determination of the critical and tricritical properties of the superconducting phase transition [23, 31]. Within the Ginzburg-Landau theory, an explanation was found only recently [32].

In the hydrodynamic Hamiltonian (5.164), the elimination of $A(x)$ leads to the Hamiltonian

$$\tilde{\beta} \tilde{H}_{\text{hy}}^{\text{dis}} = \int d^3 x \left[ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 + m_a^2 \tilde{a}_T^2 \right] + \frac{m_a^2}{2 \beta m_A^2} (\nabla \theta - \theta^\nu - \tilde{a})^2, \quad (5.169)$$

which is gauge-invariant under

$$\theta(x) \to \theta(x) + \tilde{\Lambda}(x), \quad \tilde{a}(x) \to \tilde{a}(x) + \nabla \tilde{\Lambda}(x). \quad (5.170)$$
5.3 Order versus Disorder Parameter

Since Landau’s 1947 work [3], phase transitions are characterized by an order parameter which is nonzero in the low-temperature, ordered phase, and zero in the high-temperature, disordered phase. In the 1980s, this characterization has been enriched by the dual disorder field theory of various phase transitions [7]. The expectation value of the disorder field is the disorder parameter which has the opposite temperature behavior, being nonzero in the high-temperature and zero in the low-temperature phase. Let us identify the order and disorder fields in superfluids and superconductors, and study their expectation values in the hydrodynamic theories of the two systems.

5.3.1 Superfluid $^4$He

In Landau’s original description of the superfluid phase transition with the Hamiltonian (5.6), the role of the order parameter $\mathcal{O}$ is played by the expectation value of the complex order field $\mathcal{O}(x) = \phi(x)$:

$$\mathcal{O} \equiv \langle \mathcal{O}(x) \rangle = \langle \phi(x) \rangle.$$  \hspace{1cm} (5.171)

Its behavior can be extracted from the large-distance limit of the correlation function of two order fields $\mathcal{O}(x)$

$$G_{\mathcal{O}}(x_2, x_1) \equiv \langle \mathcal{O}(x_2) \mathcal{O}^*(x_1) \rangle = \langle \phi(x_2) \phi^*(x_1) \rangle.$$  \hspace{1cm} (5.172)

This is done by taking advantage of the cluster property of the correlation functions of arbitrary local operators

$$\langle O_1(x_2) O_2(x_1) \rangle \big|_{|x_2-x_1| \to \infty} \to \langle O_1(x_2) \rangle \langle O_2(x_1) \rangle.$$  \hspace{1cm} (5.173)

Hence we obtain the large-distance limit of the correlation function (5.172)

$$G_{\mathcal{O}}(x_2, x_1) \big|_{|x_2-x_1| \to \infty} \to |\mathcal{O}|^2.$$  \hspace{1cm} (5.174)

If we go to the hydrodynamic limit of the theory where the size of $\phi(x)$ is frozen and the order field reduces to $O(x) = e^{i\theta(x)}$, the order parameter becomes

$$\mathcal{O} \equiv \langle O(x) \rangle = \langle e^{i\theta(x)} \rangle.$$  \hspace{1cm} (5.175)

This is extracted from the large-distance limit of the correlation function

$$G_{\mathcal{O}}(x_2, x_1) = \langle e^{i\theta(x_2)} e^{-i\theta(x_1)} \rangle.$$  \hspace{1cm} (5.176)

If we want to use (5.175) as an order parameter to replace (5.171), it is important that the correlation function (5.176) is vortex-gauge-invariant under the transformations (5.23). This is not immediately obvious. A quantity where the invariance is obvious is the expectation value

$$G_{\mathcal{O}}(x_2, x_1) = \langle e^{i \int_{x_1}^{x_2} d\mathbf{x} [\mathbf{\nabla} \theta(x) - \theta^v(x)]} \rangle.$$  \hspace{1cm} (5.177)
The transformations (5.23) do not change the exponent. We have, however, achieved vortex gauge invariance at the price of an apparent dependence of (5.177) on the shape of the path from $x_1$ to $x_2$. Fortunately it is possible to show that this shape dependence is not really there, so that the vortex gauge-invariant correlation function is uniquely defined, and that it is in fact the same as (5.176), thus ensuring the vortex gauge invariance of (5.176).

For the proof let us rewrite (5.177) in the form

$$G_C(x_2, x_1) = \left\langle e^{i \int d^3x \, b^m(x) [\nabla \theta(x) - \theta^v(x)]} \right\rangle,$$  \hspace{1cm} (5.178)

where the field

$$b^m(x) = \delta(x; \vec{L}^{x_2}_{x_1})$$  \hspace{1cm} (5.179)

is a $\delta$-function on an arbitrary line $\vec{L}^{x_2}_{x_1}$ running from $x_1$ to $x_2$. This field satisfies [recall (4.10) and (4.11)]

$$\nabla \cdot b^m(x) = q(x),$$  \hspace{1cm} (5.180)

where

$$q(x) = \delta^{(3)}(x - x_1) - \delta^{(3)}(x - x_2).$$  \hspace{1cm} (5.181)

It is now easy to see that the expression (5.178) is invariant under deformations of $\vec{L}^{x_2}_{x_1}$. Indeed, let $\vec{L}'^{x_2}_{x_1}$ be a different path running from $x_1$ to $x_2$. Then the difference between the two is a closed path $\vec{L}$, and the exponents in (5.178) differ by an integral

$$i \int d^3x \, \delta(x; \vec{L}) \left[ \nabla \theta(x) - \theta^v(x) \right].$$  \hspace{1cm} (5.182)

The first term vanishes after a partial integration due to Eq. (4.9). The second term becomes, after inserting (5.21),

$$-2\pi i \int d^3x \, \delta(x; \vec{L}) \delta(x; S) = -2\pi ik, \quad k = \text{integer}.$$  \hspace{1cm} (5.183)

The integer $k$ counts how many times the line $\vec{L}$ pierces the surface $S$. Since $-2\pi ik$ appears in the exponential it does not contribute to the correlation function (5.178). Thus we have proved that the expectation value (5.177) is independent of the path along which the integral runs from $x_1$ to $x_2$.

We recognize the analogy to the discussion of magnetic monopoles in Section 4.4. For this reason we shall speak of $q(x)$ as a charge density of a monopole-antimonopole pair located at $x_2$ and $x_1$, respectively. In the description of monopoles in Section 4.4, a monopole at $x_2$ is accompanied by a Dirac string $L^{x_2}$ along which the flux is imported from infinity, the antimonopole at $x_1$ is accompanied by a Dirac string $L^{x_1}$ along which the flux is exported to infinity. Since the shape of the two lines is irrelevant, we may distort them into a single line $L^{x_2}_{x_1}$ connecting $x_1$ with $x_2$ along an arbitrary path, which in the present context becomes the line $\vec{L}^{x_2}_{x_1}$ in (5.179).
The field $b^m(x)$ is a gauge field with the same properties as the monopole gauge field in Eq. [33]. A change of the shape of the line $\tilde{L}_{x_1}^{x_2}$ is achieved by a monopole gauge transformation [recall (4.63)]

$$b^m(x) \rightarrow b^m(x) + \nabla \times \delta(x; \tilde{S}). \quad (5.184)$$

Note that the invariant field strength of this gauge field is the divergence (5.180) rather than a curl [compare (5.25) for a vortex gauge field].

In this way, the independence of the manifestly vortex gauge-invariant correlation function (5.178) on the shape of the line connecting $x_1$ with $x_2$ is expressed as an additional invariance under monopole gauge transformations. The correlation function is thus a double-gauge-invariant object.

After this discussion we are able to define a manifestly vortex gauge-invariant formulation of the order parameter (5.175). It is given by the expectation value

$$O = O(x) = \langle e^{i \int x' d x' [\nabla \theta(x') - \nabla \vartheta(x')]} \rangle = \langle e^{i \int d^3 x' \delta(x'; L^x) [\nabla \theta(x') - \nabla \vartheta(x')]} \rangle, \quad (5.185)$$

where $\delta(x; L^x)$ is the $\delta$-function on an arbitrary line as defined in Eq. (4.57). It comes from infinity along an arbitrary path ending at $x$.

Let us now study the large-distance behavior (5.173) of the correlation function (5.177) at low and high temperatures. At low temperature where vortices are rare, the $\theta(x)$-field fluctuates almost harmonically. By Wick's theorem, according to which harmonically fluctuating variables $\theta_1, \theta_2$ satisfy the equation [34]

$$\langle e^{i \theta_1} e^{i \theta_2} \rangle = e^{-\frac{1}{2} \langle \theta_1 \theta_2 \rangle}, \quad (5.186)$$

we can approximate

$$G_{O}(x_2, x_1) \approx e^{-\frac{1}{2} \langle \theta(x_2) - \theta(x_1) \rangle^2} = e^{-\frac{1}{2} \langle \theta(x_2) - \theta(x_1) \rangle^2} = e^{-\frac{1}{2} \langle \theta(x_2) - \theta(x_1) \rangle^2}. \quad (5.187)$$

The correlation function of two $\theta(x)$-fields is

$$\langle \theta(x_2) \theta(x_1) \rangle \approx T v_0(|x_2 - x_1|), \quad (5.188)$$

where $v_0(r)$ is the Coulomb potential (5.43) which goes to zero for $r \rightarrow \infty$. The correlation function (5.187) is then equal to

$$G_{O}(x_2, x_1) \approx e^{-T v_0(0)} e^{-T v_0(|x_2-x_1|)}. \quad (5.189)$$

This is finite only after remembering that we are studying the superfluid in the hydrodynamic limit which is correct only for length scales larger than the coherence length $\xi$. In He this is of the order of a few $\text{rA}$. Thus we should perform all wave vector integrals only for $|k| \leq \Lambda \equiv 1/\xi$, which makes $v_0(0)$ as finite quantity

$$v_0(0) = 1/2\xi^2. \quad (5.190)$$
As a result, the correlation function (5.187) has a nonzero large-distance limit
\[ G_\mathcal{O}(x_2, x_1) \xrightarrow{|x_2 - x_1| \to \infty} \text{const}, \] (5.191)
implying via Eq. (5.173) that the order parameter \( \mathcal{O} = \langle e^{i\theta(x)} \rangle \) is nonzero.

Let us now calculate the large-distance behavior in the high-temperature phase. To find the correlation function \( G_\mathcal{O}(x_2, x_1) \), we insert into the partition function (5.28) the extra source term
\[ e^{i\theta(x_2)}e^{-i\theta(x_1)} = e^{-i\int d^3x q(x)\theta(x)}. \] (5.192)

This term enters the canonical representation (5.79) of the energy as follows:
\[ \beta H = \int d^3x \left[ \frac{1}{2\beta} b^2 - i b (\nabla \theta - \theta^v) + \frac{\beta\epsilon_c}{2}(\nabla \times \theta^v)^2 + iq(x)\theta(x) \right], \] (5.193)
where we have allowed for an extra core energy, for the sake of generality. Integrating out the \( \theta \)-field in the partition function gives the constraint
\[ \nabla \cdot b(x) = -q(x). \] (5.194)
The constraint is solved by the negative of the monopole gauge field (5.179), and has the general solution
\[ b(x) = \nabla \times a(x) - b^m(x), \] (5.195)
so that the energy (5.193) can be replaced by [using once more (5.183)]
\[ \beta H = \int d^3x \left[ \frac{1}{2\beta} (\nabla \times a - b^m)^2 - i a \cdot j^v + \frac{\beta\epsilon_c}{2} j^v \right]. \] (5.196)
Under a monopole gauge transformation (5.184), this remains invariant if the gauge field of superflow is simultaneously transformed as
\[ a(x) \rightarrow a(x) + \delta(x; \tilde{S}). \] (5.197)

The correlation function (5.178) is now calculated from the functional integral over the Boltzmann factor with the Hamiltonian (5.196).

In the transformed energy (5.196), the presence of the source term (5.192) in the functional integral is accounted for by the \( b^m \)-dependent integrand
\[ e^{-i\Sigma x q(x)\theta(x)} \equiv e^{-\frac{1}{\beta} \int d^3x \left\{ \frac{1}{2} b^m(x)^2 - b^m(x)\nabla \times a(x) \right\}}. \] (5.198)

It is instructive to calculate the large-distance behavior (5.191) of the correlation function in the low-temperature phase once more in this canonical formulation. At low temperatures, the vortex lines are frozen out and we can omit the last two terms.
in (5.196). We integrate out the gauge field $\mathbf{a}$ of superflow in the associated partition function and find that the partition function contains $b^m$ in the form of a factor

$$e^{-\frac{1}{2\pi} \int d^3x \left\{ b^m(x)^2 - \frac{1}{m^2} \nabla^2 b^m(x) \right\} m^\pm}$$

From this we obtain the correlation function

$$G_O(x_1, x_2) = e^{-\frac{1}{2\pi} \int d^3x q(x) \cdot \nabla q(x)} = e^{-\frac{1}{2\pi} \int d^3x d^3x' q(x) n_0(x-x') q(x').}$$

Inserting (5.181), this becomes

$$G_O(x_1, x_2) = e^{-v_0(0)/\beta} e^{v_0(x_1-x_2)/\beta}$$

in agreement with the previous result (5.189).

The canonical formulation (5.196) of the energy enables us to calculate the large-distance behavior of the correlation function in the high-temperature phase. The prolific vortex fluctuations produce a transverse mass term $m^2_a a^2$ which changes (5.199) to (see also Ref. [35])

$$e^{-\frac{1}{2\pi} \int d^3x \left\{ b^m(x)^2 - \frac{1}{m^2} \nabla^2 b^m(x) \right\}} = e^{-\frac{1}{2\pi} \int d^3x \left[ \nabla b^m(x) \cdot \frac{1}{m^2 + m^2_a} \nabla b^m(x) \right].}$$

Using (5.180), we factorize this as

$$e^{-\frac{1}{2\pi} \int d^3x q(x) \cdot \nabla q(x)} \times e^{-\frac{1}{2\pi} \int d^3x b^m(x) \frac{m^2}{m^2 + m^2_a} b^m(x).}$$

The first exponent contains the Yukawa potential

$$v_{m_a}(r) \equiv \int \frac{d^3k}{(2\pi)^3} e^{ikx} \frac{1}{k^2 + m^2_a} = \frac{e^{-m_a r}}{4\pi r}$$

between the monopole-antimonopole pair at $x_2$ and $x_1$, respectively, in the same form as in (5.201), $e^{-v_{m_a}(0)/\beta} e^{v_{m_a}(|x_1-x_2|)/\beta}$, and goes to zero for large distances, i.e., the exponential tends towards a constant. The second factor, on the other hand, has the form [recall (5.179)]

$$e^{-\frac{1}{2\pi} \int d^3x d^3x' \bar{b}(x, \tilde{L}^{\pm}_{x_1}) v_{m_a}(|x-x'|) \bar{b}(x', \tilde{L}^{\pm}_{x_2})}.$$
Due to the cluster property (5.173) of correlation functions, this shows that at high temperatures, the expectation value $O = \langle O(x) \rangle = \langle e^{i\theta(x)} \rangle$ vanishes, so that $O$ is indeed a good order parameter.

The mechanism which gives an energy to the initially irrelevant line $\tilde{L}_{x_1}^{x_2}$ connecting monopole and antimonopole is completely analogous to the generation of surface energy in the previous Eq. (5.75). There the energy arose from a mass of the $\theta$-fluctuations, here from a mass of the $a$-field fluctuations which was caused by the proliferation of infinitely long vortex lines in the high-temperature phase.

Note that an exponential falloff is also found within Landau’s complex order field theory where

$$\langle \psi(x_1)\psi(x_2) \rangle \propto \int \frac{d^3k}{(2\pi)^3} e^{ik(x_1-x_2)} \frac{1}{k^2 + m^2} = \frac{1}{4\pi} e^{-m|x_1-x_2|},$$

(5.207)

However, here the finite range arises in a different way. In calculating (5.207), the size fluctuations of the order field play an essential role. In the partition function (5.28), their role is taken over by the fluctuations of the vortex gauge field $\Theta'(x)$, as pointed out at the end of Section 5.1.4. The proliferation of the vortex lines produces the finite range $1/m_a$ of the Yukawa potential and the exponential falloff (5.206).

### 5.3.2 Superconductor

In contrast to the expectation value (5.175) for superfluid helium, the expectation value of the order field $\psi(x)$ of the Ginzburg-Landau Hamiltonian (5.140) cannot be used as an order parameter since it is not invariant under the ordinary magnetic gauge transformations (5.150). The expectation of all non-gauge-invariant quantities vanishes for all temperatures. This intuitively obvious fact is known as Elitzur’s theorem [36]. The theorem applies also to the hydrodynamic limit of $\psi(x)$, so that the expectation value of the exponential $e^{i\tilde{\theta}(x)}$ cannot serve as an order parameter. Let us search for other possible candidates to be extracted from the large-distance limit of various gauge-invariant correlation functions.

**a) Schwinger Candidate for Order Parameter**

As a first possible candidate, consider the following gauge-invariant version of the expectation value of $\langle e^{i\tilde{\theta}(x_2)}e^{-i\tilde{\theta}(x_1)} \rangle$:

$$G_{\tilde{O}}(x_2,x_1) = \langle e^{i\tilde{\theta}(x_2)}e^{-i\int_{x_1}^{x_2} dx A(x)} e^{-i\tilde{\theta}(x_1)} \rangle,$$

(5.208)

which can also be written as

$$G_{\tilde{O}}(x_2,x_1) = \langle e^{i\tilde{\theta}(x_2)}e^{-i\int d^3x b^m(x)A(x)} e^{-i\tilde{\theta}(x_1)} \rangle,$$

(5.209)

where $b^m(x)$ is the $\delta$-function (5.179) along the line $L_{x_1}^{x_2}$ connecting $x_1$ with $x_2$. This expression is obviously invariant under magnetic gauge transformations (5.150), due to Eqs. (5.180) and (5.181).
5.3 Order versus Disorder Parameter

As before, we must make the correlation function (5.209) manifestly invariant under vortex gauge transformations (5.151). This can be done by adding, as in (5.192), a vortex gauge field:

\[ G(\mathbf{x}_2, \mathbf{x}_1) = \left\langle e^{i \int d^3x \, b_m(x) [\nabla \delta(x) - A(x) - \delta^*(x)]} \right\rangle. \]  
(5.210)

The associated order parameter would be [compare (5.185)]

\[ \hat{O} \equiv \langle \hat{O}(\mathbf{x}) \rangle = \left\langle e^{i \int d^3x' \, \delta(x'; \mathcal{L}_x) [\nabla \delta(x') - A(x') - \delta^*(x')] \right\rangle. \]  
(5.211)

We now observe that in contrast to the correlation function in the superfluid (5.176), this is not invariant under deformations of the shape of the line \( \tilde{L}_{x_1}^2 \) connecting the points \( \mathbf{x}_1 \) and \( \mathbf{x}_2 \). Indeed, if we apply the associated monopole gauge transformation (5.154) to (5.209), we see that

\[ e^{-i \int d^3x \, b_m(x) A(x) - \int d^3x \, \{b_m(x) A(x) + [\nabla \times \delta(x; \mathcal{S})] A(x)\} = e^{-i \int d^3x \, \{b_m(x) A(x) + B(x) \delta(x; \mathcal{S})\}} \]  
(5.212)

where \( \mathcal{S} \) is the surface over which \( \tilde{L}_{x_1}^2 \) has swept. Thus, the correlation function (5.209) changes under monopole gauge transformations by a phase

\[ G(\mathbf{x}_2, \mathbf{x}_1) \rightarrow e^{-i \int d^3x \, B(x) \delta(x; \mathcal{S})} G(\mathbf{x}_2, \mathbf{x}_1), \]  
(5.213)

which depends on the fluctuating magnetic flux through the surface \( \mathcal{S} \). For this reason, we must first remove the freedom of choosing the shape of \( \tilde{L}_{x_1}^2 \) which connects \( \mathbf{x}_1 \) with \( \mathbf{x}_2 \). The simplest choice made by Schwinger [37] is the straight path between from \( \mathbf{x}_1 \) to \( \mathbf{x}_2 \).

Still, the correlation function (5.210) does not supply us with an order parameter when taking the limit of large \( |\mathbf{x}_2 - \mathbf{x}_1| \). In order to verify this, we go to the partition function with the canonical representation (5.155) of the Hamiltonian, and insert the expression (5.210). Then we change field variables from \( \tilde{b} \) to \( \tilde{b} - b^m \), and use (5.157) to obtain (5.158) with \( (\nabla \times \tilde{a})^2 \) replaced by \( (\nabla \times \tilde{a} - b^m)^2 \):

\[ \tilde{\beta} H^{by}_{SC} = \int d^3x \left[ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a} - b^m)^2 - i \tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 - i \tilde{a} \cdot \tilde{j} + \frac{\tilde{\beta} \tilde{\epsilon}}{2} \tilde{j}_v^2 \right]. \]  
(5.214)

This is quadratic in the magnetic vector potential \( A \) which can be integrated out in the associated partition function, leading to a Hamiltonian

\[ \tilde{\beta} H^{by}_{SC} = \int d^3x \frac{1}{2 \beta m_A^2} \left[ (\nabla \times \tilde{a} - b^m)^2 + m_A^2 \tilde{a}^2 - i \tilde{a} \cdot \tilde{j} + \frac{\tilde{\beta} \tilde{\epsilon}}{2} \tilde{j}_v^2 \right]. \]  
(5.215)

With this Hamiltonian, the correlation function (5.209) can be calculated from the expectation value [compare (5.198)]:

\[ G(\mathbf{x}_2, \mathbf{x}_1) = \left\langle e^{-\frac{1}{\beta m_A} \int d^3x \{ \frac{i}{2} b^m(x)^2 - b_m(x) [\nabla \times \tilde{a}(x)] \}} \right\rangle. \]  
(5.216)
Consider first the low-temperature phase where the vortices in the superconductor are frozen out, and we may omit the last two terms in (5.214). Then the massive field $\tilde{a}$ can be integrated out trivially in the partition function, leading to

$$G_{\tilde{O}}(x_1, x_2) \sim e^{-\frac{\tilde{\beta}m^2}{2} \int d^3x \left[ b^{m^2} \cdot (\nabla \times b^{m^2}) \right]}.$$  

(5.217)

This is the same expression as in the high-temperature phase of the superfluid in Eq. (5.202), except that the relevant mass is now the Meissner mass $m_A$ of the superconductor rather than $m_a$. The mass $m_A$ makes the line $\tilde{L}_{x_1}^{x_2}$ between $x_1$ and $x_2$ in $b^m(x)$ energetic, and leads to the same type of exponential long-distance falloff of the correlation function as in Eq. (5.206):

$$G_{\tilde{O}}(x_1, x_2) \sim e^{-\text{const} \cdot |x_1 - x_2|} \to 0.$$  

(5.218)

This implies a vanishing of the candidate (5.211) for the order parameter:

$$\tilde{O} = \langle \tilde{O}(x) \rangle = 0.$$  

(5.219)

Thus $\tilde{O}$ fails to indicate the order of the low-temperature phase.

Could $\tilde{O}$ be a disorder parameter? To see this we go to the high-temperature phase where the vortex lines are prolific. In the Hamiltonian (5.214), this corresponds to integrating out $\tilde{j}^x$ like an ordinary Gaussian variable. This produces a Hamiltonian

$$\tilde{\beta}H_{\text{hy}}^{\text{SC}} = \int d^3x \left[ \frac{1}{2\beta m_A^2} \left[ (\nabla \times \tilde{a} - b^m)^2 + m_a^2 \tilde{a}^2 \right] - i\tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}}{2} (\nabla \times A)^2 \right].$$  

(5.220)

If we now integrate out the magnetic vector potential $A$, the mass term changes from $m_a^2$ to $m_a^2 + m_A^2$, causing the correlation function to fall off even faster than in (5.218). Hence $\tilde{O}$ is again zero and does not distinguish the different phases.

b) Dirac Candidate for Order Parameter

An alternative to Schwinger’s choice of a straight line connection from $x_1$ to $x_2$ in Eq. (5.208) we may choose a different monopole gauge field in Eq. (5.210) which possesses the same divergence $\nabla \cdot b^m(x) = q(x)$ as $b^m(x)$ in Eq. (5.210), but has a longitudinal gauge [38, 39, 40]:

$$\nabla \times b^m(x) = 0.$$  

(5.221)

Such a choice exists. We simply take

$$b^m(x) = \nabla \frac{1}{\nabla^2} q(x) = -\frac{1}{4\pi} \nabla \left[ \frac{1}{|x - x_1|} - \frac{1}{|x - x_2|} \right].$$  

(5.222)
The monopole gauge field (5.222) is the associated Coulomb field which is longitudinal. Now the exponent in (5.217) simplifies and we obtain the limit
\[
G_\tilde{O}(\mathbf{x}_1, \mathbf{x}_2) \sim e^{-\frac{\bar{m}^2}{2} \int d^3x \, b^m \cdot \mathbf{A} \int d^3x \, \left( \nabla^2 + \frac{1}{\mathbf{x}_1 - \mathbf{x}_2} \right) \mathbf{a} = e^{-\frac{\bar{m}^2}{8\pi} \int d^3x \, \mathbf{x}_1 - \mathbf{x}_2 \rightarrow \infty} \mathbf{1}. (5.223)
\]
Actually, this result could have been deduced directly from the energy (5.215). In the longitudinal gauge, \(b^m\) is orthogonal to the purely transversal field \(\nabla \times \tilde{a}\), so that it decouples:
\[
(\nabla \times \tilde{a} - b^m)^2 = (\nabla \times \tilde{a})^2 + b^m, (5.224)
\]
and leads directly to (5.223).

The nonzero long-distance limit (5.223) is what we expect in the ordered phase, giving rise to the hope that (5.211), with \(\delta(\mathbf{x}' - L^x)\) replaced by the field (5.222) of a single monopole at \(\mathbf{r}\):
\[
b^m(\mathbf{x}') \equiv -\nabla' \frac{1}{\mathbf{r}'} \delta^{(3)}(\mathbf{x}' - \mathbf{r}), (5.225)
\]
can supply us with an order parameter:
\[
\bar{O} = \langle \hat{O}(\mathbf{x}) \rangle = \langle \exp \left\{ i\theta(\mathbf{x}) - \int d^3x' \, b^m(\mathbf{x}') \cdot \left[ A(\mathbf{x}') - \tilde{\Phi}(\mathbf{x}') \right] \right\} \rangle. (5.226)
\]
The important question is whether this is zero in the high-temperature, disordered phase of the superconductor [39, 40]. The answer is, unfortunately, negative. We have observed before that the vortex lines merely change the mass square in (5.217) from \(m^2\) to \(m_A^2 + m_{\tilde{a}}^2\). This does not modify the expression (5.223). Hence the correlation function has the same type of large-distance limit as in (5.223) as before, implying that (5.226) is again nonzero and thus capable of distinguishing the disordered from the ordered phase.

The reason why (5.223) is the same in both phases is very simple: It lies in the decoupling of the transverse \(\nabla \times \tilde{a}\) from the longitudinal field \(b^m\) in Eq. (5.223), so that the asymptotic behavior is unaffected by a change in the mass of \(\tilde{a}\).

\textit{c) Disorder Parameter}

The only way to judge the order of the superconductor is to use the disorder field theory and define a disorder parameter whose expectation value is zero for the low-temperature, ordered phase and nonzero for the high-temperature, disordered phase. For a superconductor, the disorder Hamiltonian was written down in (5.165). Recalling (5.172) we might at first consider extracting the disorder parameter from a large-distance limit of the correlation function
\[
G_D(\mathbf{x}_2, \mathbf{x}_1) = \langle \phi(\mathbf{x}_2) \phi^*(\mathbf{x}_1) \rangle. (5.227)
\]
This, however, would not possess the gauge invariance (5.167) of the disorder Hamiltonian (5.165). An invariant expression is obtained by inserting a factor of the type used in (5.209)
\[
G_D(\mathbf{x}_2, \mathbf{x}_1) = \langle \phi(\mathbf{x}_2) e^{-i \int d^3x \, b^m(\mathbf{x}) \tilde{a}(\mathbf{x}) \phi^*(\mathbf{x}_1)} \rangle, (5.228)
\]
where $b^m(x)$ is again the $\delta$-function (5.179) along the line $L_{x_1}^{x_2}$ connecting $x_1$ with $x_2$. The phase factor ensures the gauge invariance under (5.167). In the hydrodynamic limit, (5.228) becomes

$$G_{\tilde{D}}(x_2, x_1) = \langle e^{i \int d^3 x \, b^m(x) \tilde{a}(x)} e^{-i \int d^3 x \, \tilde{a}(x)} \rangle,$$

where $b^m(x)$ is again the $\delta$-function (5.179) along the line $L_{x_1}^{x_2}$ connecting $x_1$ with $x_2$. This can be rewritten as similar to (5.210) as

$$G_{\tilde{D}}(x_2, x_1) = \langle e^{i \int d^3 x \, b^m(x) [\nabla \theta(x) - \tilde{a}(x) - \theta^v(x)]} \rangle,$$

which now defines a disorder parameter of the superconductor [compare (5.185)]

$$\tilde{D} \equiv \langle \tilde{D}(x) \rangle = \langle e^{i \int d^3 x' \, \delta(x'; L) [\nabla \theta(x) - \tilde{a}(x) - \theta^v(x)]} \rangle,$$

where the line $L$ imports the flux from infinity to $x$.

Thus we must study the energy

$$\tilde{\beta} H_{SC}^{hy, D} = \int d^3 x \left\{ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 + \frac{i}{2} \tilde{a} \cdot (\nabla \times A) + \frac{m_a^2}{2 \beta m_A^2} (\nabla \theta - \theta^v - \tilde{a})^2 + b^m \cdot (\nabla \theta - \tilde{a}) \right\},$$

Integrating out the $A$-field in (5.232) makes the $\tilde{a}$-field massive and the Hamiltonian becomes

$$\tilde{\beta} H_{SC}^{hy, D} = \int d^3 x \left\{ \frac{1}{2 \beta m_A^2} (\nabla \times \tilde{a})^2 + m_a^2 \tilde{a}^2 \right\}$$

$$+ \frac{m_a^2}{2 \beta m_A^2} (\nabla \theta - \theta^v - \tilde{a})^2 + b^m \cdot (\nabla \theta - \tilde{a}) \right\},$$

where $m_a$ is the mass parameter in Eq. (5.161), although it does not coincide with the mass of the $\tilde{a}$-field as it did there.

As usual, we introduce an auxiliary field $b$ to rewrite the last two terms of (5.233) in the form

$$\int d^3 x \left[ \frac{\tilde{\beta} m_a^2}{2 m_a^2} (b - b^m)^2 + i b \cdot (\nabla \theta - \tilde{a}) \right],$$

and further as

$$\int d^3 x \left[ \frac{\tilde{\beta} m_a^2}{2 m_a^2} (\nabla \times a - b^m)^2 + i a \cdot (\nabla \times \tilde{a}) + j^v \right] + \frac{\tilde{\beta} \epsilon}{2 j^v}.$$

We have added a core energy to simplify the following discussion. 
In the low-temperature phase, there are no vortices in the superconductor but prolific vortices in the dual formulation whose vortex density is $j^v$, so that we can integrate out $j^v$ in (5.235) as if it were an ordinary Gaussian field. This gives rise to a mass term for the $a$-field, so that the Hamiltonian (5.233) becomes

$$\tilde{\beta} \tilde{H}_{\text{hy},D'}^{\text{SC}} = \int d^3x \left\{ \frac{1}{2\beta m_A^2} \left[ (\nabla \times \tilde{a})^2 + m_a^2 \tilde{a}^2 \right] + \frac{\beta m_A^2}{2m_a^2} \left[ (\nabla \times a - b^m)^2 + m_a^2 a^2 \right] + i \tilde{a} \cdot \nabla \times a \right\}. \quad (5.236)$$

Upon integrating out the $\tilde{a}$-field, we obtain

$$\tilde{\beta} \tilde{H}_{\text{hy},D'}^{\text{SC}} = \int d^3x \left\{ \frac{1}{2\beta m_A^2} \left[ (\nabla \times a - b^m)^2 + m_a^2 a^2 \right] \right\} + \Delta H, \quad (5.237)$$

where

$$\Delta H = \frac{\beta m_A^2}{2} \int d^3x \nabla \times a \left[ 1 + \sum_{n=1}^{\infty} \frac{\nabla^2}{m_A^2} \right] \nabla \times a. \quad (5.238)$$

If we forget this term for a moment we derive from the Hamiltonian (5.237) the correlation function

$$G_{\tilde{D}}(x_1, x_2) \sim e^{-\beta m_A^2} \int d^3x \left[ b^m \nabla \times a - \frac{\nabla \cdot b^m}{\nabla \times b^m} \right], \quad (5.239)$$

As in Eq. (5.202), the mass of $a$ gives the line $L_{x_1}^{x_2}$ in $b^m(x)$ an energy proportional to its length, so that the disorder correlation function (5.230) (5.239) goes to zero at large distances.

This result is unchanged by the omitted term (5.238). By expanding this in powers of $\nabla^2$, it becomes

$$\Delta H = \frac{\beta}{2} \int d^3x \nabla \times a \left[ 1 + \sum_{n=1}^{\infty} \frac{\nabla^2}{m_A^2} \right] \nabla \times a, \quad (5.240)$$

we see that this term changes only the dispersion of the $a$-field, but not its mass.

In the high-temperature phase, there are no dual vortices so that the $a$-field remains massless, and the correlation function is given by an expression like (5.199):

$$G_{\tilde{D}}(x_2, x_1) \approx e^{-\frac{\beta m_A^2}{2} \int d^3x \left[ b^m \nabla \times b^m - \frac{\nabla \cdot b^m}{\nabla \times b^m} \right]} = e^{-\frac{\beta m_A^2}{2} \int d^3x \nabla \cdot b^m - \frac{\nabla \times b^m}{\nabla \times b^m}}. \quad (5.241)$$

This has the same constant large-distance behavior as (5.200) which is independent of the shape of $L_{x_1}^{x_2}$, implying a nonzero disorder parameter (5.231). The monopole gauge invariance is unbroken in this phase.

Thus (5.231) is a good disorder parameter for the superconducting phase transition.
c) Another Disorder Parameter

At this point we are reminded of the discussion of the behavior of the right-hand side of (5.217) and realize that the same large-distance behaviors as in (5.239) and (5.241) would arise if the Meissner mass \( m_A \) of the vector potential in (5.217) would not be replaced, in the high-temperature, disordered phase, by the mass \( m_a \) of the \( a \)-field, but by the zero mass of the magnetic vector potential \( A(x) \). Then the correlation function (5.217) in the disordered phase would be the same as in (5.241)

\[
G_D(x_1, x_2) \sim e^{-\frac{\tilde{\beta} m_A^2}{2} \int dx \left[ \frac{1}{2} b^m(x)^2 - \frac{1}{2} \nabla \times (\nabla \times b^m) \right]}
\]

and thus have the same long-distance behavior as (5.239), i.e., go to a nonzero constant for \( |x_1 - x_2| \to \infty \). This behavior would be found if the correlation function is defined by an expression like (5.216), but with the field \( \tilde{a}(x) \) replaced by the magnetic vector potential \( A(x) \):

\[
G_{\tilde{a}}(x_2, x_1) = \left\langle e^{-\frac{\tilde{\beta} m_A^2}{2} \int dx \left\{ \frac{1}{2} b^m(x)^2 - b^m(x)(\nabla \times A(x)) \right\}} \right\rangle.
\]

The singular line \( L_{x_1}^{x_2} \) in \( b^m(x) \) [recall (5.179)] is taken to be the straight line connecting \( x_1 \) with \( x_2 \), as in (5.216).

The corresponding Hamiltonian looks like (5.214), but with the magnetic gauge field \( b^m(x) \) inserted into the magnetic gradient term rather than the gradient term of the field \( \tilde{a}(x) \):

\[
\tilde{\beta} H_{\text{hy}}^{\text{SC}} = \int d^3x \left[ \frac{1}{2\tilde{\beta} m_A^2} (\nabla \times \tilde{a})^2 - i\tilde{a} \cdot (\nabla \times A) + \frac{\tilde{\beta}^2}{2} (\nabla \times A - b^m)^2 - i\tilde{a} \cdot \tilde{j} + \frac{\tilde{\beta} c}{2} \tilde{\epsilon} \tilde{j}^2 \right].
\]

The correlation function (5.245) defines a disorder parameter as the expectation value

\[
D = \langle D(x) \rangle = \left\langle e^{-\frac{1}{\tilde{\beta} m_A} \int dx \left\{ \frac{1}{2} b^m(x)^2 - b^m(x)(\nabla \times A(x)) \right\}} \right\rangle,
\]

where

\[
b_m^m(x) = \delta(x; \tilde{L}_x)
\]

is singular on a straight line \( \tilde{L}_x \) from \( x \) to infinity.

5.4 Order of Superconducting Phase Transition and Tricritical Point

Most experimental data obtained for the superconducting phase transition since its discovery by Kamerlingh Onnes in 1908 are fitted very well by the BCS theory (recall Fig. 5.8). In the neighborhood of the critical point, this is approximated by the Ginzburg-Landau Hamiltonian (5.143) [41]. One may usually neglect fluctuations of the Ginzburg-Landau order field \( \phi(x) \), which is why one speaks of mean-field results.
5.4 Order of Superconducting Phase Transition and Tricritical Point

5.4.1 Size of Fluctuations

The reason why these are so accurate was first explained by Ginzburg [42] who estimated the temperature interval $\Delta T_G$ around $T_c$ for which fluctuations can be important. Actually, his criterion cannot be applied to superconductors, as has often been done, but only to systems with a real order parameter. Since superconductors have a complex order parameter, one must apply a different criterion which has only been found recently [43]. If the order parameter has a symmetry $O(N)$, the true fluctuation interval $\Delta T_{GK}$ is by a factor $N^2$ larger than Ginzburg’s estimate $T_G$. The fluctuations cause a divergence in the specific heat at $T_c$ very similar to the divergence observed in the $\lambda$-transition of superfluid helium (recall Fig. 5.1). This interval is in all transitions of traditional superconductors too small to be resolved [42, 43], so that it was not astonishing that no fluctuations were observed [44, 45].

5.4.2 First- or Second-Order Transition?

In 1972, however, the order of the superconducting phase transition became a matter of controversy after a theoretical paper by Halperin, Lubensky, and Ma [46] predicted that the transition should really be of first order. The argument was based on an application of renormalization group methods [47] to the partition function

$$Z_{GL} = \int D\psi D\psi^* DA \Phi_T[A] e^{-\tilde{\beta}H_{GL}[\psi,\psi^*,A]}$$

(5.247)

associated with the Ginzburg-Landau Hamiltonian (5.140) in $4 - \epsilon$ dimensions. The technical signal for the first-order transition was the nonexistence of an infrared-stable fixed point in the renormalization group flow [48] of the coupling constants $\epsilon$ and $g$ as a function of the renormalization scale. The fact that all experimental observations indicated a second-order transition was explained by the fact that the fluctuation interval $\Delta T_{GK}$ was too small to be detected. Since then there has been much work [49] trying to find an infrared-stable fixed point by going to higher loop orders or by different resummations of the divergent perturbation expansions, with little success. This controversy was resolved only 10 years later in 1982 by the author [50] who demonstrated that superconductors can have first- and second-order transitions, separated by a tricritical point.

With the advent of modern high-$T_c$ superconductors, the experimental situation has been improved. The temperature interval of large fluctuations is now broad enough to observe critical properties beyond the mean-field approximation. Several experiments have found a critical point of the XY universality class [52]. In addition, there seems to be recent evidence for an additional critical behavior associated with the so-called charged fixed point [53]. In future experiments it will be important to understand the precise nature of the critical fluctuations.

Starting point of the theoretical discussion is the Ginzburg-Landau Hamiltonian (5.140). It contains the field $\psi(x)$ describing the Cooper pairs, and the vector potential $A(x)$. Near the critical temperature, but outside the narrow interval $\Delta T_{GK}$
of large fluctuations, the energy (5.140) describes well the second-order phase transition of the superconductor. It takes place when $\tau$ drops below zero where the pair field $\psi(x)$ acquires the nonzero expectation value $\rho_0 = \sqrt{-\tau/g}$. The properties of the superconducting phase are approximated well by the energy (5.145). The Meissner-Higgs mass term in (5.145) gives rise to a finite penetration depth of the magnetic field $\lambda = 1/m_A = 1/\rho_0 q$.

By expanding the Hamiltonian (5.143) around (5.145) in powers of the fluctuations $\delta \rho \equiv \rho - \rho_0$, we find that the $\rho$-fluctuations have a quadratic energy

$$ H_{\delta \rho} = \frac{1}{2} \int d^3x \left[ (\delta \rho)^2 - 2\tau (\delta \rho)^2 \right], \quad (5.248) $$

implying that these have a finite coherence length $\xi = 1/\sqrt{-2\tau}$.

The ratio of the two length scales

$$ \kappa \equiv \lambda / \sqrt{2} \xi, \quad (5.249) $$

which for historic reasons [54] carries a factor $\sqrt{2}$, is the Ginzburg parameter whose mean field value is $\kappa_{MF} \equiv \sqrt{g/q^2}$. Type I superconductors have small values of $\kappa$, type-II superconductors have large values. At the mean-field level, the dividing line lies at $\kappa = 1/\sqrt{2}$.

### 5.4.3 Partition Function of Superconductor with Vortex Lines

The higher operating temperatures in the new high-$T_c$ superconductors make field fluctuations important. These can be taken into account by calculating the partition function and field correlation functions from the functional integral [compare (5.149)] or, after the field decomposition (5.142),

$$ Z_{GL} = \int \mathcal{D}\rho \mathcal{D}A \Phi [\rho, A] \sum_{\{S\}} \Phi [\theta^S] \int \mathcal{D}\theta e^{-\beta H_{GL}[\rho, \theta, A]} \quad (5.250) $$

This can be approximated by the hydrodynamic formulation (5.149). From now on we use natural temperature units where $k_B T = 1$ and omit all tildes on top of $\tilde{\rho}$, $\tilde{\theta}$, $T$, etc., for brevity, so that we shall rewrite (5.149):

$$ Z_{hy}^{SC} = \int \mathcal{D}A \Phi [A] \sum_{\{S\}} \Phi [\theta^S] \int_{-\infty}^{\infty} D\theta e^{-\beta H_{hy}^{SC}}. \quad (5.251) $$

As described above, all analytic approximations to $Z_{GL}$ investigated since the initial work [46] have had difficulties in accounting for the order of the superconducting phase transition. Let us recall the simplest argument suggesting a first-order transition. One performs a mean-field approximation in the pair field $\rho$ and ignores the effect of vortex fluctuations, setting $\theta^S \equiv 0$ in the Hamiltonian (5.143), so that it becomes (written without wiggles on top of $\rho$, $\theta$, and $\theta^S$)

$$ H_{GL}^{app} \approx \int d^3x \left[ \frac{\rho^2}{2} (\nabla \theta - qA_L)^2 + \frac{1}{2} (\nabla \rho)^2 + V(\rho) + \frac{1}{2} (\nabla \times A)^2 + \frac{\rho^2}{2} q^2 A_T^2 \right]. \quad (5.252) $$

H. Kleinert, GRAVITY WITH TORSION
The approximate sign has the following reason. We have found it useful to perform another approximation: separate $A$ into longitudinal and transverse parts $A_L$ and $A_T$ as defined in Eq. (5.124). If $\rho$ were a constant and not a field this separation would be exactly possible. Due to the $x$-dependence, however, there will be corrections proportional to the gradient of $\rho(x)$ which we shall ignore, assuming sufficiently smooth fields $\rho(x)$.

After these approximations we can integrate out the Gaussian phase fluctuations $\theta(x)$ in the partition function (5.247) and obtain

$$Z_{GL}^{app'} = \text{Det}^{-1/2}[ - \nabla^2] \int D\rho \; D\Phi T[A] e^{-\bar{\beta} H_{GL}^{app'}} ,$$

(5.253)

with

$$H_{GL}^{app'} = \int d^3x \left[ \frac{1}{2} (\nabla \rho)^2 + V(\rho) + \frac{1}{2} (\nabla \times A)^2 + \frac{\rho^2 q^2}{2} A_T^2 \right] .$$

(5.254)

The fluctuations of the vector potential are also Gaussian and can be integrated out in (5.253) yielding

$$\tilde{Z}_{GL}^{app'} = \text{Det}^{-1/2}[ - \nabla^2] \text{Det}^{-1}[ - \nabla^2 + \rho^2 q^2] \int D\rho e^{-\bar{\beta} H_{GL}^{app'}} ,$$

(5.255)

where

$$H_{GL}^{app'} = \int d^3x \left[ \frac{1}{2} (\nabla \rho)^2 + V(\rho) \right] .$$

(5.256)

### 5.4.4 First-Order Regime

Assuming again that $\rho$ is smooth, the functional determinant $\text{Det}^{-1}[ - \nabla^2 + \rho^2 q^2]$ may be done in the Thomas-Fermi approximation [55] where it yields

$$\text{Det}^{-1}[ - \nabla^2 + \rho^2 q^2] = e^{ - \text{Tr log}[ - \nabla^2 + \rho^2 q^2]} \approx e^{-V \int [d^3k/(2\pi)^3](k^2 + \rho^2 q^2)} = e^{\rho^3 q^3/6\pi} .$$

(5.257)

From now on we shall use natural units for the energy and measure energies in units of $k_B T$. Then we can set $\beta$ in the Boltzmann factors (5.254) and (5.256). Thus the $A$-fluctuations contribute simply a cubic term to the potential $V(\rho)$ in Eq. (5.144), changing it to

$$\hat{V}(\rho) = \frac{\tau}{2} \rho^2 + \frac{g}{4} \rho^4 - \frac{c}{3} \rho^3 , \quad c \equiv \frac{q^3}{2\pi} .$$

(5.258)

The cubic term generates, for $\tau < c^2/4g$, a second minimum in the potential $\hat{V}(\rho)$ at

$$\hat{\rho}_0 = \frac{c}{2g} \left( 1 + \sqrt{1 - \frac{4\tau g}{c^2}} \right) ,$$

(5.259)

as illustrated in Fig. 5.9.
Figure 5.9 Potential for the order parameter \( \rho \) with cubic term. A new minimum develops around \( \rho_1 \) causing a first-order transition for \( \tau = \tau_1 \).

If \( \tau \) decreases below

\[
\tau_1 = \frac{2c^2}{9g},
\]

the new minimum drops below the minimum at the origin, so that the order parameter jumps from \( \rho = 0 \) to

\[
\rho_1 = \frac{2c}{3g}
\]

in a phase transition. At this point, the coherence length of the \( \rho \)-fluctuations

\[
\xi = \frac{1}{\sqrt{\tau + 3g\rho^2 - 2c\rho}}
\]

has the finite value (the same as the fluctuations around \( \rho = 0 \))

\[
\xi_1 = \frac{3}{c} \sqrt{\frac{g}{2}}.
\]

The fact that the transition occurs at a finite \( \xi = \xi_1 \neq 0 \) indicates that the phase transition is of first order. In a second-order transition, \( \xi \) would go to infinity as \( T \) approaches \( T_c \).

This conclusion is reliable only if the jump of \( \rho_0 \) is sufficiently large. For small jumps, the mean-field discussion of the energy density (5.258) cannot be trusted. At a certain small \( \rho_0 \), the transition becomes second-order. The change of the order is caused by the neglected vortex fluctuations in (5.254). We must calculate the partition function (5.253) including the sum over vortex gauge fields \( \theta'(x) \), with a Hamiltonian equal to (5.143) but with omitted wiggles:

\[
H_{\text{GL}} = \int d^3x \left[ \frac{\rho^2}{2} (\nabla \theta - \theta' - qA)^2 + \frac{1}{2} (\nabla \rho)^2 + V(\rho) + \frac{1}{2} (\nabla \times A)^2 \right].
\]

If we now integrate out the \( \theta \)-fluctuations, and assume smooth \( \rho \)-fields, we obtain the partition function (5.253) extended by the sum over vortex gauge fields \( \theta'(x) \), and with the Hamiltonian

\[
H'_{\text{GL}} = \int d^3x \left[ \frac{1}{2} (\nabla \rho)^2 + V(\rho) + \frac{1}{2} (\nabla \times A)^2 + \frac{\rho^2}{2} (qA - \theta')^2 \right].
\]

We may now study the vortex fluctuations separately by defining a partition function of vortex lines in the presence of a fluctuating \( A \)-field for smooth \( \rho(x) \):

\[
Z_{\theta',A}[\rho] \equiv \int D\theta' D\rho \exp \left\{ -\frac{1}{2} \int d^3x \left[ (\nabla \times A)^2 + \frac{\rho^2}{2} (qA - \theta')^2 \right] \right\}.
\]
The transverse part of the vortex gauge field $\mathbf{v}$ is defined as in (5.124). We have abbreviated the sum over the jumping surfaces $S$ with vortex gauge fixing $\sum_{\{S\}} \Phi[\mathbf{v}]$ defined in (5.30) as $\int \mathcal{D}\mathbf{v}_T$. In addition, we have fixed the vector potential to be transverse and indicated this by the functional integration symbol $\int \mathcal{D}A_T$. 

5.4.5 Vortex Line Origin of Second-Order Transition

The important observation is now that for smooth $\rho$-fields, this partial partition function possesses a second-order transition of the XY-model type if the average value of $\rho$ drops below a critical value $\rho_c$. To see this we integrate out the $A$-field and obtain

$$Z_{\Phi,A}[\rho] \approx \exp \left[ \int d^3x \frac{\rho^3}{6\pi} \right] \int \mathcal{D}\mathbf{v}_T \exp \left[ \frac{\rho^2}{2} \int d^3x \left( \frac{1}{2} \mathbf{v}_T^2 - \mathbf{v}_T \cdot \frac{\rho^2 q^2}{\nabla^2 + \rho^2 q^2} \mathbf{v}_T \right) \right].$$

(5.266)

The first factor yields, once more, the cubic term of the potential (5.258). The second factor accounts for the vortex loops. The integral in the exponent can be rewritten as

$$\frac{\rho^2}{2} \int d^3x \left( \mathbf{v}_T \cdot \frac{-\nabla^2}{\nabla^2 + \rho^2 q^2} \mathbf{v}_T \right).$$

(5.267)

Integrating this by parts, and using identity

$$\int d^3x \nabla_i A \nabla_i B = \int d^3x \left[ (\nabla \times A)(\nabla \times B) + (\nabla \cdot A)(\nabla \cdot B) \right],$$

(5.268)

together with the transversality property $\nabla \cdot \mathbf{v}_T = 0$ and the curl relation $\nabla \times \mathbf{v}_T = j^v$ of Eq. (5.25), the partition function (5.266) without the prefactor takes the form

$$\tilde{Z}_{\Phi,A}[\rho] \approx \int \mathcal{D}\mathbf{v}_T \exp \left[ -\frac{\rho^2}{2} \int d^3x \left( \frac{1}{\nabla^2 + \rho^2 q^2} j^v \right) \right].$$

(5.269)

This is the partition function of a grand-canonical ensemble of closed fluctuating vortex lines. The interaction between them is of the Yukawa type with a finite range equal to the penetration depth $\lambda = 1/\rho q$.

It is well-known how to compute pair and magnetic fields of the Ginzburg-Landau theory for a single straight vortex line from the extrema of the energy density [44]. In an external magnetic field, there exist triangular and various other regular arrays of vortex lines, such as vortex lattices. In the presence of impurities, there are vortex glasses, etc. The study of such phases and the transitions between them is an active field of research [56].

In the core of each vortex line, the pair field $\rho$ goes to zero over a distance $\xi$. If we want to sum over a grand-canonical ensemble of fluctuating vortex lines of any shape in the partition function (5.269), the space dependence of $\rho$ causes complications. These can be avoided by an approximation, in which the system is...
placed on a simple-cubic lattice of spacing $a = \alpha \xi$, with $\alpha$ of the order of unity, and a fixed value $\rho = \tilde{\rho}_0$ given by Eq. (5.259). Thus we replace the partial partition function (5.269) approximately by

$$Z_{\Theta^\alpha \Lambda}[\tilde{\rho}_0] \approx \sum_{\{l: \nabla \cdot l = 0\}} \exp \left[ -\frac{4\pi^2 \tilde{\rho}_0^2 a}{2} \sum_x l(x) v_{\rho_0q}(x - x') l(x') \right]. \quad (5.270)$$

The sum runs over the discrete versions of the vortex density $j^v$ in (5.269). Recalling (5.29) and (5.41), these are $2\pi$ times the integer-valued vectors $l(x) = (l_1(x), l_2(x), l_3(x)) = \nabla \times n(x)$, where $\nabla$ denotes the lattice derivative (5.33). Being lattice curls of the integer vector field $n(x) = (n_1(x), n_2(x), n_3(x))$, they satisfy $\nabla \cdot l(x) = 0$. This condition restricts the sum over $l(x)$-configurations in (5.270) to all non-selfbacktracking integer-valued closed loops. The partition function (5.265) has precisely the form discussed before in Eq. (5.47) with $\rho_0q$ playing the role of the Yukawa mass $m$ in (5.47). The lattice partition function (5.270) has therefore a second-order phase transition in the universality class of the XY-model. The transition temperature was plotted in Fig. 5.7.

5.4.6 Tricritical Point

Comparing (5.270) with the partition function (5.98) of the Yukawa loop gas, we conclude that there is a second-order phase transition when [compare (5.100)]

$$4\pi^2 a \tilde{\rho}_0^2 v_{\rho_0q}(0) \approx T_c \approx 3. \quad (5.271)$$

Using the analytic approximation (5.103), we may write this as

$$4\pi^2 a v_0(0) \frac{\tilde{\rho}_0^2}{\sigma a^2 \rho_0q^2 / 6 + 1} \approx T_c \approx 3, \quad (5.272)$$

or

$$\frac{\tilde{\rho}_0^2 a}{\sigma a^2 \rho_0q^2 / 6 + 1} \approx \frac{r}{3}, \quad (5.273)$$

where $r = 9/4\pi^2 v_0(0) \equiv 0.90$. The solution is

$$\tilde{\rho}_0 \approx \frac{1}{\sqrt{3} \sigma} \sqrt{\frac{r}{1 - \sigma r^2e^2 / 18}}. \quad (5.274)$$

Replacing here $a$ by $\alpha \xi_1 = \alpha(3/c)\sqrt{g/2}$ of Eq. (5.262), and $\tilde{\rho}_0$ by $\rho_1 = 2c/3g$ of Eq. (5.261). Inserting further $c = q^3/2\pi$ of Eq. (5.258), we find the equation for the mean-field Ginzburg parameter $\kappa_{\text{MF}} = \sqrt{g/q^2}$ [recall (5.249)]:

$$\kappa_{\text{MF}}^3 + \frac{\alpha^2 \sigma \kappa_{\text{MF}}^2}{3} - \frac{\sqrt{2\alpha}}{\pi r} = 0. \quad (5.275)$$

H. Kleinert, GRAVITY WITH TORSION
5.4 Order of Superconducting Phase Transition and Tricritical Point

For the best value $\sigma \approx 1.6$ in the approximation (5.103), and $r \approx 0.9$, and the rough estimate $\alpha \approx 1$, the solution of this equation yields the tricritical value

$$\kappa_{MF}^{\text{tric}} \approx 0.82/\sqrt{2}. \quad (5.276)$$

In spite of the roughness of the approximations, this result is very close to the value

$$\kappa_{MF}^{\text{tric}} = \frac{3\sqrt{3}}{2\pi} \sqrt{1 - \frac{4}{9} \left( \frac{\pi}{3} \right)} \approx \frac{0.80}{\sqrt{2}} \quad (5.277)$$

derived from the dual theory in [23]. The approximation (5.276) has three uncertainties. First, the identification of the effective lattice spacing $a = \alpha \xi$ with $\alpha \approx 1$; second the associated neglect of the $x$-dependence of $\rho$ and its fluctuations, and third the localization of the critical point of the XY-model type transition in Eqs. (5.104) and the ensuing (5.273).

5.4.7 Disorder Theory

In the disorder theory (5.166) it is much easier to prove that superconductors can have a first- and a second-order phase transition, depending on the size of the Ginzburg parameter $\kappa$ defined in Eq. (5.249). Before we start let us rewrite the disorder theory in a more convenient way. As before, we decompose the complex disorder field $\phi$ as $\phi = \rho e^{i\theta}$. In the partition function (5.168), this changes the measure of functional integration from $\int D\phi \int D\phi^*$ to $\int D\rho \int D\theta$. Now we fix the gauge by absorbing the phase $\theta$ of the field into $\tilde{a}$ by a gauge transformation (5.167). This brings the Hamiltonian (5.166) to the form

$$\tilde{\beta} H_{SC}^{\text{hy}} \sim \int d^3x \left[ \frac{1}{2\beta m_A^2} \left[ (\nabla \times \tilde{a})^2 + m_A^2 \tilde{a}_T^2 \right] + \frac{\rho^2}{2} \left( \tilde{a}_T^2 + \tilde{a}_L^2 \right) + \frac{1}{2} (\nabla \rho)^2 + \frac{\tau}{2} \rho^2 + \frac{g}{4} \rho^4 \right], \quad (5.278)$$

where we have again assumed a smooth $\rho$-field to separate $\rho^2 \tilde{a}^2$ into $\rho^2 \tilde{a}_T^2 + \rho^2 \tilde{a}_L^2$.

The partition function (5.168) reads now

$$Z_{SC}^{\text{dual}} = \int D\rho D\tilde{a} e^{-\tilde{\beta} H_{SC}^{\text{hy}}}.$$ \hfill (5.279)

We may integrate out $\tilde{a}_L$ to obtain a factor $\text{Det}[\rho^2]^{-1/2}$ which removes the factor $\rho$ in the measure of path integration $D\rho$.

Next we integrate out $\tilde{a}_L$ and obtain

$$Z_{SC}^{\text{dual}} = \int D\rho \text{Det}[- \nabla^2 + m_A^2 (1 + \tilde{\beta} \rho^2)] e^{-\tilde{\beta} H_{SC}^{\text{hy}}}, \quad (5.280)$$

with

$$\tilde{\beta} H_{SC}^{\text{hy}} = \int d^3x \left[ \frac{1}{2} (\nabla \rho)^2 + \frac{\tau}{2} \rho^2 + \frac{g}{4} \rho^4 \right] + \text{Tr} \log[- \nabla^2 + m_A^2 (1 + \tilde{\beta} \rho^2)]. \quad (5.281)$$
In the superconducting phase, there are only a few vortex lines and the disorder field $\rho$ of vortex lines fluctuates around zero. In this phase we may expand the trace log into powers of $\rho^2$. The first expansion term is proportional to $\rho^2$ and renormalizes $\tau$ in the Hamiltonian (5.281), corresponding to a shift in the critical temperature.

The second expansion term is approximately given by

$$-\beta^2 m_A^4 \int d^3x \int \frac{d^3k}{(2\pi)^3} \frac{1}{(k^2 + m_A^2)^2} \rho^4 \propto -m_A^3 \int d^3x \rho^4. \quad (5.282)$$

This term lowers the interaction term $(g/4)\rho^4$ in the Hamiltonian (5.281). An increase in $m_A$ corresponds to a decrease of the penetration depth in the superconductor, i.e. to materials moving towards the type-I regime. At some larger value of $m_A$, the $\rho^4$-term vanishes and the disorder field theory requires a $\rho^6$-term to stabilize the fluctuations of the vortex lines. In such materials, the superconducting phase transition turns from second to first order.

A more quantitative version of this argument was used in Ref. [50] to show the existence of the tricritical point and its location at the Ginzburg parameter $\kappa \equiv g/q^2$ in Eq. (5.277), which agrees well with a recent Monte Carlo value $(0.76 \pm 0.04)/\sqrt{2}$ of Ref. [51].

### 5.5 Vortex Lattices

The model action (5.22) represents the gradient energy in superfluid $^4$He correctly only in the long-wavelength limit. The neutron scattering data yield the energy spectrum $\omega = \epsilon(k)$ shown in Fig. 5.2.

To account for this, the energy should be taken as follows:

$$H_{NG} = \frac{1}{2} \int d^3x (\nabla \theta - \theta^\nu) \frac{\epsilon^2(-i \nabla)}{-\nabla^2} (\nabla \theta - \theta^\nu). \quad (5.283)$$

The roton peak near $2\pi A^{-1}$ gives rise to a repulsion between opposite vortex line elements at the corresponding distance. If a layer of superfluid $^4$He is diluted with $^3$He, the core energy of the vortices decreases, the fugacity $g$ and the average vortex number increases. For a sufficiently high average spacing, a vortex lattice forms. In this regime, the superfluid has three transitions when passing from zero temperature to the normal phase. There is first a condensation process to a vortex lattice, then a melting transition of this lattice into a fluid of bound vortex-antivortex pairs, and finally a pair-unbinding transitions of the Kosterlitz-Thouless type [57, 58]. The latter two transitions have apparently been seen experimentally [59] (see Figs. 5.10 and (5.11).
Appendix 5A  Vortex Lines in Superfluid

Figure 5.10 Phase diagram of a two-dimensional layer of superfluid $^4$He. At a higher fugacity $y > y^*$, an increase in temperature causes the vortices to first condense to a lattice and to undergo a Kosterlitz-Thouless vortex unbinding transition only after a melting transition.

Figure 5.11 Experimental phase diagram of a two-dimensional layer of superfluid $^4$He diluted by $^3$He which decreases the fugacity and separates the vortex melting transition from the Kosterlitz-Thouless transition.

Appendix 5A  Vortex Lines

Here we derive some properties of vortex lines obtained by extremizing the Hamiltonian (5.6). For simplicity, we shall focus attention only upon straight vortex line. Such a line can be obtained as a cylindrical solution to the field equation

$$-\nabla^2 \phi + \tau \phi + \lambda|\phi|^2 \phi = 0,$$

which minimizes the energy (5.6).
Decomposing $\phi$ into its polar components as in (5.10), $\phi(x) = \rho e^{i\theta(x)}$, and ignoring for a moment the vortex gauge field, the real and imaginary parts of this equation read

$$[-\nabla^2 + (\nabla \theta)^2 + \tau + \lambda \rho^2] \rho = 0$$

(5A.2)

and

$$\nabla j_s(x) = 0, \quad j_s(x) \equiv \rho^2 \nabla \theta(x) = 0.$$  

(5A.3)

The latter equation is the statement of current conservation for the current density of superflow $j_s(x)$.

Current conservation can be ensured by a purely circular flow in which $\rho$ depends on the distance $r$ from the cylindrical axis and the phase $\theta$ is an integer multiple of the azimuthal angle in space, $\theta = n \arctan(x_2/x_1)$. Then (5A.2) reduces to the radial differential equation

$$-\left(\frac{\partial^2}{\partial \tilde{r}^2} + \frac{1}{\tilde{r}} \frac{\partial}{\partial \tilde{r}} - \frac{n^2}{\tilde{r}^2}\right) \rho + \lambda (\rho^2 - \rho_0^2) \rho = 0,$$

(5A.4)

where $\rho_0 = \sqrt{-\tau/\lambda} = \sqrt{(1-T/T_c)\tau/\lambda}$ [compare (5.7), (5.8)].

In order to solve this equation it is convenient to go to reduced quantities $\tilde{r}, \tilde{\rho}$, which measure the distance $r$ in units of the coherence length

$$\xi = \sqrt{\frac{1}{\mu_0} \frac{1}{\sqrt{1-T/T_c}}}$$

(5A.5)

and the size of the order parameter $\rho$ in units of $\rho_0$, i.e., we introduce

$$\bar{x} = x/\xi, \quad \tilde{r} = r/\xi$$

(5A.6)

$$\tilde{\rho} = \rho/\rho_0.$$  

(5A.7)

Then (5A.4) takes the form

$$\left[-\left(\frac{\partial^2}{\partial \tilde{r}^2} + \frac{1}{\tilde{r}} \frac{\partial}{\partial \tilde{r}} - \frac{n^2}{\tilde{r}^2}\right) + (\tilde{\rho}^2 - 1)\right] \tilde{\rho}(\tilde{r}) = 0.$$

(5A.8)

Multiplying this with the phase factor

$$e^{i n \theta} = e^{i n \tan^{-1}(x_2/x_1)},$$

(5A.9)

we see that the complex field $\phi(x)$ has the following small $|x|$ behavior

$$\phi(x) \propto \tilde{r}^n e^{i n \tan^{-1}(x_2/x_1)} = (x_1 + ix_2)^n,$$

(5A.10)
which corresponds to a zero of \( n \)-th order in \( \phi(x) \).

For large \( \bar{r} \gg 1 \), \( \bar{\rho}(\bar{r}) \) approaches the asymptotic value \( \bar{\rho} = 1 \). In fact, Eq. (5A.8) is solved by a large-\( \bar{r} \) expansion:

\[
\bar{\rho}_n(\bar{r}) = 1 - \frac{n^2}{2\bar{r}^2} - \left( n^2 + \frac{1}{8} n^4 \right) \frac{1}{\bar{r}^4} - \left( 8 + 2n^2 + \frac{1}{16} n^4 \right) \frac{n}{\bar{r}^6} + O \left( \frac{1}{\bar{r}^8} \right). \tag{5A.11}
\]

Integrating the differential equation numerically inward, we find the solution displayed on (5.12).

Let us now study the energy of these vortex lines. The calculation can be simplified by a scaling argument: if \( \phi(x) \) is the solution of the differential equation (5A.1), the rescaled solution

\[
\phi_\delta(x) \equiv e^\delta \cdot \phi(x) \tag{5A.12}
\]

must extremize the energy for \( \delta = 0 \). Inserting (5A.12) into (5.6), we calculate

\[
E = \frac{1}{2} \int d^3x \left[ e^{2\delta} |\nabla \phi|^2 + e^{2\delta} \tau |\phi|^2 + \frac{\lambda}{2} e^{4\delta} |\phi|^4 \right]. \tag{5A.13}
\]

Setting the derivative with respect to \( \delta \) equal to zero gives at \( \delta = 0 \)

\[
\frac{1}{2} \int d^3x \left[ |\nabla \phi|^2 + \tau |\phi|^2 + \lambda |\phi|^4 \right] = 0. \tag{5A.14}
\]

Subtracting this from (5A.13) for \( \delta = 0 \) we see that the energy of a solution of the

\[
\text{Figure 5.12} \quad \text{Order parameter } \bar{\rho} = |\phi|/|\phi_0| \text{ around a vortex line of strength } n = 1, 2, 3, \ldots \text{ as a function of the reduced distance } \bar{r} = r/\xi, \text{ where } r \text{ is the distance from the axis and } \xi \text{ the healing length.}
\]

field equation is simply given by

\[
E = -\frac{\lambda}{4} \int d^3x |\phi|^4. \tag{5A.15}
\]

Most of this energy is due to the asymptotic regime where \( \phi \to \phi_0 \), and the energy becomes the condensation energy. Subtracting this from \( E \) we find the energy above the background: due to the presence of the vortex line

\[
E_v = -\frac{\lambda}{4} \int d^3x (|\phi|^4 - |\phi_0|^4). \tag{5A.16}
\]
In terms of the natural units introduced in (5A.6) and (5A.7), this is simply

$$E_v = -f \xi^3 \int d^3 \hat{x} [1 - \bar{\rho}^4(x)].$$  \hspace{1cm} (5A.17)

Going over to cylindrical coordinates $\bar{r}, \theta, z$, the integral becomes, for line of reduced length $L$ along the $z$-direction:

$$2\pi L \int_0^\infty d\bar{r} \bar{r} [1 - \bar{\rho}^4(\bar{r})].$$  \hspace{1cm} (5A.18)

Before inserting the numerical solutions for $\bar{\rho}(\bar{r})$ shown in Fig. 5.12 we note that due to the factor $\bar{r}$, the additional energy comes mainly from the large-$\bar{r}$ regime, i.e., the far zone. In fact, if we insert the leading asymptotic behavior (5A.11), we obtain an integral

$$2\pi n^2 L \int_0^\infty d\bar{r} / \bar{r},$$  \hspace{1cm} (5A.19)

which diverges logarithmically for large $\bar{r}$. An immediate conclusion is that a single vortex line can have a finite energy only in a finite container. If this container is cylindrical of radius $R$, the integral is finite and becomes $4\pi n^2 L \log(R/\xi)$. In an infinite container, straight vortex lines can only exist in pairs of opposite circulation.

Consider now the small-$\bar{r}$ behavior. From (5A.8) we see that close to the origin, $\bar{\rho}(\bar{r})$ behaves like $\bar{r}^n$. Hence $1 - \bar{\rho}^4$ and the energy of a thin cylindrical section of radius $\bar{r}$ grows like $\bar{r}^2$. For increasing $\bar{r}$, the rate of growth rapidly slows down and settles at the asymptotic rate $4\pi n^2 L \times \log(r/\xi)$, where $\xi$ is the coherence length. The proper inclusion of the non-asymptotic behavior gives simply a finite correction to the asymptotic energy and the energy of a vortex line in a container of radius $R$ becomes $4\pi n^2 L \log(R/\xi) + c$. The same result would have been obtained by replacing the integrand $\bar{r} [1 - \bar{\rho}^4(\bar{r})]$ by its asymptotic form $2n^2 / \bar{r}$ and integrating from a radius

$$r_c = \xi e^{-c}$$  \hspace{1cm} (5A.20)

to $R$. The quantity $r_c$ is called the core radius of the vortex line.

The precise numerical evaluation of the differential equation (5A.8) and the integral (5A.18) shows that for the lowest vortex line, $c$ has the value

$$c = 0.385.$$  \hspace{1cm} (5A.21)

Hence the energy of the vortex line per unit length becomes

$$\frac{E_v}{L} = f_c \xi^2 4\pi n^2 [\log(R/\xi) + 0.385].$$  \hspace{1cm} (5A.22)

The logarithmic divergence of the energy has a simple physical meaning. In order to see this let us calculate this energy once more using the original expression (5.6), i.e., without invoking the property (5A.14). It reads, for a cylindrical solution:

$$\frac{E_v}{L} = f_c \xi^2 4\pi \int d\bar{r} \bar{r} \left\{ (\bar{\rho}^2) + \frac{1}{2} (1 - \bar{\rho}^2)^2 + \frac{\bar{\rho}^2}{\bar{r}^2}\right\}.$$  \hspace{1cm} (5A.23)
The first two terms are rapidly converging. Thus the energy of the far-zone resides completely in the last term

\[ \frac{n^2}{\bar{r}^2} \rho^2 \approx \frac{n^2}{\bar{r}^2}. \]  

(5A.24)

This energy is a consequence of the angular behavior of the condensate phase \( \theta = n \tan^{-1}(x_2/x_1) \) around a vortex line. In fact, the term (5A.25) is entirely due to the azimuthal part of the gradient energy \((1/2)\rho^2(\nabla \theta)^2\), i.e., the term which describes the Nambu-Goldstone modes.

The dominance of the energy carried by the phase gradient can also be described in a different and more physical way. In Eq. (5A.3) we have seen that \( j_s(x) = \rho^2 \nabla \theta(x) \) is the current density of superflow, and we may identify the gradient \( \nabla \theta(x) \) as the superflow velocity \( v_s(x) \). In physical units, it is given by \( \hbar \nabla \theta(x) \). Far away from the line it reads explicitly

\[ v_s = \frac{\hbar}{M} n \nabla \arctan \left( \frac{x_2}{x_1} \right) = \frac{\hbar}{M} n \frac{1}{r^2} \left( -x_2, x_1, 0 \right) = \frac{\hbar}{M} n \frac{1}{r} e_\phi, \]  

(5A.25)

where \( e_\phi \) is the unit vector in azimuthal direction. Thus, around every vortex line, there is a circular flow of the superfluid whose velocity decreases like the inverse distance from the line. With the notation (5.15) for the superfluid density, the hydrodynamic energy density of this flow is

\[ E(x) = \frac{\rho_s}{2} v_s^2(x) = \frac{\rho_s}{2} \frac{\hbar^2}{M^2} \frac{n^2}{r^2}. \]  

(5A.26)

This is precisely the dominant third Nambu-Goldstone term in the energy integral (5A.23). Thus the energy of the vortex line is indeed mainly due to the hydrodynamic energy of the superflow around the line.

For the major portion of the fluid, the limiting hydrodynamic limiting expressions (5A.25), (5A.26) give an excellent approximation to these quantities. Only in the neighborhood of the line, i.e., for small radii \( r \leq \xi \), the energy density differs from (5A.26) due to gradients of the size of the field \( |\phi| \). It is therefore suggestive to idealize the superfluid and assume the validity of the pure gradient energy density

\[ E(x) = \frac{\rho_s}{2} v_s^2(x) = \frac{\hbar^2 \rho_s}{2M^2} |\nabla \theta(x)|^2. \]  

(5A.27)

\( \textit{everywhere} \) in space.

The deviations from this law, which become significant only very close to a vortex line, i.e., at distances of the order of the coherence length \( \xi \), are treated \textit{approximately} by simply cutting off the energy integration at the core radius \( r_c \) (5A.20) around the vortex line. In other words, we pretend as though there is no superflow at all within the thin tubes of radius \( \xi \), and assume a sudden onset of idealized flow outside \( r_c \), moving with the limiting velocity (5A.25).
Although the internal part of the thin tube carries no superflow, it nevertheless carries rotational energy. Within the present approximation, this energy is associated with the number $c = 0.385$ in (5A.22). This piece will be called the core energy. The core energy has a physical interpretation. At distances smaller than the core radius, the different parts of the liquid can no longer slip past each other freely. Hence the core of a vortex line is expected to rotate roughly like a solid rod, rather than with the diverging velocity $v_s \sim 1/r$. Indeed, if we use the approximation

$$v_s \approx n \begin{cases} 1/r, & r > \xi \\ (r/\xi)^2, & r \geq \xi, \end{cases}$$

(5A.28)

for a line of vortex strength $n$, the energy density has, for small $r$, precisely the behavior proportional to $r^2$ observed before in the exact expressions (5A.17). Moreover, the energy integration gives

$$n^2 \left[ \int_{1}^{R/\xi} d\rbar \, \frac{1}{r^2} + \int_{0}^{1} d\rbar \, \rbar^2 \right] = n^2[\log(R/\xi) + 0.25],$$

(5A.29)

and we see that the number $c$ for the core energy emerges with this approximation as 0.25, which is of the correct order of magnitude.

To complete our discussion of the hydrodynamic picture, let us calculate the circulation of the superfluid velocity field around the vortex lines:

$$\kappa \equiv \oint_B dx_i v_s^i = \frac{\hbar}{M} n \oint_B dx_i \partial_i \theta = \frac{\hbar}{M} 2\pi n = n \frac{\hbar}{M} = n\kappa_1.$$  

(5A.30)

This integral is the same for any size and shape of the circuit $B$ around the vortex line. Thus the circulation is quantized and always appears in multiplets of $\kappa_1 = \hbar/M \approx 10^{-3} \text{ cm}^2/\text{sec}$. The number $n$ is called vortex strength.

The integral (5A.30) can be transformed into a surface integral, via Stokes’ theorem:

$$\int_{S^B} d^2x \, \nabla \times v_s = \frac{\hbar}{M} 2\pi n,$$

(5A.31)

where $S^B$ is some surface spanned by the circuit $B$ in (5A.30). This integral is the same for any size and shape of $S^B$. From this result we conclude that the third component of the curl of $v_s$ must vanish everywhere except at the origin. There it must have a singularity of such a strength that the two-dimensional integral gives the correct vortex strength. Hence

$$\nabla \times v_s = \frac{\hbar}{M} 2\pi n \delta(2)(\mathbf{x}_\perp) \hat{z},$$

(5A.32)

are the coordinates orthogonal to the vortex line. This is the two-dimensional version of (5.25).

If the nonlinearities of the field are taken into account, the $\delta(2)$-function is really smeared out over a circle whose radius is of order $\xi$. Typically, the term $1/r^2$ in
(5A.26) will be softened to $1/(r^2 + \varepsilon^2)$, in which case the curl of the superfluid velocity becomes

$$\nabla \times \mathbf{v}_s = \frac{\hbar}{M} \frac{2\varepsilon^2}{(r^2 + \varepsilon^2)^2} \hat{z}. \quad (5A.33)$$

The right-hand side is non-zero only within a small radius $r \geq \varepsilon$, where it diverges with the total strength

$$\int d^2x \frac{2\varepsilon^2}{(r^2 + \varepsilon^2)^2} = 2\pi \varepsilon^2 \int dr \frac{2r}{(r^2 + \varepsilon^2)^2} = 2\pi. \quad (5A.34)$$

This shows that (5A.33) is, indeed, a smeared-out version of the singular relation (5A.32).

Because of their rotational properties, vortex lines can be generated experimentally by rotating a vessel with an angular velocity $\Omega$. Initially, the lack of friction will cause the superfluid part of the liquid to remain at rest. This situation cannot, however, persist forever since it is not in a state of thermal equilibrium. After some time, vortex lines form on the walls which migrate into the liquid and distribute evenly. This goes on until their total number is such that the rotational Helmholtz free energy

$$E_\Omega = H - \mathbf{\Omega} \cdot \mathbf{L} \approx \int d^3x \left( \frac{\rho_s}{2} \mathbf{v}_s^2 - \mathbf{\Omega} \cdot \mathbf{x} \times \rho_s \mathbf{v}_s \right) \quad (5A.35)$$

is minimal. This equilibration process has been observed in the laboratory and has even been photographed. This was done using the property that vortex lines trap ions which can be accelerated against a photographic plate.

If we evaluate the energy (5A.35) with the circular velocity field $\mathbf{v}_s(x)$ of a single vortex of Eq. (5A.25), we find that in a cylindrical vessel of radius $R$, the first vortex line $n = 1$ appears at a critical angular velocity

$$\Omega_c = \frac{\kappa_1}{\pi R^2} \log \frac{R}{\xi} \quad (5A.36)$$

and settles on the axis of rotation. It is useful to observe that the vortex lines of higher $n$ are all unstable. Since the energy increases quadratically with $n$, it is favorable for a single line with $n > 1$ to decay into $n$ lines with $n = 1$. When generating vortex lines by stirring a vessel, one may nevertheless be able to create, for a short time, such an unstable line, and to observe its decay.

Notes and References


[10] Note that this configurational entropy cannot be properly accounted for by a model restricted only to circular vortex lines proposed by G. Williams, Phys. Rev. Lett. 59, 1926 (1987).


[19] See p. 503 in the textbook [7] (kl/b1/gifs/v1-503s.html). The high-temperature expansions of the partition function (5.31) and the associated free energy are given in Eqs. (7.42a) and (7.42b), the low-temperature expansion of the free energy in Eq. (7.43).


[23] H. Kleinert, Lett. Nuovo Cimento 35, 405 (1982) (kl/97). The tricritical value $\kappa \approx 0.80/\sqrt{2}$ derived in this paper was confirmed only recently by Monte Carlo simulations [51].

[24] The equation $j^s = -a_T$ is the disorder version of the famous first London equation for the superconductor $j^s = -(q^2n_0/Mc)A_T$ to be discussed further in Section 7.2.

[25] Recall that the density of states per spin direction in the energy interval $E, E + dE$ is given, in proper physical units, by $\mathcal{N} = \int [d^3p/(2\pi\hbar)^3] \delta(p^2/2m - E) = (2me^2/\hbar^2)^{1/2}m/2\pi^2\hbar^2$. On the surface of the Fermi sphere it becomes $mp_F/2\pi^2\hbar^2$. The total density of electrons $n_e$ is $2 \int_0^{E_F} dE \mathcal{N} = p_F^3/3\pi^2$, so that we obtain indeed (5.137).


[28] Gorkov’s derivation was valid only at the mean field level. A modern derivation based on functional integrals [29] permits the inclusion of fluctuations to all orders.


[30] This property of the disorder theory was demonstrated in detail in Ref. [7].


[34] For a derivation see Section 3.10 of the textbook [20].

[35] This is an analog of the Meissner effect in the dual description of superfluid helium.


[43] H. Kleinert, *Criterion for Dominance of Directional over Size Fluctuations in Destroying Order*, Phys. Rev. Lett. 84, 286 (2000) (cond-mat/9908239). The Ginzburg criterion estimates the energy necessary for hopping over an energy barrier for a real order parameter. For a superconductor, however, the size of directional fluctuations is relevant which gives rise to vortex loop proliferation. In general, fluctuations with symmetry O(N) are important in an $N^2$-times larger temperature interval than Ginzburg’s. See also pp. 18–23 in the textbook [48].


[49] A small selection of papers on this subject is:
J. Tessmann, *Two Loop Renormalization of Scalar Electrodynamics*, MS thesis 1984 (the pdf file is available on the internet at (kl/MS-Tessmann.pdf), where kl is short for www.physik.fu-berlin.de/˜kleinert;

[50] H. Kleinert, *Disorder Version of the Abelian Higgs Model and the Superconductive Phase Transition*, Lett. Nuovo Cimento **35**, 405 (1982). See also the more detailed discussion in Ref. [7], Part 2, Chapter 13 (kl/b1/gifs/v1-716s.html), where the final disorder theory was derived [see, in particular, Eq. (13.30)].

The Monte Carlo simulations of these authors yield the tricritical value (0.76±0.04)/√2 for the Ginzburg parameter κ = √g/q^2.

Scientific, Singapore, 2000; *Evidence for 3D-xy critical properties in under-doped YBa2Cu3O7+x*, (cond-mat/0610289).


[54] There is also a good physical reason for the factor $\sqrt{2}$: In the high-temperature, disordered phase the fluctuations of real and imaginary part of the order field $\psi(x)$ have equal range for $\kappa = 1/\sqrt{2}$.

[55] See Section 4.10 of the textbook [34].


Dynamics of Superfluids

It has been argued by Feynman [1] that at zero temperature, the time dependence of the $\phi$-field in the Hamiltonian (5.6) is governed by the action

$$A = \int dt \int d^3x \mathcal{L} = \int dt \left\{ \int d^3x i\hbar \phi^* \partial_t \phi - H[\phi] \right\},$$

so that the Lagrangian density is

$$\mathcal{L} = i\hbar \phi^* \partial_t \phi - \frac{\hbar^2}{2M} |\nabla \phi|^2 - c_0^2 M (\phi^* \phi - n_0)^2 + \frac{c_0^2 M n_0}{2},$$

where $n_0 = |\phi_0|^2 = -\tau/\lambda$ is the density $\phi^* \phi$ of the superfluid particles in the ground state, i.e., the superfluid density (5.15) which we name $n(x)$ to avoid confusion with the field size $\rho(x) = |\phi(x)|$ in Eq. (5.10)–(5.20). The last term is the negative condensation energy density $-c_0^2 M n_0/2$ in the superfluid phase. The interaction strength $\lambda$ in (5.6) has been reparametrized as $2c_0^2 M n_0/2$ and $\tau$ as $-2c_0^2 M$ for reasons to be understood below.

The equation of motion of the time-dependent field $\phi(t, x) \equiv \phi(x)$ is

$$i\hbar \partial_t \phi(x) = \left[ -\frac{\hbar}{2M} \nabla^2 - c_0^2 M + \frac{c_0^2 M}{n_0} \phi^*(x) \phi(x) \right] \phi(x).$$
6.1 Hydrodynamic Description

After substituting $\phi(x)$ by $\rho(x)e^{i\theta(x)}$ as in Eq. (5.10), and further $\rho(x)$ by $\sqrt{n(x)}$, the Lagrangian density in (6.3) becomes

$$L = n(x)\left\{-\hbar [\partial_t \theta(x) + \theta^\nu(x)] - \frac{\hbar^2}{2M} [\nabla \theta(x) - \theta^\nu(x)]^2 - e_{\nabla n}(x) - e_n(x)\right\}$$

where

$$e_n(x) \equiv \frac{c_0^2 M}{2n_0 n(x)} \left\{ [n(x) - n_0]^2 - n_0^2 \right\}$$

is the energy per particle associated with the density of the fluctuating condensate, and

$$e_{\nabla n}(x) \equiv \frac{\hbar^2 [\nabla n(x)]^2}{8M n^2(x)}$$

the gradient energy of the condensate. This energy may be also be written with

$$e_{\nabla n}(x) = \frac{p^{\text{os2}}(x)}{2M}$$

where

$$p^{\text{os}}(x) \equiv M v^{\text{os}}(x) \equiv \frac{\hbar \nabla n(x)}{2 n(x)}$$

is $i$ times the quantum-mechanical momentum associated with the expansion of the condensate, the so-called osmotic momentum. The vector $v^{\text{os}}(x)$ is the associated osmotic velocity.

If the particles move in an external trap potential $V(x)$, this is simply added to $e(x)$, so that the two last terms in (6.5) are replaced by

$$e_{\text{tot}}(x) = e_{\nabla n}(x) + e_n(x) + V(x).$$

The field $\theta^\nu(x)$ is the time component of the vortex gauge field. Together with $\theta^\nu(x)$ it forms the four-vector

$$\theta^{\mu\nu}(x) = (c\theta^\nu(x), \theta^\nu(x)), \quad (\mu = 0, 1, 2, 3),$$

which is the spacetime extension of the vortex gauge field (5.21). If the jumping surface $S$ in Eq. (5.17) moves along the time axis, it becomes a volume $V$, for which we define a $\delta$-function as follows:

$$\delta_{abc}(x; V) \equiv \int d\sigma d\tau d\lambda \left[ \sum_{P(abc)} \epsilon_p \frac{\partial \vec{x}_b}{\partial \sigma} \frac{\partial \vec{x}_c}{\partial \tau} \frac{\partial \vec{x}_d}{\partial \lambda} \right] \delta^{(4)}(x - \vec{x}(\sigma, \tau, \lambda)).$$

H. Kleinert, GRAVITY WITH TORSION
where the sum runs over all 6 permutations $P$ of the indices and $\epsilon_P$ denotes their parity ($\epsilon_P = +1$ for even and $-1$ for odd permutation $P$). From this we form the dual $\delta$-function

$$\tilde{\delta}_a(x; V) \equiv \epsilon_{abcd} \delta_{abc}(x; V).$$ (6.13)

We may conveniently chose the axial gauge of the vortex gauge field where the time component $\theta^v_t(x)$ vanishes and only the spatial part $\theta^v(x)$ is nonzero. Then the spatial components of $\delta_a(x; V)$ can be written as $\delta(x; S(t))$.

After gauge fixing, the field $\theta(x)$ runs from $-\infty$ to $\infty$ rather than $-\pi$ to $\pi$ [recall the steps leading from the partition function (5.31) to (5.39)].

We now introduce the velocity field with vortices

$$v(x) \equiv \hbar [\nabla \theta(x) - \theta^v(x)]/M,$$ (6.14)

and the local deviation of the particle density from the ground-state value $\delta n(x) \equiv n(x) - n_0$, so that (6.5) can be written as

$$L = -n(x) \left[ \hbar \partial_t \theta(x) + \frac{M}{2} v^2(x) + e_{\text{tot}}(x) \right].$$ (6.15)

The Lagrangian density (6.15) is invariant under changes of $\theta(x)$ by an additive constant $\Lambda$. According to Noether’s theorem, this implies the existence of a conserved current density. We can calculate the charge and particle current densities from the rule (3.101) as

$$n(x) = -\frac{1}{\hbar} \frac{\partial L}{\partial \partial_t \theta(x)}; \quad j(x) = -\frac{1}{\hbar} \frac{\partial L}{\partial \nabla \theta(x)} = n(x)v(x).$$ (6.16)

To find the second expression we must remember (6.14). The prefactor $1/\hbar$ is chosen to have the correct physical dimensions. The associated conservation law reads

$$\partial_t n(x) = -\nabla \cdot [n(x)v(x)],$$ (6.17)

which is the continuity equation of hydrodynamics. This equation is found from the Lagrangian density (6.15) by extremizing the associated action with respect to $\theta(x)$.

Functional extremization with respect to $\delta n(x)$ yields

$$\hbar \partial_t \theta(x) + \frac{M}{2} v^2(x) + V(x) + h_{\nabla n}(x) + h_n(x) = 0,$$ (6.18)

where we have included a possible external potential $V(x)$ as in Eq. (6.10). The last term is the enthalpy per particle associated with the energy density $e_n(x)$. It is defined by

$$h_n(x) \equiv \frac{\partial [n(x)e_n(x)]}{\partial n(x)} = e_n(x) + n(x) \frac{\partial e_n(x)}{\partial n(x)} = e_n(x) + \frac{p_n(x)}{n(x)},$$ (6.19)
where \( p_n(x) \) is the pressure due to the energy \( e_n(x) \):

\[
p_n(x) \equiv n^2(x) \frac{\partial}{\partial n} e_n(x) = \left( n \frac{\partial}{\partial n} - 1 \right) [n(x)e_n(x)].
\]

(6.20)

For \( e_n(x) \) from Eq. (6.6), and allowing for an external potential \( V(x) \) as in (6.10), we find

\[
h_n(x) = \frac{\hbar^2 M}{n_0} \delta n(x), \quad p_n(x) = \frac{\hbar^2 M n_0}{2n_0} n^2(x).
\]

(6.21)

The term \( h_{\nabla n}(x) \) is the so-called quantum enthalpy. It is obtained from the energy density \( e_{\nabla n}(x) \) as a contribution from the Euler-Lagrange equation:

\[
h_{\nabla n}(x) \equiv \frac{\partial [n(x)e_{\nabla n}(x)]}{\partial n(x)} - \nabla \frac{\partial [n(x)e_{\nabla n}(x)]}{\partial \nabla n(x)}.
\]

(6.22)

This can be written as

\[
h_{\nabla n}(x) = e_{\nabla n}(x) + \frac{p_{\nabla n}(x)}{n(x)},
\]

(6.23)

where

\[
p_{\nabla n}(x) = n^2(x) \left[ \frac{\partial}{\partial n} - \nabla \frac{\partial}{\partial \nabla n} \right] e_{\nabla n}(x) = \left\{ n(x) \left[ \frac{\partial}{\partial n} - \nabla \frac{\partial}{\partial \nabla n} \right] - 1 \right\} [n(x)e_{\nabla n}(x)].
\]

(6.24)

is the so-called quantum pressure.

Inserting (6.7) yields

\[
h_{\nabla n}(x) = \frac{\hbar^2}{8M} \left\{ \frac{[\nabla n(x)]^2}{n(x)} - 2\nabla^2 n(x) \right\}, \quad p_{\nabla n}(x) = -\frac{\hbar^2}{4M} \nabla^2 n(x).
\]

(6.25)

The two equations (6.17) and (6.18) were found by Madelung in 1926 [2].

The gradient of (6.18) yields the equation of motion

\[
M \partial_t v(x) + \hbar \partial_t \theta^x + \frac{M}{2} \nabla v^2(x) = -\nabla V_{\text{tot}}(x) - \nabla h_{\nabla n}(x) - \nabla h_n(x),
\]

(6.26)

where

\[
V_{\text{tot}}(x) \equiv V(x) \equiv +h_{\nabla n}(x) + \nabla h_n(x).
\]

(6.27)

We now use the vector identity

\[
\frac{1}{2} \nabla v^2(x) = v(x) \times [\nabla \times v(x)] + [v(x) \cdot \nabla]v(x),
\]

(6.28)

and rewrite Eq. (6.26) as

\[
M \partial_t v(x) + M[v(x) \cdot \nabla]v(x) = -\nabla V_{\text{tot}}(x) + f^v(x),
\]

(6.29)
where

\[ f^v(x) \equiv -\hbar \partial_t \theta^v(x) - M v(x) \times [\nabla \times v(x)] = -\hbar \partial_t \theta^v(x) + \hbar v(x) \times [\nabla \times \theta^v(x)] \] (6.30)

is a force due to the vortices. The classical contribution to the second term is the important Magnus force [3] acting upon a rotating fluid:

\[ f^v_{\text{Magnus}}(x) \equiv -M v(x) \times [\nabla \times v(x)]. \] (6.31)

The important observation is now that this force is in fact zero for the vortices in the superfluid. Let us prove this. Consider first the two-dimensional situation with a point-like vortex which lies at the origin at a given time \( t \). This can be described by a vortex gauge field

\[ \theta^v_1(x) = 0, \quad \theta^v_2(x) = 2\pi \Theta(x_1) \delta(x_2), \] (6.32)

where \( \Theta(x_1) \) is the Heaviside step function which is zero for negative and unity for positive \( x_1 \). The curl of (6.32) is the vortex density, which is proportional to a \( \delta \)-function at the origin:

\[ \nabla \times v(x) = \nabla_1 \theta^v_2(x) - \nabla_2 \theta^v_1(x) = 2\pi \delta^{(2)}(x), \] (6.33)

in agreement with the general relation (5.25). Suppose that the vortex moves, after a short time \( \Delta t \), to the point \( x + \Delta x = (\Delta x_1, 0) \), where

\[ \theta^v_1(x) = 0, \quad \theta^v_2(x) = \Theta(x_1 + \Delta x_1) \delta(x_2), \quad \nabla \times \theta^v(x) = 2\pi \delta^{(2)}(x + \Delta x). \] (6.34)

Since \( \Theta(x_1 + \Delta x_1) = \Theta(x_1) + \Delta x_1 \delta(x_1) \), we see that \( \Delta \theta^v(x) = \Delta x \times [\nabla \times \theta^v(x)] \) which becomes

\[ \partial_t \theta^v(x) = v(x) \times [\nabla \times \theta^v(x)] \] (6.35)

after dividing it by \( \Delta t \) and taking the limit \( \Delta t \to 0 \), thus proving the vanishing of \( f^v(x) \).

The result can easily be generalized to a line with wiggles by approximating it as a sequence of points in closely stacked planes orthogonal to the line elements. As long as the line is smooth, the change in the direction is of higher order in \( \Delta x \) and does not influence the result in the limit \( \Delta t \to 0 \). Thus we can omit the last term in (6.29).

Equation (6.35) is the equation of motion for the vortex gauge field. The time dependence of this field is governed by quantum analog of the Magnus force (6.31).

Note that for a vanishing force \( f^v(x) \) and quantum pressure \( p_{\gamma n}(x) \), Eq. (6.29) coincides with the classical Euler equation of motion for an ideal fluid

\[ M \frac{d}{dt} v(x) = M \partial_t v(x) + M [v(x) \cdot \nabla] v(x) = -\nabla V(x) - \frac{\nabla p_{\gamma n}(x)}{n(x)}. \] (6.36)
The last term is initially equal to \(-\nabla h_n(x)\). However, since \(e_n(x)\) depends only on \(n(x)\) (such systems are referred to as baryotropic), we see that (6.19) implies

\[
\nabla h_n(x) = \nabla e_n(x) - \frac{p_n(x)}{n^*(x)} \nabla n(x) + \frac{\nabla p_n(x)}{n(x)} = \left[ \frac{\partial e_n(x)}{\partial n(x)} - \frac{p_n(x)}{n^*(x)} \right] \nabla n(x) + \frac{\nabla p_n(x)}{n(x)}
\]

(6.37)

There are only two differences between (6.29) with \(\mathbf{f}'(x) = 0\) and the classical equation (6.37). One is the extra quantum part \(-\nabla h_{\Psi_n}(x)\) in (6.29). The other lies in the nature of the vortex structure. In a classical fluid, the vorticity\(^1\)

\[
\mathbf{w}(x) \equiv \nabla \times \mathbf{v}(x)
\]

(6.38)
can be an arbitrary function of \(x\). For instance, a velocity field \(\mathbf{v}(x) = (0, x_1, 0)\) has the constant vorticity \(\nabla \times \mathbf{v}(x) = 1\). In a superfluid, such vorticities do not exist. If one performs the integral over any closed contour \(\oint d\mathbf{x} \cdot \mathbf{v}(x)\), one must always find an integer multiple of \(\hbar\) to ensure the uniqueness of the wave function around the vortex line. This corresponds to the Sommerfeld quantization condition \(\oint d\mathbf{x} \cdot \mathbf{p}(x) = \hbar n\). In a superfluid, there exists no continuous regions of nonzero vorticity, only infinitesimally thin lines. This leaves only vorticities which are superpositions of \(\delta\)-functions \(2\pi \hbar \delta(\mathbf{x}; L)\), which is guaranteed here by the expression (6.14) for the velocity.

By taking the curl of the right-hand side of the vanishing force \(\mathbf{f}'(x)\), we obtain an equation for the time derivative of the vortex density

\[
\partial_t [\nabla \times \mathbf{\theta'}(x)] = \nabla \times \mathbf{v}(x) \times [\nabla \times \mathbf{\theta'}(x)].
\]

(6.39)

Such an equation was first found in 1942 by Ertel [4] for the vorticity \(\mathbf{w}(x)\) of a classical fluid, rather than \(\nabla \times \mathbf{\theta'}(x)\). Using the vector identity

\[
\nabla \times \mathbf{v}(x) \times \mathbf{w}(x) = -\mathbf{w}(x)[\nabla \cdot \mathbf{v}(x)] - [\mathbf{v}(x) \cdot \nabla]\mathbf{w}(x)
\]

\[
+ \mathbf{v}(x)[\nabla \cdot \mathbf{w}(x)] + [\mathbf{w}(x) \cdot \nabla]\mathbf{v}(x),
\]

(6.40)

and the identity \(\nabla \cdot \mathbf{w}(x) = \nabla \cdot [\nabla \times \mathbf{v}(x)] \equiv 0\), we can rewrite the classical version of (6.39) in the form

\[
\frac{d}{dt} \mathbf{w}(x) = \partial_t \mathbf{w}(x) + [\mathbf{v}(x) \cdot \nabla] \mathbf{w}(x) = -\mathbf{w}(x)[\nabla \cdot \mathbf{v}(x)] + [\mathbf{w}(x) \cdot \nabla] \mathbf{v}(x),
\]

(6.41)

which is the form stated by Ertel. This and Eq. (6.36) are the basis for deriving the famous Helmholtz-Thomson theorem of an ideal perfect classical fluid which states that the vorticity is constant along a vortex line if the forces possess a potential.

Equation (6.39) is the quantum version of Ertel’s equation where the vorticity occurs only in infinitesimally thin lines satisfying the quantization condition \(\oint d\mathbf{x} \cdot \mathbf{p}(x) = \hbar n\).

\(^1\)The letter \(\mathbf{w}\) stems from the German word for vorticity=“Wirbelstärke”.

H. Kleinert, GRAVITY WITH TORSION
Inserting the vortex density (5.25) into Eq. (6.39), we obtain for a line \( L(t) \) moving in a fluid with a velocity field \( \mathbf{v}(x) \) the equation

\[
\partial_t \delta(x; L(t)) = \nabla \times \mathbf{v}(x) \times \delta(x; L(t)). \tag{6.42}
\]

It has been argued by L. Morati [5, 6] on the basis of a stochastic approach to quantum theory by E. Nelson [7] that the force \( f^\nu(x) \) is not zero but equal to quantum force

\[
f^\nu(x) \equiv -\frac{\hbar}{2} \left[ \nabla n(x) \right]_n + \nabla \times \left[ \nabla \times \mathbf{v}(x) \right]. \tag{6.43}
\]

Our direct derivation from the superfluid Lagrangian density (6.2) and the equivalent (6.2) does not produce such a term.

### 6.2 Velocity of Second Sound

Consider the Lagrangian density (6.15) and omit the trivial constant condensation energy density \(-c_0^2 M n_0/2\) as well as external potential \( V(x) \). The result is

\[
\mathcal{L} = -[n_0 + \delta n(x)] \left[ \hbar \partial_t \theta(x) + \frac{M}{2} \mathbf{v}^2(x) \right] - \frac{\hbar^2}{8M} \left[ \nabla \delta n(x) \right]^2_\mathbf{n} - \frac{c_0^2 M}{2n_0} [\delta n(x)]^2. \tag{6.44}
\]

For small \( \delta n(x) \ll n_0 \), this is extremal at

\[
\delta n(x) = \frac{n_0}{c_0^2 M} \frac{1}{1 - \xi^2 \nabla^2} \left[ \hbar \partial_t \theta(x) + \frac{M}{2} \mathbf{v}^2(x) \right], \tag{6.45}
\]

where

\[
\xi \equiv \frac{\hbar}{2 c_0^2 M} = \frac{1}{2} c_0^2 \lambda_M \tag{6.46}
\]

is the range of the \( \delta n(x) \)-fluctuations, i.e., the coherence length of the superfluid, and \( \lambda_M = \hbar/Mc \) the Compton wavelength of the particles of mass \( M \).

Reinserting (6.45) into (6.44) leads to the alternative Lagrange density

\[
\mathcal{L} = -n_0 \left[ \hbar \partial_t \theta(x) + \frac{M}{2} \mathbf{v}^2(x) \right] + \frac{n_0}{2 c_0^2 M} \left[ \hbar \partial_t \theta(x) + \frac{M}{2} \mathbf{v}^2(x) \right] \frac{1}{1 - \xi^2 \nabla^2} \left[ \hbar \partial_t \theta(x) + \frac{M}{2} \mathbf{v}^2(x) \right]. \tag{6.47}
\]

The first term is an irrelevant surface term and can be omitted. The quadratic fluctuations of \( \theta(x) \) are governed by the Lagrange density

\[
\mathcal{L}_0 = \frac{n_0 \hbar^2}{2M} \left\{ \frac{1}{c_0^2} [\partial_t \theta(x) - \theta'(x)] - \frac{1}{1 - \xi^2 \nabla^2} [\partial_t \theta(x) - \theta'(x)]^2 \right\}. \tag{6.48}
\]
For the sake of manifest vortex gauge invariance we have reinsered the time-component $\theta^i_v(x)$ of the vortex gauge field which was omitted in (6.48) where we used the axial gauge.

In the absence of vortices, and in the long-wavelength limit, the Lagrangian density (6.48) leads to the equation of motion

$$(-\partial^2_t + c_0^2 \nabla^2)\theta(x) = 0.$$  \hfill (6.49)

This is a Klein-Gordon equation for $\theta(x)$ which shows that the parameter $c_0$ is the propagating velocity of phase fluctuations, which form the second sound in the superfluid.

Note the remarkable fact that although the initial equation of motion (6.4) is nonrelativistic, the sound waves follow a Lorentz-invariant equation in which the sound velocity $c_0$ playing the role of the light velocity. If there is a potential, the velocity of second sound will no longer be a constant but depend on the position.

### 6.3 Vortex-Electromagnetic Fields

It is useful to carry the analogy between the gauge fields of electromagnetism further and define the vortex-electric and vortex-magnetic fields as

$$E^v(x) \equiv -[\nabla \theta^v(x) + \partial_t \theta^v(x)], \quad B^v(x) \equiv \nabla \times \theta^v(x).$$  \hfill (6.50)

These are the analogs of the definitions (2.73) and (2.74) where $\theta^v_t(x)$ corresponds to $A_t(x) \equiv c\phi(x) \equiv A_0(x)$ defined so that $dx^\mu A_\mu \equiv dx^0 A_0 - d\mathbf{x} \cdot \mathbf{A} = dt A_t - d\mathbf{x} \cdot \mathbf{A}$. The $B^v$-field has the dimension cm/sec$^2$, the $E^v$-field 1/sec. They satisfy the same type of Bianchi identities as the electromagnetic fields in (1.182) and (1.183):

$$\nabla \cdot B^v(x) = 0,$$

$$\nabla \times E^v(x) + \partial_t B^v(x) = 0.$$  \hfill (6.51)\hfill (6.52)

With these fields, the vortex force (6.31) becomes

$$f^v(x) = E^v(x) + v(x) \times B^v(x),$$

which has the same form as the electromagnetic force upon a moving particle of unit charge. The corresponding vortex force vanishes.

Recall that the vanishing of this force implies that the time dependence of the vortex gauge field is driven by the Magnus force [see Eq. (6.35)].

By substituting Eq. (6.50) for $B^v(x)$ into (6.39), we find the following equation of motion for the vortex magnetic field:

$$\partial_t B^v(x) = \nabla \times v(x) \times B^v(x).$$  \hfill (6.54)

Note that due to Eq. (6.14),

$$\nabla \times v(x) = -\frac{\hbar}{M} \nabla \times \theta^v(x) = -\frac{\hbar}{M} B^v(x).$$  \hfill (6.55)
6.4 Simple Example

As a simple example illustrating the above extension of Madelung’s theory consider a harmonic oscillator in two dimensions with the Schrödinger equation in cylindrical coordinates \((r, \phi)\) with \(r \in (0, \infty)\) and \(\phi \in (0, 2\pi)\):

\[
\left( -\frac{1}{2} \nabla^2 + \frac{1}{2} r^2 \right) \psi_{nm}(r, \phi) = E_{nm} \psi_{nm}(r, \phi),
\]

where \(n, m\) are the principal quantum number the azimuthal quantum numbers, respectively. For simplicity, we have set \(M = 1\) and \(\hbar = 1\). In particular, we shall focus on the state

\[
\psi_{11}(r, \theta) = \pi^{-1/2} r e^{-r^2/2} e^{i\phi}.
\]

The Hamiltonian of the two-dimensional oscillator corresponding to the field formulation (5.20) reads

\[
H[\phi] = \frac{1}{2} \int d^2 x \phi^* \left( -\nabla^2 + x^2 \right) \phi,
\]

where we have done a nabla integration to replace \(|\nabla \phi|^2\) by \(-\phi^* \nabla^2 \phi\). The wave function (6.57) corresponds to the specific field configuration

\[
\rho(r) = \pi^{-1/2} r, \quad \theta = \arctan(x_2/x_1).
\]

Thus we calculate the energy (6.61) in cylindrical coordinates, where \(\phi^* (-\nabla^2) \phi = -\phi^* (r^{-1} \partial_r r \partial_r - r^{-2} \partial_\phi^2) \phi\) becomes for \(\phi(x) = \psi_{11}(r, \varphi)\):

\[
-\phi^* (-\nabla^2) \phi = \frac{1}{\pi} \left( 4 - r^2 \right) r e^{-r^2},
\]

so that we find

\[
E_{11} = \pi \int_0^\infty dr r \left[ \frac{1}{\pi} (4 - r^2) e^{-r^2} + \frac{1}{\pi} r^4 e^{-r^2} \right] = 1 + 1 = 2.
\]

Let us now calculate the same energy from the hydrodynamic expression for the energy which we read directly off the Lagrangian density (6.5) as

\[
H = \int d^2 x H = \int d^2 x n(x) \left\{ \frac{1}{2} \left[ \nabla \theta(x) - \theta^\ast(x) \right]^2 + \frac{P_{\theta^\ast} x^2}{2} + \frac{x^2}{2} \right\}.
\]

The gradient of \(\theta(x) = \arctan(x_2/x_1)\) has the jump at the cut of \(\arctan(x_2/x_1)\), which runs here from zero to infinity in the \(x_1, x_2\)-plane:

\[
\nabla_1 \arctan(x_2/x_1) = -x_2/r^2, \quad \nabla_2 \arctan(x_2/x_1) = x_1/r^2 + 2\pi \Theta(x_1) \delta(x_2).
\]

The vortex gauge field is the same as in the example (6.32).
When forming the superflow velocity (6.14), the second term in $\nabla^2 \arctan(x_2/x_1)$ is removed by the vortex gauge field (6.32), and we obtain simply

$$v_1(x) = -x_2/r^2, \quad v_2(x) = x_1/r^2. \quad (6.64)$$

Since the wave function has $n(x) = r^2 e^{-r^2}/\pi$, the osmotic momentum (6.9) is

$$p^{\text{os}} = \frac{1}{2} \frac{\nabla n(x)}{n(x)} = \frac{1}{2} \frac{\nabla (r^2 e^{-r^2})}{r^2 e^{-r^2}} \frac{x}{r} = \left( \frac{1}{r} - r \right) \frac{x}{r}. \quad (6.65)$$

Inserting this into (6.62) yields the energy

$$H = \pi \int_0^\infty dr \int_0^\infty dr r^2 e^{-r^2} \left[ \frac{1}{r^2} + \left( \frac{1}{r} - r \right)^2 + r^2 \right], \quad (6.66)$$

which gives the same value 2 as in the calculation (6.61).

Let us check the Ertel equation in the form (6.54). The vortex magnetic field $B^v(x)$ is according to Eq. (6.50) in the present natural units

$$B^v(x) = 2\pi \delta(2)(x). \quad (6.67)$$

This is time-independent, so that the right-hand side of Eq. (6.54) must vanish. Indeed, from (6.64) we see that

$$\mathbf{v}(x) \times B^v(x) = 2\pi \left( \frac{x_1}{r}, \frac{x_2}{r} \right) \delta^2(x), \quad (6.68)$$

so that its curl gives

$$2\pi \nabla \times \left( \frac{x_1}{r}, \frac{x_2}{r} \right) \delta^2(x) = 2\pi \left( \nabla_1 \frac{x_2}{r} - \nabla_2 \frac{x_1}{r} \right) \delta^2(x). \quad (6.69)$$

This vanishes identically due to the rotational symmetry of the $\delta$-function in two dimensions

$$\delta^2(x) = \frac{1}{2\pi r} \delta(r). \quad (6.70)$$

After applying the chain rule of differentiation to (6.69) one obtains zero.

It is interesting to note that the extra quantum force (6.43) happens to vanish as well in this atomic state. In terms of the vortex magnetic field it reads

$$f^{\text{qu}}(x) \equiv \frac{\hbar^2}{2M} \left[ \frac{\nabla n(x)}{n(x)} + \nabla \right] \times B^v(x). \quad (6.71)$$

The $B^v(x)$-field (6.67) has a curl

$$\nabla \times B^v = 2\pi (\nabla_2 \delta^2(x), -\nabla_1 \delta^2(x)). \quad (6.72)$$
If we insert here the rotationally symmetric expression \((6.70)\) for \(\delta^{(2)}(x)\) and rewrite \((6.72)\) and rewrite it as
\[
\nabla \times B^{v} = \left(\frac{x_2}{r}, -\frac{x_1}{r}\right) = -\left(\frac{x_2}{r}, -\frac{x_1}{r}\right) \left[\frac{1}{r^2}\delta(r) - \frac{1}{r}\delta'(r)\right].
\]
(6.73)
The osmotic term adds to this
\[
\nabla n(x) \times B^{v}(x) = 2 \left(\frac{1}{r} - r\right) \nabla \left[\frac{1}{r^2}\delta(r) - \frac{1}{r}\delta'(r)\right],
\]
(6.74)
The two distributions \((6.73)\) and \((6.74)\) are easily shown to cancel each other. Since they both point in the same direction, we remove the unit vectors \((x_2, -x_1)/r\) and compare the two contributions in the force \((6.71)\) which are proportional to
\[
-\left[\frac{1}{r^2}\delta(r) - \frac{1}{r}\delta'(r)\right], \quad 2 \left(\frac{1}{r^2} - 1\right) \delta(r).
\]
(6.75)
Multiplying both expressions by an arbitrary smooth rotation-symmetric test function \(f(r)\) and integrating we obtain
\[
2\pi \int_0^\infty dr \, r \, f(r) \left\{ -\left[\frac{1}{r^2}\delta(r) - \frac{1}{r}\delta'(r)\right], \quad 2\pi \int_0^\infty dr \, r \, f(r) \left[\frac{1}{r^2} - 1\right] \delta(r) \right\}.
\]
(6.76)
These integrals are finite only if \(f(0) = 0\), \(f'(0) = 0\), so that \(f(r)\) must have the small-\(r\) behavior \(f''(0)r^2/2! + f^{(3)}(0)r^3/3! + \ldots\). Inserting this into the two integrals and using the formula \(\int_0^\infty dr \, r^n \delta(r) = -\delta_{n,1}\) for \(n \geq 1\), we obtain the values \(-2\pi f''(0)/2\) and \(2\pi f''(0)/2\), respectively, so that the force \((6.71)\) is indeed equal to zero.

### 6.5 Eckart Theory of Ideal Quantum Fluids

It is instructive to compare the above equations with those for an ideal isentropic quantum fluid without vortices which is described by a Lagrange density due to Eckart [8]:
\[
\mathcal{L} = \frac{1}{2} \nabla^2 n(x) + \Theta(x) M \{ \partial_t n(x) + \nabla \cdot [n(x) \mathbf{v}(x)] \} - n(x)e_{\text{tot}}(x),
\]
(6.77)
where \(e_{\text{tot}}(x)\) is the internal energy \((6.10)\) per particle.

If we extremize the action \((6.77)\) with respect to \(\Theta(x)\), we obtain once more the continuity equation \((6.17)\). Extremizing \((6.77)\) with respect to \(\mathbf{v}(x)\), we see that the velocity field is given by the gradient of the Lagrange multiplier:
\[
\mathbf{v}(x) = \nabla \Theta(x).
\]
(6.78)
Reinserting this into \((6.77)\), the Lagrange density of the fluid becomes
\[
\mathcal{L} = \frac{1}{2} \nabla \Theta(x)^2 - n(x)e(x) + \Theta(x) \{ \partial_t M n(x) + M \nabla \cdot [n(x) \nabla \Theta(x)] \} - n(x)e(x),
\]
(6.79)
or, after a partial integration in the associated action,

\[ \mathcal{L} = -n(x) \left\{ M \partial_t \Theta(x) + \frac{M}{2} \left[ \nabla \Theta(x) \right]^2 + e_{\text{tot}}(x) \right\}. \]  

(6.80)

Since the gradient of a scalar field has no curl, this implies that these actions describe only a vortexless flow.

Comparing (6.78) with (6.14) in the absence of vortices, we identify the velocity potential as

\[ \Theta(x) \equiv \hbar \theta(x)/M. \]  

(6.81)

### 6.6 Rotating Superfluid

If we want to study a superfluid in a vessel which rotates with a constant angular velocity \( \Omega \), we must add to the Lagrangian density (6.3) a source term

\[ \mathcal{L}_n = -x \times j(x) \cdot \Omega = \frac{i}{2} \hbar \phi^*(x) [x \times \vec{\nabla}] \phi(x) \cdot \Omega, \]  

(6.82)

where \( j(x) \) is the current density (2.70), and \( x \times j(x) \) the density of angular momentum of the fluid. After substituting \( \phi(x) \) by \( \rho(x)e^{i\theta(x)} \), this becomes

\[ \mathcal{L}_n = -i\hbar \nabla \rho \cdot \mathbf{v} + Mn(x) \cdot \mathbf{v}_n, \]  

(6.83)

where \( \nabla \rho \) denotes the azimuthal derivative of the density around the direction of the rotation axis \( \Omega \), and

\[ \mathbf{v}_n(x) \equiv \Omega \times x \]  

(6.84)

is the velocity which the particles at \( x \) would have if the fluid would rotate as a whole like a solid. The action associated with the first term vanishes by partial integration since \( n(x) \) is periodic around the axis \( \Omega \). Adding this to the hydrodynamic Lagrangian density (6.5) and performing a quadratic completion in \( \mathbf{v}_n(x) \), we obtain

\[ \mathcal{L} = n(x) \left\{ -\hbar \left[ \partial_t \theta(x) + \bar{\theta}^\prime(x) \right] - \frac{\hbar^2}{2M} \left[ \nabla \theta(x) - \bar{\theta}^\prime(x) - \frac{M}{\hbar} \mathbf{v}_n(x) \right] \right] \right)^2 - e_{\text{tot}}(x) - n(x) V_\mathbf{n}(x) \right\}, \]  

(6.85)

where \( V_\mathbf{n}(x) \) is the harmonic potential

\[ V_\mathbf{n}(x) \equiv -\frac{M}{2} \mathbf{v}_n^2(x) = -\frac{M}{2} \Omega^2 r_\perp^2, \]  

(6.86)

depending quadratically on the distance \( r_\perp \) from the rotation axis.

The velocity \( \mathbf{v}_n(x) \) has a constant curl

\[ \nabla \times \mathbf{v}_n(x) = 2 \Omega. \]  

(6.87)

It can therefore not be absorbed into the wave function by a phase transformation \( \phi(x) \rightarrow e^{i\alpha(x)} \phi(x) \), since this would make the wave function multivalued. The energy
of the rotating superfluid can be minimized only by a triangular lattice of vortex lines. Their total number $N$ is such that the total circulation equals that of a solid body rotation with $\Omega$. Thus, if we integrate along a circle $C$ of radius $R$ around the rotation axis, the number of vortices enclosed is given by

$$M \oint_C dx \cdot v = 2\pi \hbar N. \quad (6.88)$$

In this way the average of the vortex gauge field $\theta'(x)$ cancels the rotation field $v_\Omega(x)$ of constant vorticity.

Triangular vortex lattices have been observed in rotating superfluid $^4$He [11], and recently in Bose-Einstein condensates [12]. The theory of these lattices was developed in the 1960’s by Tkachenko and others for superfluid $^4$He [13], and recently by various authors for Bose-Einstein condensates [14].

Notes and References


See also


[3] The Magnus force is named after the German physicist Heinrich Magnus who described it in 1853. According to J. Gleick, *Isaac Newton*, Harper Fourth Estate, London (2004), Newton observed the effect 180 years earlier when watching tennis players in his Cambridge college. The Magnus force makes airplanes fly due to a circulation of air around the wings. The circulation forms at takeoff, leaving behind an equal opposite circulation at the airport. The latter has caused crashes of small planes starting too close to a jumbo jet. The effect was used by the German engineer Anton Flettner in the 1920’s to drive ships by a rotor rather than a sail. His ship *Baden-Baden* crossed the Atlantic in 1926. Presently, only the French research ship *Alcyone* built in 1985 uses such a drive with two rotors shaped like an airplane wing.


[9] For detailed reviews and references see

[10] For a detailed review and references see


V. Bretin, S. Stock, Y. Seurin, and J. Dalibard, Fast Rotation of a Bose-Einstein Condensate, Phys. Rev. Lett. 92, 050403 (2004);

D. Stauffer and A.L. Fetter, Distribution of Vortices in Rotating Helium II, Phys. Rev. 168, 156 (1968);

7

Dynamics of Charged Superfluid and Superconductor

In the presence of electromagnetism, we simply extend the vortex-covariant derivative in the Lagrangian density (6.3) by a minimally coupled vector potential \(A^\mu(x) = (A_0^\mu(x), \mathbf{A}(x))\), thus forming the fully covariant derivatives

\[
D_t \phi(x) \equiv [\partial_t - \theta^\nu(x) - e\phi(x)]\phi(x).
\]

This amounts to replacing

\[
\theta^\nu_\mu(x) \rightarrow \theta^\nu_\mu(x) + \frac{e}{c} A_\mu(x),
\]

or with \(A_t(x) = cA_0^t(x), \quad \theta^\nu_t(x) = c\theta^\nu_0(x), \quad \theta^\nu_t(x) \rightarrow \theta^\nu_t(x) + \frac{q}{\hbar c} A_t(x), \quad \Theta^\nu(x) \rightarrow \Theta^\nu(x) + \frac{q}{\hbar c} \mathbf{A}(x),
\]

so that we obtain from (6.5):

\[
L = n(x) \left\{ -\hbar \left[ \partial_t \theta(x) + \theta_t(x) + \frac{q}{\hbar c} A_t(x) \right] - \frac{\hbar^2}{2M} \left[ \nabla \theta(x) - \Theta(x) - \frac{q}{\hbar c} \mathbf{A}(x) \right]^2 - e_{\text{tot}}(x) \right\}.
\]

(7.4)

This has to be supplemented by Maxwell’s electromagnetic Lagrangian density (2.84).

Conversely, we may simply take the electromagnetically coupled equation, and replace the vector potential by

\[
A_t(x) \rightarrow \tilde{A}_t(x) \equiv A_t(x) + q_m \theta_t^\nu(x), \quad A(x) \rightarrow \tilde{A}(x) \equiv A(x) + q_m \Theta^\nu(x),
\]

(7.5)

where we have introduced a magnetic charge associated with the electric charge \(q\):

\[
q_m = \frac{\hbar c}{q}.
\]

(7.6)
Then we can rewrite (7.4) in the shorter form

\[ \mathcal{L} = n(x) \left\{ - \hbar \partial_t \theta(x) + \frac{q}{c} \tilde{A}_t(x) - \frac{1}{2M} \left[ \hbar \nabla \theta(x) - \frac{q}{c} \tilde{A}(x) \right]^2 - e_{\text{tot}}(x) \right\}. \]  

(7.7)

The equation of motion of the time-dependent field \( \phi(t, x) \equiv \phi(x) \) is

\[ i \left[ \hbar \partial_t + \frac{q}{c} \tilde{A}_t(x) \right] \phi(x) = \left\{ - \frac{1}{2M} \left[ \hbar \nabla \theta(x) - \frac{q}{c} \tilde{A}(x) \right]^2 - c_0^2 M + \frac{q^2 M}{n_0} \phi^*(x) \phi(x) \right\} \phi(x). \]  

(7.8)

### 7.1 Hydrodynamic Description of Charged Superfluid

For a charged superfluid, the velocity field is given by

\[ v(x) \equiv \frac{1}{M} \left[ \hbar \nabla \theta(x) - \frac{q}{c} \tilde{A}(x) \right] = \frac{\hbar}{M} \left[ \nabla \theta(x) - \theta v(x) - \frac{q}{\hbar c} A(x) \right]. \]  

(7.9)

It is invariant under both magnetic and vortex gauge transformations. In terms of the local deviation of the particle density from the ground-state value \( \delta n(x) \equiv n(x) - n_0 \), the hydrodynamic Lagrangian density (7.7) can be written as

\[ \mathcal{L} = - n(x) \left[ \hbar \partial_t \theta(x) + \hbar \theta v(x) + \frac{q}{c} qA_t(x) + \frac{M}{2} v^2(x) + e_{\text{tot}}(x) \right]. \]  

(7.10)

The electric charge and current densities are simply \( q \) times the particle and current densities (6.16). They can now be derived alternatively from the Noether rule (3.118):

\[ \rho(x)q n(x) = - \frac{1}{c} \frac{\partial \mathcal{L}}{\partial A_t(x)} = q n(x), \quad J(x) = \frac{1}{c} \frac{\partial \mathcal{L}}{\partial A(x)} = q n(x) v(x). \]  

(7.11)

This satisfies the continuity equation

\[ q \partial_t n(x) = - \nabla \cdot J(x) = 0, \]  

(7.12)

which can again be found by extremizing the associated action with respect to \( \theta(x) \).

Functional extremization of the action with respect to \( n(x) \) yields

\[ \hbar \partial_t \theta(x) + \hbar \theta v(x) + qA_0(x) + \frac{M}{2} v^2(x) + h_{\text{tot}}(x) = 0, \]  

(7.13)

where \( \rho^{\text{qu}}(x) \) is the quantum pressure defined in Eq. (6.24). These are the extensions of the Madelung equations (6.17) and (6.18) by vortices and electromagnetism. The last term may, incidentally, be replaced by the enthalpy per particle \( h(x) \) of Eq. (6.19).

The gradient of the second equation yields the equation of motion

\[ M \partial_t v(x) + q \left[ \frac{1}{c} \partial_t A(x) + \nabla A_0(x) \right] + \hbar [ \partial_t \theta v(x) + \nabla \theta v(x)] + \frac{M}{2} \nabla v^2(x) = - \nabla h_{\text{tot}}(x). \]  

(7.14)
Inserting here the identity (6.28) we obtain

\[ M \partial_t \mathbf{v}(x) + M \mathbf{v}(x) \cdot \nabla \mathbf{v}(x) = -\nabla h_{\text{tot}}(x) + \mathbf{f}^{\text{em}}(x) + \mathbf{f}^\nu(x), \quad (7.15) \]

where the sum of the two forces on the right-hand side is

\[ \mathbf{f}^{\text{em}}(x) + \mathbf{f}^\nu(x) = -q \left[ \frac{1}{c} \partial_t \mathbf{A}(x) + \nabla A_0(x) \right] - \hbar [ \partial_t \theta^\nu(x) + \nabla \theta^\nu(x) ] - \hbar \mathbf{v}(x) \times [ \nabla \times \mathbf{v}(x) ] \].

(7.16)

Inserting here the velocity (7.9), further the defining equation (2.73) and (2.79) for the electromagnetic fields, and finally Eqs. (6.50) for the vortex electromagnetic fields, we see that \( \mathbf{f}^{\text{em}}(x) + \mathbf{f}^\nu(x) \) is the sum of the vortex force \( \mathbf{f}^\nu(x) \) of Eq. (6.53), which was shown to vanish in Eq. (6.35), and the electromagnetic Lorentz force acting upon a charged moving particle

\[ \mathbf{f}^{\text{em}}(x) = \mathbf{E}^\nu(x) + \frac{\mathbf{v}(x)}{c^2} \times \mathbf{B}^\nu(x). \quad (7.17) \]

The classical limit of Eq. (7.16) is the well-known equation of motion of magneto-hydrodynamic [5].

The additional Maxwell action adds equations for the electromagnetic field

\[ \nabla \cdot \mathbf{E} = \rho \quad \text{(Coulomb’s law),} \quad (7.18) \]
\[ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c} \mathbf{J} \quad \text{(Ampère’s law),} \quad (7.19) \]
\[ \nabla \cdot \mathbf{B} = 0 \quad \text{(absence of magnetic monopoles),} \quad (7.20) \]
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{(Faraday’s law).} \quad (7.21) \]

\section*{7.2 London Theory of Charged Superfluid}

If we ignore the vortex gauge field in Eq. (7.9), the current density (7.11) is

\[ \mathbf{J}(x) \equiv qn(x)\mathbf{v}(x) = \frac{\hbar qn(x)}{M} \nabla \theta(x) - \frac{q^2 n(x)}{Mc} \mathbf{A}(x). \quad (7.22) \]

The charge \( q \) is equal to \(-2e\) since the charge carriers in the superconductor are Cooper pairs of electrons.

The brothers Heinz and Fritz London [1] considered superconductors with constant density \( n(x) \equiv n_0 \) which is the \textit{London-limit} introduced before in Eq. (5.12). Then they absorbed the phase variable \( \theta(x) \) into the vector potential \( \mathbf{A}(x) \) by a gauge transformation

\[ A'_\mu(x) = A_\mu(x) - \frac{c}{q} \partial_\mu \theta, \quad (7.23) \]

H. Kleinert, GRAVITY WITH TORSION
so that
\[
J(x) \equiv qn_0 \mathbf{v}(x) = -\frac{q^2 n_0}{cM} \mathbf{A}'(x),
\] (7.24)
where \( \mathbf{A}'(x) \) satisfies the gauge condition
\[
\nabla \cdot \mathbf{A}'(x) = 0,
\] (7.25)
to make (7.24) compatible with the current conservation law (7.12).

Taking the time derivative of this and using the defining equation (2.73) for the electric field in terms of the vector potential, they obtained
\[
\partial_t J(x) = \frac{q^2 n_0}{M} [\mathbf{E}(x) + \nabla A'_0(x)].
\] (7.26)

At this place they postulated that the electric potential \( A'_0(x) \) vanishes in a superconductor, which led them to their famous \textit{first London equation}:
\[
\partial_t J(x) = \frac{q^2 n_0}{M} \mathbf{E}(x).
\] (7.27)

In a second step they formed, at a constant \( n(x) = n_0 \), the curl of the current (7.22), and obtained the \textit{second London equation}:
\[
\nabla \times J(x) + \frac{q^2 n_0}{Mc} \mathbf{B}(x) = 0.
\] (7.28)

To check the compatibility of the two London equations one may take the curl of (7.26) and use Faraday’s law of induction (7.21) to find
\[
\partial_t \left[ \nabla \times J(x) + \frac{q^2 n_0}{Mc} \mathbf{B}(x) \right] = 0,
\] (7.29)
in agreement with (7.28).

From the second London equation (7.28) one derives immediately the Meissner effect. First one recalls how the electromagnetic waves are derived from the combination of Ampère’s and Faraday’s laws (1.181) and (1.183), and the magnetic source condition (1.182):
\[
\nabla \times \nabla \times \mathbf{B}(x) + \frac{1}{c^2} \partial_t^2 \mathbf{B}(x) = -\nabla^2 \mathbf{B}(x) + \frac{1}{c^2} \partial_t^2 \mathbf{B}(x) = \frac{1}{c} \nabla \times J(x).
\] (7.30)

In the absence of currents, this equation describes electromagnetic waves propagating with light velocity \( c \). In a superconductor, the right-hand side is replaced by the second London equation (7.28), and leads to
\[
\left[ \frac{1}{c^2} \partial_t^2 \mathbf{B}(x) - \nabla^2 + \lambda_L^{-2} \right] \mathbf{B}(x) = 0,
\] (7.31)
with
\[ \lambda_L = \sqrt{\frac{M c^2}{n_0 q^2}} = \frac{1}{2} \sqrt{\frac{M c^2}{n_0 e^2}} = \frac{1}{2 n_0 \lambda_M^4 \pi \alpha}, \] (7.32)

where \( \alpha \approx 1/137.0359 \ldots \) is the fine-structure constant (1.189), and \( \lambda_M = \hbar/Mc \) the Compton wavelength of the particles of mass \( M \).

Equation (7.31) shows that inside a superconductor, the magnetic field has a finite London penetration depth \( \lambda_L \).

### 7.3 Including Vortices in London Equations

The development in the last section allows us to correct the London equations. First we add the vortex gauge field, so that (7.22) becomes
\[ J(x) \equiv q n(x) \mathbf{v}(x) = \frac{\hbar q n_0(x)}{M} \left[ \nabla \theta(x) - \mathbf{v}^\tau(x) \right] - \frac{q^2 n_0(x)}{c M} \mathbf{A}(x). \] (7.33)

In the London limit where \( n(x) \approx n_0 \), we take again the time derivative of (7.33) and recalling Eq. (6.50), we obtain
\[ \partial_t J(x) = \frac{\hbar q n_0}{M} \left[ \nabla \partial_t \theta(x) + \mathbf{E}^\tau(x) + \nabla \mathbf{v}^\tau(x) \right] + \frac{q^2 n_0}{M} \left[ \mathbf{E}(x) + \nabla A_0(x) \right]. \] (7.34)

As before, we fix the vortex gauge to have \( \theta^v_t(x) = 0 \), and absorb the phase variable \( \theta(x) \) in the vector potential \( \mathbf{A} \) by a gauge transformation (7.23). Thus we remain with the vortex-corrected first London equation
\[ \partial_t J(x) = \frac{q^2 n_0}{M} \left[ \mathbf{E}(x) + q_m \mathbf{E}^\tau(x) + \nabla A_0(x) \right], \] (7.35)

where \( q_m \) is the magnetic charge (7.6) associated with the electric charge \( q \).

Taking at the curl of Eq. (7.33) in the London limit with the same fixing of vortex and electromagnetic gauge, we obtain the vortex-corrected second London equation
\[ \nabla \times J(x) + \frac{q^2 n_0}{Mc} \left[ \mathbf{B}(x) + q_m \mathbf{B}^\tau(x) \right] = 0. \] (7.36)

The compatibility with (7.35) is checked by forming the curl of (7.35) and using the Faraday law of induction (1.183) and its vortex analog (6.52). The result is the statement that the time derivative of (7.36) vanishes, which is certainly true.

Inserting (7.36) into the combined Maxwell equation (7.30) yields the vortex-corrected Eq. (7.31):
\[ \left[ \frac{1}{c^2} \partial_t^2 - \nabla^2 + \lambda_L^{-2} \right] \mathbf{B}(x) = -\lambda_L^{-2} q_m \mathbf{B}^\tau(x). \] (7.37)
7.4 Hydrodynamic Description of Superconductor

From this we can directly deduce the interaction between vortex lines

\[ \mathcal{A}_{\text{int}} = -\frac{\hbar}{2} \int d^4 x d^4 x' B^y(x) G^R_{\lambda L}(x-x') B^y(x), \quad (7.38) \]

where \( G^R_{\lambda L}(x-x') \) is the retarded Yukawa Green function

\[ G^R_{\lambda L}(x-x') = \frac{1}{-c^2 \partial_t^2 + \nabla^2 + \lambda_L^{-2}} \delta(t-t' - R/c), \quad (7.39) \]

in which \( R \) denotes the spatial distance \( R \equiv |x - x'| \).

In the limit \( \lambda_L \to \infty \) this goes over to the Coulomb version which is the origin of the well-known Liénard-Wiechert potential of electrodynamics.

For slowly moving vortices, the retardation can be neglected and, after inserting \( q_m \) from (7.6) and \( B^y(x) \) from (6.50), and performing the time derivatives in (7.40), we find

\[ \mathcal{A}_{\text{int}} = -\frac{\hbar^2 c^2}{2q_m^2} \int dt \int d^3 x j^y(x, t) \frac{1}{-\nabla^2 + \lambda_L^{-2}} j^y(x, t). \quad (7.40) \]

This agrees with the previous static interaction energy in the partition function (5.269), if we go to natural units \( \hbar = c = M = 1 \).

7.4 Hydrodynamic Description of Superconductor

For a superconductor, the above theory of a charged superfluid is not applicable since the initial Ginzburg-Landau Lagrangian density can be derived [3] only near the phase transition where it has, moreover, a purely damped temporal behavior. Hence there is no time derivative term as in Eq. (6.2). At zero temperature, however, the superflow can be described by simple hydrodynamic equations. From the BCS theory, one can derive [2] a Lagrangian of the type (6.48) in the harmonic approximation

\[ \mathcal{L}_0 = -n_0 \hbar \partial_t \theta(x) + \frac{n_0 \hbar^2}{2M} \left\{ \frac{1}{c_0^2} |\partial_t \theta(x) + \theta^v_t(x)|^2 - |\nabla \theta(x) - \theta^v(x)|^2 \right\} + \ldots. \quad (7.41) \]

with the second sound velocity [3]

\[ c_0 = \frac{v_F}{\sqrt{3}}, \quad (7.42) \]

where \( v_F = p_F/M = \sqrt{2M E_F} \) is the velocity of electrons on the surface of the Fermi sphere, which is calculated from the density of electrons (which is twice as big as the density of Cooper pairs \( n_0 \)):

\[ n_{el} = 2 \int \frac{d^3 p}{(2\pi \hbar)^3} = \frac{p_F^3}{3\hbar^2 \pi^2} = \frac{v_F^3}{3\hbar^2 M^4 \pi^2}. \quad (7.43) \]
The dots in (7.41) indicate terms which can be ignored in the long-wavelength limit. These are different from those of the Bose Lagrangian density in (6.48). There we see that the energy spectrum of second sound excitations has a first correction term of the form
\[ \epsilon(k) = c_0|k|(1 - \gamma k^2 + \ldots), \]
with a negative \( \gamma = -\xi < 0 \). In a superconductor at \( T = 0 \), on the other hand, the BCS theory yields a positive \( \gamma \):
\[ \gamma = \frac{\hbar^2}{45 \Delta^2} \frac{4}{l^2}, \quad l \equiv \hbar \sqrt{1/M \Delta}, \]
where \( \Delta \) is the energy gap of the quasiparticle excitations of the superconductor, which is of the order of the transition temperature (times \( k_B \)) (see Appendix 7A). The length scale \( l \) is of the order of the zero-temperature coherence length.

The positivity of \( \gamma \) ensures the stability of the long-wavelength excitations against decay since it makes \( |k_1 + k_2|[1 - \gamma (k_1 + k_2)^2] < |k_1|(1 - \gamma k_1^2) + |k_2|(1 - \gamma k_2^2) \).

We now add the electromagnetic fields by minimal coupling [recall (7.2) and (7.3)] and find
\[ \mathcal{L}_0 = -n_0 \hbar \left[ \partial_t \theta(x) + \frac{\mathbf{q}}{\hbar c} \mathbf{A}_t(x) \right] + \frac{n_0 \hbar^2}{2M} \left\{ \frac{1}{c_0^2} \left[ \partial_t \theta(x) + \frac{\mathbf{q}}{\hbar c} \mathbf{A}_t(x) \right]^2 - \frac{1}{\hbar^2 c^2 M} \nabla \theta(x) - \frac{\mathbf{q}}{\hbar c} \mathbf{A}(x)^2 \right\}, \]

(7.46)
to be supplemented by the Maxwell Lagrangian density (2.84).

The derivative of \( \mathcal{L}_0 \) with respect to \( -\mathbf{A}(x)/c \) yields the current density [recall (3.118)]:
\[ \mathbf{J}(x) = qn_0 \mathbf{v}(x) = \frac{qn_0 \hbar}{M} \left[ \nabla \theta(x) - \mathbf{v}(x) \right] - \frac{q^2 n_0}{cM} \mathbf{A}(x). \]

(7.47)
From the derivative of \( \mathcal{L}_0 \) with respect to \( -\mathbf{A}_t(x)/c \) we obtain the charge density:
\[ q[n(x) - n_0] = \frac{qn_0 \hbar}{M} \frac{1}{c_0^2} \left[ \partial_t \theta(x) + \mathbf{v}(x) \right] - \frac{q^2 n_0}{c_0^2 cM} \mathbf{A}_t(x). \]

(7.48)

If we absorb the field \( \theta(x) \) in the vector potential, we find the same supercurrent as in (7.22):
\[ \mathbf{J}(x) = -\frac{q^2 n_0}{cM} \mathbf{A}(x) \]

(7.49)
whereas the charge density becomes
\[ qn(x) = -\frac{q^2 n_0}{c_0^2 cM} \mathbf{A}_t(x). \]

(7.50)
The current conservation law implies that
\[ \nabla \cdot \mathbf{A}(x) + \frac{c^2}{c_0^2 c^2} \partial_t \mathbf{A}_t = 0. \]

(7.51)
Note the difference by the large factor $c^2/c_0^2$ of the time derivative term with respect to the Lorentz gauge (2.105):

$$\partial_a A^a(x) = \nabla \cdot A(x) + \partial_0 A^0(x) = \nabla \cdot A(x) + \frac{1}{c^2} \partial_t A^t(x) = 0. \quad (7.52)$$

Since the velocity $c_0 = v_F/\sqrt{3}$ is much smaller than the light velocity $c$, typically by a factor $1/100$, the ratio $c^2/c_0^2$ is of the order of $10^4$.

At this point, we recall that according to the definition (5.24), the ratio of the penetration depth $\lambda_L$ of Eq. (7.32) and the coherence length $\xi$ of Eq. (6.46) define the Ginzburg parameter $\kappa \equiv \lambda_L/\sqrt{2} \xi$. This allows us to express the ratio $c_0/c$ in terms of $\kappa$ as follows:

$$\frac{c_0}{c} = \frac{\kappa}{\sqrt{2} \sqrt{n_0 \lambda^2_M q^2}}. \quad (7.53)$$

If the current density is inserted into the combined Maxwell equation (7.30) to obtain Eq. (7.31) for the screened magnetic field $B$ and its vortex-corrected version (7.31), the field equation for $A_0$, however, has quite different wave propagation properties, due to the factor $c^2/c_0^2$. It is obtained by varying the action $A = \int dtd^3x [L^\text{em}] + L_0$, with respect to $-A_0(x) = -A_t(x)/c$, which yields

$$\nabla \cdot E(x) = qn(x). \quad (7.54)$$

Inserting $E(x)$ from (2.73), and $qn(x)$ from Eq. (7.50) in the axial vortex gauge, we find

$$-\nabla^2 A^0(x) - \frac{1}{c} \partial_t \nabla \cdot A(x) = -\frac{q^2 n_0}{c_0^2 M} A^0(x). \quad (7.55)$$

Eliminating $\nabla \cdot A(x)$ with the help of Eq (7.51), we obtain

$$\left(-\nabla^2 + \lambda_{L,0}^{-2}\right) A^0(x) - \frac{c^2}{c_0^2 c^2} \partial_t^2 A^0 = 0. \quad (7.56)$$

This equation shows that the field $A^0(x)$ penetrates a superconductor over the distance

$$\lambda_{L,0} = \frac{c_0}{c} \lambda_L = \frac{c_0}{c} \frac{1}{\sqrt{n_0 \lambda^2_M q^2}}. \quad (7.57)$$

which is typically two orders of magnitude smaller than the penetration depth $\lambda_L$ of the magnetic field. Moreover, the propagation velocity of $A^0(x)$ is not the light velocity $c$ but the much smaller velocity $c_0 = v_F/\sqrt{3}$.

Note that Eq. (7.56) for $A^0(x)$ has no gauge freedom left, the gauge being fixed by Eq. (7.51). This is best seen by expressing $A^0(x)$ in terms of the charge density $qn(x)$ via Eq. (7.50) which yields

$$\left[-\frac{1}{c^2} \partial_t^2 + \frac{c_0^2}{c^2} \left(-\nabla^2 + \lambda_{L,0}^{-2}\right)\right] n(x) = 0. \quad (7.58)$$
Using the vanishing of $f^v(x)$ of Eq. (6.53), and Eq. (6.55), we can rewrite the divergence of $E^v(x)$ as

$$\nabla \cdot E^v(x) = -\nabla \cdot [v(x) \times B^v(x)] = -[\nabla \times v(x)]B^v(x) + v(x) \cdot [\nabla \times B^v(x)]$$

$$= \frac{\hbar}{M}[B^v(x)]^2 + v(x) \cdot [\nabla \times B^v(x)].$$

(7.59)

The Lagrangian density of the vector field $A^\mu(x)$ can also be written as

$$L = \frac{1}{2} A^a(x) \left( \partial^2 g_{ab} - \partial_a \partial_b \right) A^b(x) + \frac{m^2_A}{2} \{ A^0(x) ^2 - A^2(x) \}$$

(7.60)

where $m^2_A = \lambda - \frac{2}{L}$ and $a = \frac{c^2}{c^2_0}$. The field equation in the presence of an external source $j^a(x)$ is

$$\left[ \left( \partial^2 g_{ab} - \partial_a \partial_b \right) + m^2_A g_{ab} + m^2_A (a - 1) h_{ab} \right] A^b(x) = j_a(x),$$

(7.61)

where $h_{ab}$ has only one nonzero matrix element $h_{00} = 1$. By contracting (7.61) with $\partial_a$, and using current conservation $\partial_a j^a(x) = 0$, we obtain

$$\partial_a A^a(x) + (a - 1) \partial_a A^0(x) = 0,$$

(7.62)

which is the divergence equation (7.51). Reinserting this into (7.61) yields

$$(a \partial_0^2 - \nabla^2 + am_A^2) A^0(x) = j^0(x),$$

(7.63)

$$(\partial^2 + m_A^2) A(x) + \frac{a - 1}{a} \nabla \cdot A(x) = j(x).$$

(7.64)

To check the consistency, we take the divergence of the second equation and use the current conservation law to write

$$-(\partial^2 + m_A^2) [\nabla \cdot A(x)] + \frac{a - 1}{a} \nabla^2 [\nabla \cdot A(x)] = \partial_0 j^0(x).$$

(7.65)

Then we replace $\nabla \cdot A(x)$ by $-a \partial_0 A^0(x)$ and obtain

$$a \partial_0 (\partial_0^2 - \nabla^2 + m_A^2) A^0(x) + (a - 1) \partial_0 \nabla^2 A^0(x) = \partial_0 j^0(x),$$

(7.66)

which is the time derivative of Eq. (7.63).

The inverse of Eq. (7.63) is

$$A^0(x) = \frac{1}{a} \left( \partial_0^2 - a^{-1} \nabla^2 + m_A^2 \right)^{-1} j^0(x).$$

(7.67)

To invert Eq. (7.64) we rewrite it in terms of spatial transverse and longitudinal projection matrices

$$P^t_{ij} = \delta_{ij} - \frac{\nabla_i \nabla_j}{\nabla^2}, \quad P^l_{ij} = \frac{\nabla_i \nabla_j}{\nabla^2},$$

(7.68)
as
\[
\left[ (\partial^2 + m_A^2)P_t + (\partial^0_0 - a^{-1}\nabla^2 + m_A^2)P^0 \right] A(x) = j(x).
\] (7.69)

This is immediately inverted to
\[
A(x) = \left[ (\partial^2 + m_A^2)^{-1}P_t + (\partial^0_0 - a^{-1}\nabla^2 + m_A^2)^{-1}P^0 \right] j(x).
\] (7.70)

From these equations we derive the interaction between external currents
\[
A^{\text{int}} = \int d^4x \left[ \frac{1}{a}j^0(x)\partial^2 - a^{-1}\nabla^2 + m_A^2 j^0(x) - j^t(x)\partial^2 - a^{-1}\nabla^2 + m_A^2 j^t(x) \right].
\] (7.71)

Only the transverse currents interact with the relativistic retarded interaction. Due to current conservation the first two terms can be combined and we obtain
\[
A^{\text{int}} = \int d^4x \left[ j^0(x)\frac{a^{-1} - \nabla^2 a^{-1}}{\partial^2 - a^{-1}\nabla^2 + m_A^2} j^0(x) - j^t(x)\frac{1}{\partial^2 - \nabla^2 + m_A^2} j^t(x) \right].
\] (7.72)

For \(a = 1\), this reduces to the usual relativistic interaction
\[
A^{\text{int}} = \int d^4x j_\mu(x) \frac{1}{\partial^2 + m_A^2} j^\mu(x).
\] (7.73)

## Appendix 7A  Excitation Spectrum of Superconductor

In BCS theory, the electrons are bound to Cooper pairs, and the quasiparticle energies have the form
\[
E(p) = \sqrt{\xi^2(p) + \Delta^2},
\] (7A.1)

where
\[
\xi(p) \equiv \frac{p^2}{2M} - \mu
\] (7A.2)

are the free-electron energies measured from the chemical potential \(\mu\). At zero temperature, this is equal to the Fermi energy \(\epsilon_F = Mv_F^2/2\).

### 7A.1  Gap Equation

The quasiparticle energies have a gap \(\Delta\) which is determined by the gap equation
\[
\frac{1}{g} = \frac{T}{V} \sum_{\omega_m, p} \frac{1}{\omega_m^2 + E^2(p)} = \frac{1}{V} \sum_{p} \frac{1}{2E(p)} \tanh \frac{E(p)}{2T},
\] (7A.3)

where \(g\) is the attractive short-range interaction between electrons near the surface of the Fermi sea caused by the electron-phonon interaction. The sum over \(\omega_m\) runs
over the Matsubara frequencies \( \omega_m = 2\pi k_B T m \), for \( m = 0, \pm 1, \pm 2, \ldots \). The equality of the second and third expression in (7A.3) follows from the summation formula

\[
T \sum_{\omega_m} \frac{e^{i\omega_m \eta}}{i\omega_m - E} = n(E),
\]

(7A.4)

where \( 0 < \eta \ll 1 \) is an infinitesimal parameter to make the sum convergent, and \( n(E) \) is the Fermi distribution function

\[
n(E) = \frac{1}{e^{E/T} + 1} = \frac{1}{2} \left( 1 - \tanh \frac{E}{2T} \right).
\]

(7A.5)

By combining the sums (7A.4) with \( E \) and \(-E\) we obtain the important formula

\[
T \sum_{\omega_m} \frac{1}{\omega_m^2 + E^2} = \frac{1}{2E} T \sum_{\omega_m} \left( \frac{1}{i\omega_m + E} - \frac{1}{i\omega_m - E} \right) = \frac{1}{2E} \left[ n(-E) - n(E) \right]
\]

(7A.6)

In terms of \( n(E) \) of (7A.5), the gap equation (7A.3) reads

\[
\frac{1}{g} = \frac{1}{V} \sum_p \frac{1}{2E(p)} [1 - 2n(E)].
\]

(7A.7)

The momentum sum in (7A.3) are conveniently performed in an approximation which is excellent for small electron-phonon interaction, where only the neighborhood of the Fermi surface contributes significantly, as follows:

\[
\frac{1}{V} \sum_p \to \int \frac{d^3 p}{(2\pi)^3} = \int dp \int \frac{d^2 p}{2\pi^2} \approx M^2 v_F^2 \int d\xi \int d\zeta.
\]

(7A.8)

The prefactor

\[
\mathcal{N}(0) \equiv \frac{m^2 v_F^2}{2\pi^2}
\]

(7A.9)

is the density of states of one spin orientation on the surface of the Fermi sea. This brings the gap equation (7A.3) to the form

\[
\frac{1}{g} = \mathcal{N}(0) \int_0^{\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2}} \tanh \frac{\sqrt{\xi^2 + \Delta^2}}{2T}.
\]

(7A.10)

The integral is logarithmically divergent. Since the attraction is due to phonons in the crystal whose spectrum is limited by the Debye frequency \( \omega_D \) which in conventional superconductors is much smaller than the Fermi energy \( \epsilon_F \), we have cut off the integral at \( \omega_D \).

The critical temperature \( T_c \) is the place where the gap \( \Delta \) vanishes and (7A.19) reduces to

\[
\frac{1}{g} = \mathcal{N}(0) \int_0^{\omega_D} \frac{d\xi}{\xi} \tanh \frac{\xi}{2T_c}.
\]

(7A.11)
Appendix 7A

Excitation Spectrum of Superconductor

219

The integral is done as follows. It is performed first by parts to yield
\[
\int_0^{\omega_D} \frac{d\xi}{\xi} \tanh \frac{\xi}{2T_c} = \log \frac{\xi}{T_c} \tanh \frac{\xi}{2T_c} \bigg|_0^{\omega_D} - \frac{1}{2} \int_0^{\omega_D} \frac{\xi}{T_c} \log \frac{\xi}{T_c} \cosh^2 \frac{\xi}{2T_c}.
\]

(7A.12)

Since \(\omega_D/\pi T_c \gg 1\), the first term is equal to \(\log(\omega_D/2T_c)\), with exponentially small corrections from the hyperbolic tangens which can be ignored. In the second integral, we have taken the upper limit of integration to infinity since it converges. We may use the integral formula
\[
\int_0^{\infty} dx \frac{x^{\mu-1}}{\cosh^2(ax)} = \frac{4}{(2a)^{2\mu}} \left(1 - 2^{2-\mu}\right) \Gamma(\mu)\zeta(\mu - 1),
\]

(7A.13)

set \(\mu = 1 + \delta\), expand the formula to order \(\delta\), and insert the special values
\[
\Gamma'(1) = -\gamma, \quad \zeta'(0) = -\frac{1}{2} \log(2\pi) \log(4e^\gamma/\pi),
\]

(7A.14)

where \(\gamma\) is Euler’s constant
\[
\gamma = -\Gamma'(1)/\Gamma(1) \approx 0.577,
\]

(7A.15)

so that \(e^\gamma/\pi \approx 1.13\). Thus we find from the linear terms in \(\delta\):
\[
\int_0^{\infty} dx \frac{\log x}{\cosh^2(x/2)} = -2 \log(2e^\gamma/\pi),
\]

(7A.16)

and we Eq. (7A.11) becomes
\[
\frac{1}{g} = \mathcal{N}(0) \log \left(\frac{\omega_D 2e^\gamma}{\pi} \right),
\]

(7A.17)

which determines \(T_c\) in terms of the coupling strength \(g\) as
\[
T_c = \omega_D \frac{2e^\gamma}{\pi} e^{-1/g\mathcal{N}(0)}.
\]

(7A.18)

In order to find the \(T\)-dependence of the gap, we may expand the hyperbolic tangens in Eq. (7A.19) in powers of \(e^{-E(p)/T}\) and obtain
\[
\frac{1}{g} = \mathcal{N}(0) \int_0^{\omega_D} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2}} \left[1 + 2 \sum_{n=1}^{\infty} (-1)^n \exp(-n\sqrt{\xi^2 + \Delta^2}/T) - \right],
\]

(7A.19)

where \(K_0(z)\) are the modified Bessel functions of the second kind. The cutoff is needed only in the first integral, in the others it can be moved to infinity, and we obtain
\[
\frac{1}{g} = \mathcal{N}(0) \left[ \log \frac{2\omega_D}{\Delta} + 2 \sum_{n=1}^{\infty} (-1)^n K_0(n\Delta/T) \right].
\]

(7A.20)

Replacing $1/g$ by (7A.18), we find
\[
\log \left( \frac{\Delta e^\gamma}{T_c \pi} \right) = 2 \sum_{n=1}^{\infty} (-1)^n K_0(n\Delta/T).
\] (7A.21)

For small $T$, $K_0$ vanishes exponentially fast:
\[
2K_0 \left( \frac{\Delta}{T} \right) \to \frac{1}{\Delta} \sqrt{2\pi T \Delta e^{-\Delta/T}}.
\] (7A.22)

Hence we find the $T=0$ -gap
\[
\Delta(0) = 2\omega_D e^{-1/gN(0)}.
\] (7A.23)

Combining this with (7A.18) we obtain the important universal relation between critical temperature $T_c$ and energy gap at zero-temperature $\Delta(0)$:
\[
\frac{\Delta(0)}{T_c} = \pi e^{-\gamma} \approx 1.76.
\] (7A.24)

This value is approached exponentially as $T \to 0$ since from (7A.19)
\[
\log \frac{\Delta(T)}{\Delta(0)} \approx \frac{\Delta(T)}{\Delta(0)} - 1 \approx -\frac{1}{\Delta(0)} \sqrt{2\pi T \Delta(0) e^{-\Delta(0)/T}}.
\] (7A.25)

With $\Delta(0)$ determined by (7A.24), we may replace the left-hand side of (7A.21) by $\log[\Delta/\Delta(0)]$.

For $T \approx T_c$, the gap is calculated most efficiently by combining the gap equation (7A.19) with its $T = T_c$ -version to obtain
\[
\int_0^{\omega_D} d\xi \left( \tanh \frac{\xi}{2T_c} - \tanh \frac{\xi}{2T} \right) = \int_0^{\omega_D} d\xi \left( \frac{1}{\sqrt{(2\xi^2 + \Delta^2)/2T}} - \frac{1}{\xi} \tanh \frac{\xi}{2T} \right).
\] (7A.26)

The integrals on both sides are now convergent so that the cutoff frequency $\omega_D$ can be removed. If the integrals on the left-hand side are performed individually as in Eq. (7A.11)–(7A.17) they yield
\[
\log \left( \frac{\omega_D}{T_c} \frac{2e^\gamma}{\pi} \right) - \log \left( \frac{\omega_D}{T} \frac{2e^\gamma}{\pi} \right) = \log \frac{T}{T_c}.
\] (7A.27)

On the right-hand side we replace each hyperbolic tangens by a Matsubara sum according to Eq. (7A.6), and arrive at
\[
\log \frac{T}{T_c} = T \sum_{\omega_m} \int_0^{\infty} d\xi \left( \frac{1}{\omega_m^2 + \xi^2 + \Delta^2} - \frac{1}{\omega_m^2 + \xi^2} \right).
\] (7A.28)
This can be integrated over $\xi$ to yield the gap equation

$$\log \frac{T}{T_c} = 2\pi T \sum_{\omega_m>0} \left( \frac{1}{\sqrt{\omega_m^2 + \Delta^2}} - \frac{1}{\omega_m} \right).$$  \hfill (7A.29)

It is convenient to introduce the reduced gap

$$\delta \equiv \frac{\Delta}{T}$$  \hfill (7A.30)

and a reduced version of the Matsubara frequencies:

$$x_n \equiv (2n + 1)\pi/\delta.$$  \hfill (7A.31)

Then the gap equation (7A.29) takes the form

$$\log \frac{T}{T_c} = \frac{2\pi}{\delta} \sum_{n=0}^{\infty} \left( \frac{1}{x_n^2 + 1} - \frac{1}{x_n} \right).$$  \hfill (7A.32)

It can be used to calculate $T/T_c$ as a function of $\delta$, from which we obtain

$$\frac{\Delta(T)}{\Delta(0)} = (e^{\gamma}/\pi) \frac{\Delta(T)}{T_c} = (e^{\gamma} \delta/\pi) T/T_c$$  \hfill (7A.32)

as a function of $T/T_c$ as shown in Fig. 7.1.

The behavior in the vicinity of the critical temperature $T_c$ can be extracted from Eq. (7A.32) by expanding the sum under the assumption of small $\delta$ and large $x_n$. The leading term gives

$$\log \frac{T}{T_c} \approx \frac{2\pi}{\delta} \sum_{n=0}^{\infty} \frac{1}{2x_n^3} = -\frac{\delta^2}{\pi} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = -\frac{\delta^2}{\pi^2} \frac{7}{8} \zeta(3)$$  \hfill (7A.33)

so that

$$\delta^2 \approx \frac{8\pi^2}{7\zeta(3)} \left(1 - \frac{T}{T_c} \right)$$  \hfill (7A.34)

and

$$\frac{\Delta}{T_c} = \delta_c = \pi \sqrt{\frac{8}{7\zeta(3)}} \left(1 - \frac{T}{T_c} \right)^{1/2} \approx 3.063 \times \left(1 - \frac{T}{T_c} \right)^{1/2}.$$  \hfill (7A.35)
7A.2 Action of Quadratic Fluctuations

The small fluctuations $\delta \Delta(x)$ of the complex field of Cooper pairs are governed by the quadratic action

$$A_2[\delta \Delta^*, \delta \Delta] = -\frac{i}{2} \text{Tr} \left[ \mathbf{G}_\Delta \left( \begin{array}{cc} 0 & \delta \Delta \\ \delta \Delta^* & 0 \end{array} \right) \right] - \frac{1}{g} \int dx |\delta \Delta(x)|^2,$$

(7A.36)

where $\delta \Delta(x)$ is a small fluctuation of the complex gap field around the real background value $\Delta(T)$ given by the gap equation (7A.3). The gap equation is determined from the extremum of the action which ensures that the fluctuation expansion $\delta \Delta(x)$ has no linear term and is dominated by (7A.36). The matrix $\mathbf{G}_\Delta(x', x)$ denotes the free correlation functions of the electrons in a constant background pair field $\Delta(x) = \Delta$:

$$\mathbf{G}_\Delta(x, x') = i \left( [i \partial_t - \xi (-i \nabla)] \delta - \Delta \mp i \partial_t - \xi (i \nabla)] \delta \right)^{-1} (x, x').$$

(7A.37)

At finite temperature we go to Fourier space with Matsubara frequencies $\nu_n = 2\pi(m + \frac{1}{2})T$ for the electrons and $\nu_n = 2\pi n T$ for the pair field which guarantee the antiperiodicity of the Fermi and the periodicity of the Cooper pair field in the imaginary time interval $\tau \in (0, 1/T)$. If we use a four-vector notation for the Euclidean electron momenta $p \equiv (\omega_m, \mathbf{p})$ and pair momenta $k = (\nu_n, \mathbf{k})$, the action (7A.36) reads

$$A_2[\delta \Delta^*, \delta \Delta] = \frac{1}{2V} \sum_{p, k} \left\{ \left[ (\omega_m + \frac{\nu_n}{2})^2 + E_1^2 (\mathbf{p} + \mathbf{k}/2) \right] \left[ (\omega_m - \frac{\nu_n}{2})^2 + E_1^2 (\mathbf{p} - \mathbf{k}/2) \right]^{-1} \right\} \times \left\{ \left[ \omega_m^2 - \frac{\nu_n^2}{4} + \xi (\mathbf{p} + \mathbf{k}/2) \xi (\mathbf{p} - \mathbf{k}/2) \right] \left[ \Delta^*(k) \delta \Delta(k) + \delta \Delta(-k) \delta \Delta^*(-k) \right] \right. \left. - |\Delta_0|^2 \left[ \delta \Delta^*(k) \delta \Delta^*(-k) + \delta \Delta(k) \delta \Delta(-k) \right] \right\} - \frac{1}{g} \sum_k \delta \Delta^*(k) \delta \Delta(k).$$

(7A.38)

This has the generic quadratic form

$$A_2[\delta \Delta^*, \delta \Delta] = \frac{1}{2V} \sum_k \left[ \delta \Delta^*(k) L_{11}(k) \delta \Delta(k) + \delta \Delta(-k) L_{22}(k) \delta \Delta(-k) \right] \left[ \delta \Delta^*(k) L_{12}(k) \delta \Delta^*(-k) + \delta \Delta(-k) L_{21}(k) \delta \Delta(k) \right],$$

(7A.39)

with coefficients

$$L_{11}(k) = L_{22}(k) = \int \frac{d^3p}{(2\pi)^3} \sum_{\omega_m} \frac{\omega_m^2 - \nu_n^2/4 + \xi_+ \xi_-}{\left[ (\omega_m + \frac{\nu_n}{2})^2 + E_1^2 \right] \left[ (\omega_m - \frac{\nu_n}{2})^2 + E_1^2 \right]},$$

(7A.40)

$$L_{12}(k) = \left[ L_{22}(k) \right]^* = -\Delta^2 \int \frac{d^3p}{(2\pi)^3} \sum_{\omega_m} \frac{1}{\left[ (\omega_m + \frac{\nu_n}{2})^2 + E_1^2 \right] \left[ (\omega_m - \frac{\nu_n}{2})^2 + E_1^2 \right]}.$$
Appendix 7A  Excitation Spectrum of Superconductor 223

After some straightforward algebra and using the sum formula (7A.4), we replace the Euclidean pair energy \( \nu \) by \(-i\) times a real continuous energy \( \epsilon \), and obtain

\[
L_{11}(\epsilon, k) = \int \frac{d^3p}{(2\pi)^3} \left\{ \frac{E_+ E_- + \xi_+ \xi_-}{2E_+ E_-} \frac{E_+ + E_-}{(E_+ + E_-)^2 - \epsilon^2} \left[ 1 - n(E_+) - n(E_-) \right] \right\}
\]

(7A.42)

\[
- \frac{E_+ E_- - \xi_+ \xi_-}{2E_+ E_-} \frac{E_+ - E_-}{(E_+ - E_-)^2 - \epsilon^2} \left[ n(E_+) - n(E_-) \right]
\]

(7A.43)

\[
L_{12}(\epsilon, k) = -\Delta^2 \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_+ E_-}
\]

\[
\times \left\{ \frac{E_+ - E_-}{(E_+ - E_-)^2 - \epsilon^2} \left[ 1 - n(E_+) - n(E_-) \right] + \frac{E_+ - E_-}{(E_+ + E_-)^2 - \epsilon^2} \left[ n(E_+) - n(E_-) \right] \right\}.
\]

In the last term of (7A.42) we have used the gap equation in the form (7A.7) to eliminate \( 1/g \). The symbols \( \xi_\pm \) and \( E_\pm \) are defined as follows. In terms of these quantities we define

\[
\xi_\pm = \xi(p \pm k/2) = \frac{p^2}{2m} \pm \frac{1}{2} \nu k + \frac{k^2}{8m} - \mu, \quad E_\pm = E(p \pm k/2),
\]

(7A.44)

where \( m \) is the electron mass. Recalling (7A.1) and (7A.2), the energies in the above integrals read explicitly, with \( \xi \equiv \xi(p) \),

\[
\xi_+ \xi_- = \xi^2 - \frac{1}{4} (\nu k)^2 + \xi \frac{k^2}{2m} + \frac{k^4}{64m^2}, \quad \nu \equiv \frac{p}{m}, \quad \xi \equiv \xi(p).
\]

(7A.45)

\[
\left\{ \frac{E_+^2}{E_-^2} \right\} = E^2 \pm \nu k + \frac{1}{4} (\nu k)^2 + \frac{2}{4m} \nu \xi \pm \frac{k^2}{8m} + \frac{k^4}{64m^2}.
\]

(7A.46)

The excitation spectrum is determined by the vanishing of the fluctuation determinant, i.e., from

\[
L_{11}(k)L_{22}(k) - L_{12}(k)L_{12}^*(k) = 0.
\]

(7A.47)

This equation has two solutions:

\[
L_{11}(k) = \pm L_{12}(k),
\]

(7A.48)

the first giving the low-, the second the high-energy excitations.

7A.3  Long-Wavelength Excitations at Zero Temperature.

At zero temperature, \( n(E) \) vanishes and the functions (7A.42), (7A.43) reduce to

\[
L_{11}(\epsilon, k) = \frac{1}{V} \sum_p \left\{ \frac{E_+ E_- + \xi_+ \xi_-}{2E_+ E_-} \frac{E_+ + E_-}{(E_+ + E_-)^2 - \epsilon^2} - \frac{1}{2E} \right\},
\]

(7A.49)

\[
L_{12}(\epsilon, k) = -\frac{1}{V} \sum_p \frac{\Delta^2}{2E_+ E_- (E_+ + E_-)^2 - \epsilon^2} \frac{E_+ + E_-}{2E_+ E_- (E_+ + E_-)^2 - \epsilon^2}.
\]

(7A.50)
Inserting here (7A.45) and (7A.46), and expanding the integrands in powers of $k$, we can perform the directional integrals using the formula
\[ \int \frac{d\hat{p}}{4\pi} (vk)^n = \frac{v^{2n/2}}{n+1} \left\{ \begin{array}{ll} 1, & n = \text{even}, \\ 0, & n = \text{odd}, \end{array} \right. \] (7A.51)
and obtain
\[ L_{11}(\epsilon, k) = -\frac{N(0)}{2} \left( 1 - \frac{\epsilon^2}{3\Delta^2} + \frac{v_F^2 k^2}{9\Delta^2} + \frac{v_F^2 \epsilon^2 k^2}{30\Delta^4} - \frac{\epsilon^4}{20\Delta^4} - \frac{v_F^4 k^4}{100\Delta^4} + \ldots \right), \] (7A.52)
\[ L_{12}(\epsilon, k) = -\frac{N(0)}{2} \left( 1 + \frac{\epsilon^2}{6\Delta^2} - \frac{v_F^2 k^2}{18\Delta^2} + \frac{v_F^2 \epsilon^2 k^2}{45\Delta^4} + \frac{\epsilon^4}{30\Delta^4} + \frac{v_F^4 k^4}{150\Delta^4} + \ldots \right). \] (7A.53)

We have ignored terms such as $k^4/m^2\Delta^2$ compared to $v_F^2 k^4/\Delta^4$ since the Fermi energy is much larger than the gap in a superconductor, i.e., $mv_F^2/2 \gg \Delta$. This limit is most easily accommodated by replacing $k$ by $\eta k$, and $v_F$ by $v_F/\eta$, and taking the limit $\eta \to 0$. In the expressions (7A.45) and (7A.46) this procedure eliminates the last term in comparison to the one before it.

The spectrum of the long-wavelength excitation is found from the equation $L_{11}(\epsilon, k) = L_{12}(\epsilon, k)$ which leads to the small-$k$ expansion (7.44) of the energy. For higher $k$ we must solve the equation $L_{11}(\epsilon, k) = L_{12}(\epsilon, k)$ numerically. The result is shown in Fig. 7.2.

![Figure 7.2](image_url)

**Figure 7.2** Energies of the low-energy excitations in superconductor. The abscissa is the momentum in units of $mv_F$, the ordinate the energy in units of the gap energy $\Delta$. The curve approaches the energy $\epsilon = 2\Delta$ for large $k$. The dashed curve shows the analytic small-$k$ expansion (7.44).

The fluctuations of the size of the order parameter are found by solving the equation $L_{11}(\epsilon, k) = -L_{12}(\epsilon, k)$. Since $\epsilon$ remains large we may perform only a small-$k$ expansion which leads to the energies [3, 4]
\[ \epsilon^{(n)}(k) = 2\Delta + \Delta \left( \frac{v_F k}{2\Delta} \right)^2 z_n \] (7A.54)
where $z_n$ are the solutions of the integral equation
\[ \int_{-1}^{1} dx \int_{-\infty}^{\infty} dy \frac{x^2 - z}{x^2 + y^2 - z} = 0. \] (7A.55)
Setting $e^t = (\sqrt{1-z} + 1)/(\sqrt{1-z} - 1)$ this turns into the algebraic equation

$$\frac{\pi}{2 \sinh^2(t/2)} (t + \sinh t) = 0,$$

which has infinitely many solutions $t_n$. The lowest is

$$t_1 = 2.25073 + 4.21239 i,$$

and tending asymptotically to

$$t_n \approx \log[\pi(4n - 1)] + i \left(2\pi n - \frac{\pi}{2}\right).$$

See the contour plot in Fig. 7.3.

**Figure 7.3** Contour plot of zeros for energy eigenvalues following from Eq. (7A.55) which are approximately given by Eq. (7A.58). The contour lines indicate fixed values of $\text{Abs}[(t + \sinh t)/\sinh(t/2)]$.

The excitation energies are

$$\epsilon^{(n)}(k) = 2\Delta - \frac{v_F^2 k^2}{4\Delta} \frac{1}{\sinh^2 t_n/2},$$

Of these, only the first at $\epsilon^{(1)}(k) \approx 2\Delta + (0.2369 - 0.2956 i)v_F^2/4\Delta^2 k^2$ lies in the second sheet and may have observable consequences. The others are hiding under lower and lower sheets of the two-particle branch cut from $2\Delta$ to $\infty$. This cut is logarithmic due to the dimensionality $D = 2$ of the surface of the Fermi sea at $T = 0$.

**7A.4 Long-Wavelength Excitations at Nonzero Temperature**

Consider now the case of nonzero temperature where the energy gap is calculated from Eq. (7A.3).
Let us first study the static case and consider only the long-wavelength limit of small \( k \). Hence, we set \( \epsilon = 0 \) and keep the lowest orders in \( k \) only. At \( k = 0 \) we find from (7A.42) and (7A.8):

\[
L_{11}(0, 0) = N(0) \int_{-\infty}^{\infty} d\xi \left\{ \frac{E^2 + \xi^2}{4E^3} \frac{1}{1 - 2n(E)} - \frac{E^2 - \xi^2}{2E^2} n'(E) - \frac{1}{2E} [1 - 2n(E)] \right\}.
\]

(7A.60)

Inserting \( E = \sqrt{\Delta^2 + \xi^2} \), and using the reduced variable \( \delta \equiv \Delta/T \) of Eq. (7A.30), this becomes

\[
L_{11}(0, 0) = -\frac{1}{2} N(0) \phi(\delta)
\]

(7A.61)

where we have introduced the so-called Yoshida function

\[
\phi(\delta) \equiv \Delta^2 \int_{0}^{\infty} d\xi \left\{ \frac{1}{E^3} [1 - 2n(E)] + 2 \frac{1}{E^2} n'(E) \right\}.
\]

(7A.62)

Here we observe that

\[
\partial_\xi \left[ \frac{\xi}{\Delta^2 E} n(E) \right] = \frac{1}{E^3} n(E) + \left( \frac{1}{\Delta^2} - \frac{1}{E^2} \right) n'(E),
\]

(7A.63)

to bring (7A.62) to the form

\[
\phi(\delta) \equiv \Delta^2 \int_{0}^{\infty} d\xi \left\{ \frac{1}{E^3} \left[ 1 - 2n(E) \right] + 2 \frac{1}{E^2} n'(E) \right\}.
\]

(7A.64)

The surface term vanishes, and the first integral in Eq. (7A.64) can be done, so that we arrive at the more convenient form

\[
\phi(\delta) = 1 + 2 \int_{0}^{\infty} d\xi n'(E) = 1 - \frac{1}{2T} \int_{0}^{\infty} d\xi \frac{1}{\cosh^2(E/2T)}.\]

(7A.65)

For \( T \approx 0 \), this function approaches zero exponentially.

For the function \( L_{12}(\epsilon, k) \) in Eq. (7A.43) we find at \( \epsilon = 0, k = 0 \):

\[
L_{12}(0, 0) = -N(0) \Delta^2 \int_{-\infty}^{\infty} d\xi \left\{ \frac{1}{4E^3} [1 - 2n(E)] + \frac{1}{2E^2} n'(E) \right\}.
\]

(7A.66)

This can again be expressed in terms of the Yoshida function (7A.65) as

\[
L_{12}(0, 0) = -\frac{1}{2} N(0) \phi(\delta).
\]

(7A.67)

This implies that the modes following from the equation \( L_{11}(\epsilon, k) = L_{12}(\epsilon, k) \) which at \( T = 0 \) have zero energy for \( k = 0 \), keep this property also for nonzero \( T \). This is a consequence of the Nambu-Goldstone theorem.

The full temperature behavior is best calculated by using the Matsubara sum (7A.4) to obtain

\[
T \sum_{\omega_n} \frac{1}{\omega_n^2 + E^2} = \frac{1}{2E} [n(-E) - n(E)] = \frac{1}{2E} \tanh \frac{E}{2T}.
\]

(7A.68)
Taking the derivative of this with respect to the energy we see that (7A.65) can be rewritten as

$$\phi(\delta) = 2T \sum_{\omega_m} \frac{\Delta^2}{(\omega_m^2 + E^2)^2} = -2\Delta^2 T \sum_{\omega_m} \frac{\partial}{\partial \Delta^2} \int d\xi \frac{1}{\omega_m^2 + \xi^2 + \Delta^2}$$

$$= -2\Delta^2 T \sum_{\omega_m} \frac{\partial}{\partial \Delta^2} \frac{\pi}{\omega_m^2 + \Delta^2} = 2T \pi \sum_{\omega_m > 0} \frac{\Delta^2}{\sqrt{\omega_m^2 + \Delta^2}}. \quad (7A.69)$$

Using the reduced variable (7A.31), this becomes

$$\phi(\delta) = \frac{2\pi}{\delta} \sum_{n=0}^{\infty} \frac{1}{\sqrt{x_n^2 + 1}}. \quad (7A.70)$$

For $T$ near $T_c$ where $\delta$ is small [see Eq. (7A.34)], we can approximate

$$\phi(\Delta) \approx 2\frac{\delta^2}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^{3/2}} = 2\frac{\delta^2}{\pi^2} \frac{7\zeta(3)}{8} \approx 2 \left(1 - \frac{T}{T_c}\right). \quad (7A.71)$$

In the limit $T \to 0$, the sum turns into an integral. Using the formula

$$\int_0^\infty dx \frac{x^{\mu-1}}{(x^2 + 1)^{1/2}} = \frac{1}{2} B(\mu/2, \nu_n - \mu/2) \quad (7A.72)$$

with $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$ we verify that

$$\phi(\Delta)|_{T=0} = 1. \quad (7A.73)$$

### 7A.5 Bending Energies of Order Field

Let us now calculate the bending energies of the collective field $\Delta(x)$. For this, we expand $L_{11}(0, k)$ and $L_{12}(0, k)$ in powers of the momentum $k$ up to $k^2$. We start from the Matsubara sums (7A.40) and (7A.68):

$$L_{11}(0, k) = \frac{T}{V} \sum_{\omega_m, p} \frac{\omega_m^2 + \xi_+ \xi_-}{(\omega_m^2 + E_+^2)(\omega_m^2 + E_-^2)} - \frac{1}{g}, \quad (7A.74)$$

$$L_{12}(0, k) = -\frac{T}{V} \sum_{\omega_m, p} \Delta^2 \frac{\omega_m^2 + \xi_+ \xi_-}{(\omega_m^2 + E_+^2)(\omega_m^2 + E_-^2)}. \quad (7A.75)$$

Inserting the expansions (7A.45) and (7A.46), these become

$$L_{11}(0, k) - L_{12}(0, k) \approx \int \frac{d^3p}{(2\pi)^3} T \sum_{\omega_m} \frac{\omega_m^2 + \Delta^2 + \xi^2 - \frac{1}{4} (vk)^2}{\left(\omega_m^2 + E^2\right)^2} \left[1 + \frac{1}{2} (vk)^2 \left(\omega_m^2 + \xi^2 + \Delta^2\right) \left(\omega_m^2 + E^2\right)^{-1}\right] - \frac{1}{g} + \ldots$$

$$= \int \frac{d^3p}{(2\pi)^3} \left\{ T \sum_{\omega_m} \frac{1}{\omega_m^2 + E^2} - \frac{1}{g} \right\}$$

$$+ T \sum_{\omega_m} \left\{ \frac{1}{4} \frac{1}{\omega_m^2 + E^2} - \frac{\omega_m^2 + \Delta^2}{\left(\omega_m^2 + E^2\right)^3} \right\} (vk)^2 \right\} + \ldots (7A.76)$$
Due to the gap equation (7A.3), the first term in the curly brackets vanishes, and using the directional integral (7A.51), we find
\[ L_{11}(0, k) - L_{12}(0, k) \approx N(0) \frac{v_F k^2}{3} \times \sum_{\omega_m} \int_{-\infty}^{\infty} d\xi \left[ \frac{1}{4 (\omega_m^2 + \xi^2 + \Delta^2)^2} - \frac{\omega_m^2 + \Delta^2}{(\omega_m^2 + \xi^2 + \Delta^2)^2} \right]. \] (7A.77)

Similarly we obtain
\[ L_{12}(0, k) \approx -N(0) \sum_{\omega_m} \int_{-\infty}^{\infty} d\xi \left\{ \frac{\Delta^2}{(\omega_m^2 + \xi^2 + \Delta^2)^2} \right. \]
\[ + \frac{v_F^2 k^2 \Delta^2}{3} \left[ \frac{1}{2 (\omega_m^2 + \xi^2 + \Delta^2)^3} - \frac{\omega_m^2 + \Delta^2}{(\omega_m^2 + \xi^2 + \Delta^2)^4} \right] \} \] (7A.78)

Using the integrals
\[ \int_{-\infty}^{\infty} d\xi \frac{1}{(\omega_m^2 + \xi^2 + \Delta^2)^{2,3,4}} = \left( \frac{1}{2}, \frac{3}{8}, \frac{5}{16} \right) \frac{\pi}{\sqrt{\omega_m^2 + \Delta^2}}, \] (7A.79)
we find, using (7A.69),
\[ L_{11}(0, k) - L_{12}(0, k) \approx -N(0) \frac{v_F^2 k^2}{4\Delta^2} \phi(\delta), \] (7A.80)
\[ L_{12}(0, k) \approx -\frac{N(0)}{2} \phi(\delta) + \frac{N(0) v_F^2 k^2 \Delta^2}{3} \bar{\phi}(\delta), \] (7A.81)

where \( \phi(\delta) \) is the Yoshida function (7A.70), while \( \bar{\phi}(\delta) \) is a further gap function:
\[ \bar{\phi}(\delta) \equiv 3\Delta^4 \pi T \sum_{\omega_m \geq 0} \frac{1}{\sqrt{\omega_m^2 + \Delta^2}^4} = \frac{3\pi}{\delta} \sum_{n=0}^{\infty} \frac{1}{\sqrt{x_n^2 + 1}}, \] (7A.82)

In the limit \( T \to 0 \), the sum turns into an integral whose value is, by formula (7A.72),
\[ \bar{\phi}(\delta) \bigg|_{T=0} = 1. \] (7A.83)

Together with (7A.73) we see that (7A.80) and (7A.81) reproduces correctly the \( k^2 \)-terms of Eqs. (7A.52) and (7A.53).

For \( T \approx T_c \), where \( \delta \to 0 \), we find
\[ \bar{\phi}(\delta) \approx \frac{3\delta^4}{\pi^4} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^8} = \frac{3\delta^4}{\pi^4} \frac{31\zeta(5)}{32}, \] (7A.84)

H. Kleinert, GRAVITY WITH TORSION
and thus, by (7A.34),
\[ \tilde{\phi}(\delta) \approx \frac{3}{\pi^4} \frac{31\zeta(5)}{32} \left( \frac{8\pi^2}{7\zeta(3)} \right)^2 \left( 1 - \frac{T}{T_c} \right)^2 \approx 2.7241 \left( 1 - \frac{T}{T_c} \right)^2. \] (7A.85)

Inserting this together with (7A.71) into (7A.80) and (7A.81), and considering only long-wavelength excitations with \( v_F^2 k^2 \leq \Delta^2 \), so that \( (v_F^2 k^2 / \Delta^2) \tilde{\phi}(\delta) \) is of the order of \( (1 - T/T_c)^2 \), we obtain to lowest order in \( 1 - T/T_c \):
\[ L_{11}(0, k) - L_{12}(\epsilon, k) \approx -N(0) \frac{v_F^2}{\pi^2 T_c^2} 7\zeta(3) k^2, \] (7A.86)
\[ L_{12}(0, k) \approx -N(0) \left( 1 - \frac{T}{T_c} \right). \] (7A.87)

Using (5.137), we see that
\[ N(0) v_F^2 = \frac{3}{2} \frac{n_e}{m} = \frac{3}{2} \frac{n_\Delta}{m} = \frac{3}{2} \frac{\rho}{m^2} = 6 \frac{\rho}{M^2}, \] (7A.88)
where \( n_\Delta \) is the density of Cooper pairs whose mass is \( M = 2m \), and \( \rho \equiv M n_\Delta = 2m n_\Delta = mn_e \) their mass density. Note that the mass density of the pairs is equal to that of the electrons: \( \rho \equiv 2M n_e = M n_\Delta \). Thus we may eliminate \( N(0) \) in favor of the pair density and obtain
\[ L_{11}(0, k) - L_{12}(0, k) \approx -\rho \frac{2}{2} \frac{M^2}{\Delta^2} k^2 \phi(\delta), \] (7A.89)
\[ L_{12}(0, k) \approx -\rho \frac{2}{2} \frac{M^2}{\Delta^2} k^2 \phi(\delta) + \frac{6 \rho}{4 M^2 \Delta^2} \phi(\delta). \] (7A.90)

Returning to \( x \)-space we decompose the collective field \( \Delta(x) \) into a real size field \( |\Delta(x)| \) fluctuating around \( \Delta \) and a phase field \( \theta(x) \) fluctuating around zero as
\[ \Delta(x) = |\Delta(x)| e^{i\theta(x)} \] (7A.91)
and extract from the action (7A.39), the energy for static small fluctuations as
\[ E(x) = \frac{1}{2 M^2 \Delta^2} \{ \rho^{11} \nabla \Delta^*(x) \nabla \Delta(x) + \text{Re} \left[ \rho^{12} \nabla \Delta^*(x) \nabla \Delta^*(x) + a^{12} \Delta^*(x) \delta \Delta^*(x) \right] \}. \] (7A.92)

From (7A.89) and (7A.90) we identify the coefficients as:
\[ \rho^{11} - \rho^{12} = \rho \phi(\delta), \quad \rho^{12} = -\frac{2}{3} \frac{\rho}{v_F^2} \tilde{\phi}(\delta), \quad a^{12} = 6 \rho \frac{\Delta^2}{v_F^2} \phi(\delta). \] (7A.93)

Decomposing the collective field \( \Delta(x) \) into a real size field \( \Delta(x) \) and a phase field \( \theta(x) \) as
\[ \Delta(x) = \Delta(x) e^{i\theta(x)} , \] (7A.94)
the energy density reads
\[ E(x) = \frac{1}{2M^2} \left\{ \rho_s(\nabla \theta - \mathbf{v})^2 + \rho_\Delta(\nabla|\Delta(x)|)^2/\Delta^2 + 2a^{12}|\delta|\Delta(x)||^2 \right\}. \tag{7A.95} \]

The factor before the first gradient term is the superfluid density:
\[ \rho_s \equiv \rho^{11} - \rho^{12} = \rho \phi(\delta). \tag{7A.96} \]

By analogy we have introduced the quantities
\[ \rho_\Delta \equiv \rho^{11} + \rho^{12} = \rho_s + 2\rho^{12} = \rho - \frac{2}{3} \bar{\rho}_s, \quad \bar{\rho}_s \equiv \rho \bar{\phi}(\delta). \tag{7A.97} \]

The behavior of \( \rho_s \) and \( \bar{\rho}_s \) for all \( T \leq T_c \) is shown in Fig. 7.4.

![Figure 7.4](image.png)

**Figure 7.4** Temperature behavior of superfluid density \( \rho_s/\rho \) (Yoshida function) and the gap function \( \bar{\rho}_s/\rho \) of Eqs. (7A.96) and (7A.97).

The phase fluctuations are of infinite range, the size fluctuations have a finite range characterized by the temperature-dependent coherence length (with reinserted \( \hbar \) to have proper physical units)
\[ \xi(T) = \frac{\hbar v_F}{\Delta} \sqrt{\frac{1}{12\Delta^2} \frac{\rho_s - 2\bar{\rho}_s/3}{\rho_s}}. \tag{7A.98} \]

Near \( T_c \), we insert \( \rho_s \approx 2(1 - T/T_c) \), \( \bar{\rho}_s \approx (1 - T/T_c)^2 \), and \( \Delta \) from (7A.35) to find the limit (in physical units)
\[ \xi(T) = \frac{\xi_0}{\sqrt{2}} \left( 1 - \frac{T}{T_c} \right)^{-1/2}, \quad \xi_0 \equiv \sqrt{\frac{7\zeta(3)}{48} \frac{\hbar v_F}{\pi k_B T_c}} \approx 0.42 \frac{\hbar v_F}{\pi k_B T_c} \sim 0.27 \frac{T_F}{T_c}, \tag{7A.99} \]

The last expression we have introduced the Fermi length \( l_F = \hbar/p_F \), and the Fermi temperature \( T_F = p_F^2/2M k_B \). In old-fashioned superconductors, \( l_F \) is of the order of the lattice spacing, while \( T_F \) is usually larger than \( T_c \) by factors \( 10^3 - 10^4 \), so that the ratio of the coherence length with respect to the Fermi length is quite large. In
high-temperature superconductors, however, \( \xi_0 \) can shrink to only a few times \( l_F \), which greatly increases the effect of fluctuations.

At zero temperature we obtain, with the help of (7A.24):

\[
\xi(0) = \frac{1}{6} \frac{v_F}{\Delta(0)} \approx 0.0935 \frac{v_F}{\pi T_c},
\]

which is about a sixth of the length parameter \( \xi_0 \) in the \( T \approx T_c \) -equation (7A.99).

### 7A.6 Kinetic Terms of Pair Field at Nonzero Temperature

At nonzero temperature, we shall extract the dynamics of the kinetic term of slowly varying pair field by calculating from Eqs. (7A.42) and (7A.43) at small \( k \). We begin with \( k = 0 \) where we obtain, instead of (7A.60) and (7A.66):

\[
L_{11}(\epsilon, 0) - L_{11}(0, 0) = N(0) \frac{\epsilon^2}{4\Delta^2} \left( \frac{\Delta}{E^2 - \epsilon^2/4} \right) [1 - n(E)],
\]

(7A.101)

\[
L_{12}(\epsilon, 0) - L_{12}(0, 0) = -N(0) \frac{\epsilon^2}{4\Delta^2} \frac{1}{4} \int_{-\infty}^{\infty} d\xi \frac{1}{E^3} \left( \frac{\Delta^4}{E^2 - \epsilon^2/4} \right) [1 - n(E)].
\]

(7A.102)

They can be rewritten as

\[
L_{11}(\epsilon, 0) - L_{12}(\epsilon, 0) = N(0) \frac{\epsilon^2}{4\Delta^2} \gamma(\delta, \epsilon),
\]

(7A.103)

\[
L_{12}(\epsilon, 0) - L_{12}(0, 0) = -N(0) \frac{\epsilon^2}{4\Delta^2} \frac{1}{3} \tilde{\gamma}(\delta, \epsilon),
\]

(7A.104)

where

\[
\gamma(\delta, \epsilon) = \frac{1}{2} \int_{-\infty}^{\infty} d\xi \frac{\Delta^2}{E(E^2 - \epsilon^2/4)} [1 - 2n(E)],
\]

(7A.105)

\[
\tilde{\gamma}(\delta, \epsilon) = \frac{3}{4} \int_{-\infty}^{\infty} d\xi \frac{\Delta^4}{E^3(E^2 - \epsilon^2/4\Delta^2)} [1 - 2n(E)].
\]

(7A.106)

At zero temperature where \( n(E) = 0 \), both \( \gamma(\delta, 0) \) and \( \tilde{\gamma}(\delta, 0) \) start out with the value 1, so that the results (7A.103) and (7A.104) reproduce the \( \epsilon^2 \)-terms of (7A.52) and (7A.53). The full temperature behavior of \( \gamma(\delta, 0) \) and \( \tilde{\gamma}(\delta, 0) \) is plotted in Fig. 7.5.

For arbitrary temperature we calculate (7A.105) and (7A.106) most conveniently by expanding \( n(E) \) into a Matsubara sum via Eq. (7A.6), so that it takes the form

\[
\gamma(\delta, \epsilon) = 2T \sum_{\omega_m > 0} \int_{-\infty}^{\infty} d\xi \frac{\Delta^2}{E^2 - \epsilon^2/4} \frac{1}{\omega_m^2 + E^2},
\]

(7A.107)

\[
\tilde{\gamma}(\delta, \epsilon) = 3T \sum_{\omega_m > 0} \int_{-\infty}^{\infty} d\xi \frac{\Delta^4}{E^2(E^2 - \epsilon^2/4)} \frac{1}{\omega_m^2 + E^2}.
\]

(7A.108)
Performing the integrals over $\xi$ yields

$$
\gamma(\delta, \epsilon) = 2\pi \frac{T}{\Delta} \sum_{\omega_m > 0} \left( \frac{1}{\sqrt{1 - \epsilon^2/4\Delta^2}} - \frac{\Delta^2}{\omega_m^2 + \epsilon^2/4\Delta^2 \sqrt{\omega_m^2 + \Delta^2}} \right), \quad (7A.109)
$$

$$
\bar{\gamma}(\delta, \epsilon) = 3\pi \frac{T}{\Delta} \sum_{\omega_m > 0} \left[ \frac{4\Delta^2}{\epsilon^2} \frac{\Delta^2}{\omega_m^2 + \epsilon^2/4} \left( \frac{1}{\sqrt{1 - \epsilon^2/4\Delta^2}} - 1 \right) + \left( \frac{\Delta}{\sqrt{\omega_m^2 + \Delta^2}} - 1 \right) \frac{\Delta^2}{\omega_m^2 (\omega_m^2 + \epsilon^2/4)} \right]. \quad (7A.110)
$$

For $\epsilon = 0$, these become

$$
\gamma(\delta, 0) = 2\pi \frac{T}{\Delta} \sum_{\omega_m > 0} \left( \frac{\Delta^2}{\omega_m^2} - \frac{\Delta^3}{\omega_m^2 \sqrt{\omega_m^2 + \Delta^2}} \right), \quad (7A.111)
$$

and

$$
\bar{\gamma}(\delta, 0) = 3\pi \frac{T}{\Delta} \sum_{\omega_m > 0} \left[ \frac{1}{2\omega_m^2} + \frac{\Delta^4}{\omega_m^4} \left( \frac{\Delta}{\sqrt{\omega_m^2 + \Delta^2}} - 1 \right) \right]. \quad (7A.112)
$$

Here we can replace

$$
\sum_{\omega_m > 0} \omega_m^{-k} = (1 - 2^{-k}) \zeta(k)(\pi T)^{-k}, \quad (7A.113)
$$

which is equal to $1/8T^2$ for $k = 2$, and $1/96T^4$ for $k = 4$. In the limit $T \to 0$, the Matsubara sums $T \sum_{\omega_m > 0} \omega_m^{-k}$ become integrals $\int_0^\infty d\omega_m/2\pi$ and we recover the limits $\gamma(\delta) \to 1$, $\bar{\gamma}(\delta) \to 1$ obtained before from (7A.105) and (7A.106).

In the limit $T \to T_c$ where $\Delta \to 0$, the functions (7A.109) and (7A.110) have the limit

$$
\gamma(\delta, \epsilon) \to 2\pi \frac{T\Delta}{\sqrt{1 - \epsilon^2/4\Delta^2}} \sum_{\omega_m > 0} \frac{1}{\omega_m^2} = \frac{\pi \Delta}{4T \sqrt{1 - \epsilon^2/4\Delta^2}} \to \frac{\pi \Delta^2}{2T \sqrt{-\epsilon^2}}, \quad (7A.114)
$$

\[\text{Figure 7.5 Temperature behavior of the functions } \gamma(\delta, 0) \text{ and } \bar{\gamma}(\delta, 0) \text{ of Eqs. (7A.105) and (7A.106).}\]
\[ \frac{\epsilon^2}{4\Delta^2} \bar{\gamma}(\delta,\epsilon) \rightarrow 3\pi \frac{T\Delta}{\sqrt{1 - \epsilon^2/4\Delta^2}} \sum_{\omega_m > 0} \frac{1}{\omega_m^2} = \frac{3\pi\Delta}{8T\sqrt{1 - \epsilon^2/4\Delta^2}} \rightarrow \frac{3\pi\Delta^2}{4T\sqrt{-\epsilon^2}}, \]  

so that

\[ L_{11}(\epsilon, k) - L_{12}(\epsilon, k) \approx N(0) \frac{i\pi\epsilon}{8T}, \]  

\[ L_{12}(\epsilon, 0) - L_{12}(0, 0) \approx -N(0) \frac{i\pi\Delta^2}{4T\epsilon}. \]

For \( \epsilon \gg \Delta^2 \), the second function can be ignored in comparison with the first.

The same results could have been derived directly from Eq. (7A.103) and (7A.102) for \( \Delta = 0 \):

\[ L_{11}(\epsilon, 0) - L_{12}(\epsilon, 0) \approx N(0) \int_{-\infty}^{\infty} d\xi \left[ \frac{\xi}{2(\xi^2 - \epsilon^2/4)} - \frac{1}{2\xi} \right] \tanh \frac{\xi}{2T}, \]  

\[ L_{12}(\epsilon, 0) \approx 0. \]

Together with (7A.86) (7A.87), and (7A.100), this yields

\[ L_{11}(\epsilon, k) - L_{12}(\epsilon, k) \equiv -N(0) \left( -\frac{i\pi\epsilon}{8T} + \xi_0^2 k^2 + \ldots \right), \]  

\[ L_{12}(\epsilon, k) = -N(0) \left( 1 - \frac{T}{T_c} \right) + \ldots. \]

This shows that for \( T \approx T_c \), the excitations are purely damped with a decay rate

\[ \Gamma = 2\frac{8T}{\pi} \xi_0^2 k^2. \]  

The above results provide us with all information to set up a Ginzburg-Landau action for describing a superconductor in the neighborhood of the critical temperature. This action is a low-order expansion in powers of the Cooper pair field and its gradients

\[ A_2[\Delta, \Delta^*] \approx N(0) \int dt \int d^3x \left\{ \Delta^*(x) \left[ -\frac{\pi}{8T} \partial_t + \xi_0^2 \nabla^2 - a_2 \right] \Delta(x) - \frac{a_4}{2} |\Delta(x)|^4 + \ldots \right\}, \]

where the gradient terms follow directly from (7A.120). The dots indicate higher expansion terms which contain more powers of the field such as \( |\Delta(x)|^6 \), \( |\Delta(x)|^8 \), \ldots, or higher derivative terms such as \( |\nabla^2 \Delta(x)|^2 \), \( |\partial_t \Delta(x)|^2 \), \ldots. For the study of the phase transition these are all irrelevant.

To determine \( a_2 \) and \( a_4 \) we insert the decomposition \( \Delta(x) = \Delta + \delta\Delta(x) \), into (7A.123) and find that for \( a_2 < 0 \), the action has an extremum at \( \Delta = \sqrt{-a_2/a_4} \). The quadratic fluctuations \( \delta\Delta(x) \) possess a the same gradient terms as in (7A.123), while the potential terms \( -a_2|\Delta(x)|^2 - a_4|\Delta(x)|^4/2 \) contribute

\[ \Delta A_2 \approx - \int dt \int d^3x \left[ \left( a_2 + 2a_4 \Delta^2 \right) \delta\Delta^*(x) \delta\Delta(x) + \frac{a_4}{2} \left( \Delta^2 \right) \left( [\delta\Delta(x)]^2 + [\delta\Delta^*(x)]^2 \right) \right]. \]
At the extremum, this becomes
\[ \Delta A_2 \approx a_2 \int dt \int d^3 x \left( \delta \Delta^* (x) \delta \Delta (x) + \frac{1}{2} \left\{ [\delta \Delta (x)]^2 + [\delta \Delta^* (x)]^2 \right\} \right). \] (7A.125)

The imaginary part of \( \delta \Delta (x) \) drops out ensuring that its static infinite-wavelength fluctuations have an infinite range, in accordance with the Nambu-Goldstone theorem.

Comparing (7A.125) with (7A.39) we identify \( a_4 \Delta^2 \) with \( L_{12}(0,0) \) in Eq. (7A.81) for small \( \Delta \), i.e.,
\[ a_4 \approx \frac{\mathcal{N}(0)}{2 \Delta^2} \phi (\delta) \approx \mathcal{N}(0) \frac{1}{\pi^2 T^2} \frac{7 \zeta (3)}{8} = \mathcal{N}(0) \frac{6 \xi_0^2}{h^2 v_F}. \] (7A.126)

The constant \( a_2 \) is then [recalling (7A.35) and (7A.99)]
\[ a_2 = - \Delta^2 a_4 \approx \mathcal{N}(0) \left( \frac{T}{T_c} - 1 \right). \] (7A.127)

Inserting this into (7A.123) we see that the fluctuations of \( \Delta (x) \) around \( \Delta \) have the coherence lengths
\[ \xi (T) = \xi_0 \left( \frac{T}{T_c} - 1 \right)^{-1/2}, \quad T > T_c, \] (7A.128)
\[ \xi_{\text{size}} (T) = \frac{\xi_0}{\sqrt{2}} \left( \frac{T}{T_c} - 1 \right)^{-1/2}, \quad T < T_c. \] (7A.129)

For critical temperatures of the order of 1 to 10 K, Fermi temperature of the order \( 10^4 \) to \( 10^5 \) K, and Fermi momenta of the order \( \hbar /rA \), we obtain quite a large coherence length of the order of \( 10^3 - 10^4 \) rA.

The energy in the action (7A.123) coincides with the Ginzburg-Landau energy (5.140) if we identify:
\[ \psi (x) = \sqrt{2 \mathcal{N}(0) \xi_0} \Delta (x), \quad \tau = \frac{1}{\xi_0^2} \left( \frac{T}{T_c} - 1 \right). \] (7A.130)

Then \( a_4 \) of Eq. (7A.125) implies that the coupling constant \( g \) in (5.140) has the BCS-value
\[ g = \frac{3}{\mathcal{N}(0) \hbar^2 v_F^2 \xi_0^2}. \] (7A.131)

The condensation energy density of the superconductor is given by
\[ \mathcal{E}_c = \frac{\tau^4}{4g} = \frac{1}{4 \xi_0^4} \frac{1}{3} \frac{(k_B T_c)^2 \xi_0^4}{7 \zeta (3) / 48 \pi^2} \mathcal{N}(0) \left( 1 - \frac{T}{T_c} \right)^2 = \frac{1}{7 \zeta (3)} \left( \frac{P_F}{\hbar} \right)^3 \frac{T_c}{T_F} k_B T_c \left( 1 - \frac{T}{T_c} \right)^2, \] (7A.132)

\[ ^2 \text{Note that the dimensions of } \psi, \mathcal{N}(0), \Delta, g \text{ are (energy/length)}^{1/2}, \text{ (energy)}^{-1}, \text{ length}^{-3}, \text{ energy, and (energy} \cdot \text{length})^{-1}, \text{ respectively.} \]
which is of the order
\[ \mathcal{E}_c \approx 10^{-4} k_B \left( 1 - \frac{T}{T_c} \right)^2 k_B K/rA^3 \approx 10^4 \left( 1 - \frac{T}{T_c} \right)^2 \text{erg/cm}^3. \] (7A.133)

In order to obtain a better idea of the size of interaction strength, it is useful to go to natural units used in the fluctuation discussion after Eqs. (5.255) and measure energies in units of \( k_B T_c \). In addition, we measure distances in units of \( \xi_0 \). Then, taking a factor \( \sqrt{k_B T_c/\xi_0} \) out of \( \psi \) and \( A \), and \( \xi_0 \) out of \( x \) (i.e., \( \sqrt{\xi_0/k_B T_c \psi}, A_{\text{new}} = \sqrt{\xi_0/k_B T_c A}, x_{\text{new}} = x/\xi_0 \)) we arrive at the dimensionless Ginzburg-Landau Hamiltonian
\[ \mathcal{H}_{\text{GL}} = \int d^4x \left\{ \frac{1}{2}(|\nabla - iqA)|^2 + \frac{1}{2} \left( \frac{T}{T_c} - 1 \right) |\psi|^2 + \frac{g}{4} |\psi|^4 + \frac{1}{2} (\nabla \times A)^2 \right\}. \] (7A.134)

Here the coupling \( g \) and \( q \) are dimensionless and their magnitudes are
\[ g = \frac{3 \xi_0 k_B T_c}{N(0) \hbar^2 v_F^2 \xi_0} = \frac{3}{2} \pi^2 \frac{48 \pi^2}{\sqrt{7} \xi(3)} \left( \frac{T_c}{T_F} \right)^2 \sim 111.08 \left( \frac{T_c}{T_F} \right)^2, \] (7A.135)
\[ q = \frac{2e}{\hbar c} k_B T_c \xi_0 = 2 \sqrt{4 \pi \alpha \frac{v_F}{c}} \frac{\sqrt{7} \xi(3)}{48 \pi^2} \sim 2.59 \sqrt{\alpha \frac{v_F}{c}}, \] (7A.136)
where \( \alpha = (e^2/4\pi)/\hbar c = 1/137 \) is the fine structure constant. Since \( T_c/T_F \sim 10^{-4} \) and \( \alpha(v_F/c) \sim 10^{-4} \), both couplings are extremely small, i.e., \( g \sim 10^{-6}, q \sim 10^{-2} \).

Gorkov’s original derivation [2] was valid only for perfect crystals. In dirty materials, the mean free path of the electron has finite value, say \( \ell \). In that case, the length scale \( \xi_0^2 \) in front of the gradient term of (7A.123) receives a correction factor
\[ r = \sum_{n=1}^{\infty} \frac{1}{(2n+1)^2(2n+1 + \xi_0/2\pi \cdot 0.18\ell)} / \sum_{n=1}^{\infty} \frac{1}{(2n+1)^2}, \] (7A.137)
with the other terms remaining the same. In the clean limit \( \ell = \infty \), \( r \) is equal to 1. In very dirty materials, however, \( \ell \ll \xi_0 \) and \( r \) becomes \( \sim 0.18(\xi_0) \) which can be quite small. If \( \xi_0' = r^{1/2} \xi_0 \) is used as a new length scale and \( \sqrt{r}^{1/2} \xi_0/k_B T_c \) is taken out of the fields, instead of \( \xi_0/k_B T_c \), the correction factor \( r \) changes the constants in the reduced energy as follows:
\[ g \rightarrow gr^{-3/2}, \quad q \rightarrow qr^{-1/4}. \] (7A.138)

Note that the reduced condensation energy density
\[ \beta_0 \mathcal{E}_c = \frac{1}{4g \xi_0^3} \left( 1 - \frac{T}{T_c} \right)^2 \] (7A.139)
remains unchanged since \( \xi_0 \) and \( g \) are modified by \( r \) oppositely. This is the content of a theorem by Abrikosov which states that dirt does not change the global thermodynamics of the superconductor.
Appendix 7B  Properties of Ginzburg-Landau
Theory of Superconductivity

Let us discuss some properties of the Ginzburg-Landau field theory with Hamiltonian
(5.140). The field equations are
\[
- (\nabla - iqA)^2 + \tau + g|\psi|^2 \psi = 0,
\]
and
\[
\nabla \times \nabla \times A = qj_s,
\]
with the supercurrent (5.141), which is conserved as a consequence of Eq. (7B.1):
\[
\nabla \cdot j_s = 0.
\]
In order to show this explicitly, observe that the product rule of differentiation holds
also for the covariant derivative of a product of complex fields:
\[
\nabla (a^\dagger b) = (\nabla a^\dagger) b + a^\dagger (\nabla b) = (\nabla + iqA) a^\dagger b + a^\dagger (\nabla - iqA) b
\]
\[
= (D a^\dagger) b + a^\dagger (D b).
\]
Thus in each term of the product rule, we may directly replace the ordinary derivative
\(\nabla\) by the covariant one \(D = \nabla - iqA\). Applying this rule to (7B.3), we see
that
\[
\nabla \cdot j_s = \frac{1}{2i} \left\{ (D \psi)^\dagger (D \psi) + \psi^\dagger D^2 \psi - (D \psi)^\dagger (D \psi) - (D^2 \psi)^\dagger \psi \right\}
\]
\[
= \frac{1}{2i} \left\{ \psi^\dagger D^2 \psi - (D^2 \psi)^\dagger \psi \right\},
\]
which vanishes indeed due to the field equation (7B.1).

The invariance of the Ginzburg-Landau equations (7B.1) and (7B.1) under the
gauge transformations
\[
A(x) \rightarrow A(x) + \nabla \Lambda(x), \quad \psi(x) \rightarrow e^{iq\Lambda(x)} \psi(x)
\]
can be used to transform away the phase of the \(\psi\)-field. As in (5.142), we shall
parametrize it in terms of size and phase angle as \(\psi(x) = \rho(x)e^{i\theta(x)}\), but omit
the wiggles for brevity. We may choose \(q\Lambda(x) = -\theta(x)\), and the field equations become
\[
- (\nabla - iqA)^2 + \tau + g\rho^2 \rho = 0,
\]
\[
\nabla \times \nabla \times A(x) = qj_s = -q^2 \rho^2 A.
\]
Separating real and imaginary parts, the first equation decomposes into an equation
for \(\rho(x)\):
\[
(-\nabla^2 + q^2 A^2 + \tau + g\rho^2)\rho = 0,
\]
and another one for $\mathbf{A}(x)$:

$$\rho \nabla \cdot \mathbf{A} + 2\mathbf{A} \cdot \nabla \rho = 0. \quad (7B.10)$$

The latter expresses the current conservation law (7B.3) in terms of size and phase fields where (5.141) reads

$$j_s(x) = \rho^2 [\nabla \theta(x) - q\mathbf{A}(x)]. \quad (7B.11)$$

The Hamiltonian in these field variables was written down in Eq. (5.143). In the present notation without wiggles, the Hamiltonian density without vortices reads:

$$H_{GL} = \int d^3x \int d^3x \left\{ \frac{1}{2} (\nabla \rho)^2 + \frac{\tau}{2} \rho^2 + \frac{g}{4} \rho^4 + \frac{1}{2} q^2 \rho^2 \mathbf{A}^2 + \frac{1}{2} (\nabla \times \mathbf{A})^2 \right\}. \quad (7B.12)$$

Let us extract from this a few experimental properties. The derivation of the finite penetration depth $\lambda$ in the superconducting phase was described in Section 5.2.1. The result was, in the present natural units,

$$\lambda = 1/qp. \quad (7B.13)$$

Here we exhibit a few more important properties.

### 7B.1 Critical Magnetic Field

The Ginzburg-Landau equations explain the existence of a critical external magnetic field $H_c$ at which the Meissner effect breaks down and the field invades into the superconductor, thereby destroying all supercurrents. This is most easily derived by studying the magnetic enthalpy, whose density is

$$\mathcal{E}_H = \mathcal{H}_{GL} - \mathbf{H} \cdot \mathbf{H}^{ext}. \quad (7B.14)$$

We can then see that for $H^{ext} < H_c = (1/\sqrt{2})|\tau|/\sqrt{g}$ the enthalpy $\mathcal{E}_H$ is minimized by $\rho_0 = \sqrt{-\tau/g}$, $\mathbf{A} = 0$ with a minimal density

$$\mathcal{E}_H = -\frac{\tau^2}{4g} = \mathcal{E}_c. \quad (7B.15)$$

For $H^{ext} > H_c$, however, the minimum of (7B.14) lies at $\rho = 0$, $H = H^{ext}$, where it has the value

$$\mathcal{E}_H = -\frac{(H^{ext})^2}{2}. \quad (7B.16)$$

Since the order parameter vanishes, this state is no longer superconducting.

For $H^{ext} = H_c$ the system can be in either state. In cgs units, the critical field is given by $(H_c^{ext})^2/2 = \mathcal{E}_c$ so that its order of magnitude lies, according to (7A.132), (7A.133), in the range of a few gauss.

The interesting consequence of the Ginzburg-Landau equations is that it allows for both the superconducting and the normal phase in one and the same sample separated by domain walls. This mixed state, also called the Shubnikov phase, is experimentally of particular importance and deserves some discussion.
7B.2 Two Length Scales and Type I or II Superconductivity

In the superconducting phase with field expectation value $\rho = \rho_0 = \sqrt{-\tau/g}$, the $\rho$ fluctuations $\delta \rho = \rho - \rho_0$ have a coherence length given by Eq. (7A.129) in natural units:

$$\xi_{\text{size}}(T) = \frac{1}{\sqrt{-2\tau}}.$$  

(7B.17)

By comparison with Eq. (7B.13) we find the ratio of penetration depth and coherence length, the Ginzburg parameter $\kappa$ [recall (5.249)]

$$\kappa \equiv \frac{\lambda}{\sqrt{2} \xi_{\text{size}}} = \sqrt{\frac{g}{q^2}}.$$  

(7B.18)

For $\kappa > 1/\sqrt{2}$ or $< 1/\sqrt{2}$, the magnetic penetration depth is larger or smaller than the coherence length of the order parameter. These two cases are called type-II and type-I superconductivity, respectively.

Inserting Eq. (7A.136) we estimate

$$\kappa \approx \frac{4.06}{\sqrt{\alpha v_F/c T_F}},$$  

(7B.19)

which is of the order of 1/10. Thus, a clean superconductor is usually of type-I.

In a dirty superconductor the result is modified by a factor $r^{-1}$ as a consequence of Eq. (7A.138).

Hence impurities can bring $\kappa$ into the type-II zone. In aluminium, for instance, 0.1% of impurities are sufficient to achieve this.

Let us now study types of domain walls between normal and superconducting materials; they differ significantly for the two types of superconductors. It will be convenient to go to a further reduced field variable $\hat{\rho} = \rho / \sqrt{-\tau/g}$ which, in the superconductive state, fluctuates around unity instead of $\rho_0 = \sqrt{-\tau/g}$. Similarly we shall define a reduced vector potential $\hat{A} = A / (\kappa \sqrt{-\tau/g})$ and measure lengths in units of the temperature dependent coherence length $r^{1/2} \xi_0 / \sqrt{-\tau}$, rather than $r^{1/2} \xi_0$. Then the Hamiltonians (7A.134) and (7B.12) become, for $\tau < 0$,

$$\hat{H}_{\text{GL}} \equiv \frac{g}{r^2} H_{\text{red}} = \frac{1}{2} \int d^3 x \left\{ |(\nabla - iA) \psi|^2 - |\psi|^2 + \frac{1}{2} |\psi|^4 + \frac{\kappa^2}{2} (\nabla \times A)^2 \right\},$$  

(7B.20)

and in size-phase angle fields

$$\hat{H}_{\text{GL}} = \frac{1}{2} \int d^3 x \left\{ (\nabla \rho)^2 - \rho^2 + \frac{1}{2} \rho^4 + \left[ \rho^2 A^2 + \kappa^2 (\nabla \times A)^2 \right] \right\},$$  

(7B.21)

where we have dropped the hats on top of the fields, for brevity. The associated supercurrent density is

$$j_s = \frac{1}{2i} \bar{\psi} \nabla \psi - \bar{A} |\psi|^2 = -\rho^2 A.$$  

(7B.22)
We also define a reduced magnetic field

\[ \mathbf{H} \equiv \kappa \nabla \times \mathbf{A}, \]  

(7B.23)
such that the magnetic field energy takes the usual form \( \mathbf{H}^2/2 \), and the critical magnetic field \( H_c \) is equal to \( 1/\sqrt{2} \). In these units, the field equations (7B.7) and (7B.8) read simply

\[ \left( -\nabla^2 + \mathbf{A}^2 - 1 + \rho^2 \right) \rho = 0, \] 

(7B.24)

\[ \kappa^2 \nabla \times (\nabla \times \mathbf{A}) = \kappa \nabla \times \mathbf{H} = -\rho^2 \mathbf{A}. \] 

(7B.25)

They can be solved for an \( \mathbf{H} \) and a \( \rho \) field varying, say, along the \( x \)-direction with \( \mathbf{H} \) pointing in the \( y \)-direction. Accordingly, we choose a potential along the \( z \)-direction

\[ \mathbf{A}(x) = (0, 0, -A(x)), \] 

(7B.26)

so that [with \( ' \equiv \partial_x \)]

\[ H(x) = \kappa A'(x). \] 

(7B.27)

The field equations are

\[ -\rho''(x) + A^2 \rho(x) = \rho(x) - \rho^3(x), \] 

(7B.28)

\[ \kappa^2 A''(x) = \kappa H'(x) = \rho^2 A(x). \] 

(7B.29)

Differentiating the second equation, it reduces to an equation for the magnetic field

\[ \rho^2 \mathbf{H} = \kappa^2 (H'' - 2H' \rho' / \rho) = \kappa^2 \rho^2 \left( \frac{1}{\rho^2} H' \right)' . \] 

(7B.30)

In the first equation we can eliminate \( A \) in favor of the magnetic field by writing the second equation as

\[ A = \kappa^2 A'' / \rho^2 = \kappa H' / \rho^2 \] 

(7B.31)

so that

\[ -\rho'' + \kappa^2 H'^2 / \rho^3 = \rho - \rho^3. \] 

(7B.32)

Now we observe that for the value \( \kappa = 1/\sqrt{2} \), where magnetic and size fluctuations have equal length scales, these equations become particularly simple. For, if we make a trial ansatz

\[ H = \frac{1}{\sqrt{2}} (1 - \rho^2) \] 

(7B.33)
and insert it into Eq. (7B.30), this takes the form
\[
\frac{1}{\sqrt{2}} (1 - \rho^2) \rho^2 = -\frac{1}{\sqrt{2}} (\rho \rho'' - \rho^2).
\]  
(7B.34)

But this happens to coincide with the second field equation (7B.32). Moreover, introducing \( \sigma = 2 \log \rho \) we see that (7B.34) reduces to a differential equation of the Liouville type
\[
\frac{\sigma''}{2} = e^\sigma - 1.
\]  
(7B.35)

This can be integrated to yield
\[
\frac{\sigma^2}{4} = e^\sigma - 1 - \sigma,
\]  
(7B.36)
or
\[
x = \frac{1}{2} \int_{-1}^{\sigma} \frac{d\zeta'}{\sqrt{e^\zeta - 1 - \zeta}}.
\]  
(7B.37)

From this we see that for \( x \to -\infty \), \( \sigma \) goes to zero like \( e^{x/\sqrt{2}} \), so that \( \rho \sim \exp\left(\frac{x}{\sqrt{2}}/2\right) \to 1 \). For \( x \to -\infty \), there is superconducting order and no magnet field; for \( x \to \infty \) there is no order, \( \rho = 0 \), and the critical magnetic field \( H = H_c = 1/\sqrt{2} \). The important point about a domain wall for \( \kappa = 1/\sqrt{2} \) is that in an external magnetic field \( H^{\text{ext}} = 1/\sqrt{2} \), it can be formed without any cost in energy. In order to see this we calculate, in reduced units, the magnetic enthalpy (for any \( \kappa \))
\[
\hat{E}_H = \int d^3 x \hat{H}_{GL} - \int d^3 x \mathbf{H} \cdot \mathbf{H}^{\text{ext}}
\]  
\[= \frac{1}{2} \int d^3 x \left[ (\nabla \rho)^2 - \rho^2 + \frac{1}{2} \rho^4 + \left( \rho^2 A^2 + \mathbf{H}^2 \right) \right] - \int d^3 x \mathbf{H} \cdot \mathbf{H}^{\text{ext}}. \]  
(7B.38)

Inserting the field equations (7B.24) and subtracting off the condensation energy \( \hat{E}_c = -(1/4) \int d^3 x \), this can be rewritten as
\[
\hat{E}_H - \hat{E}_c = \int d^3 x \left[ (1 - \rho^4) + \frac{1}{2} \mathbf{H}^2 - \mathbf{H} \cdot \mathbf{H}^{\text{ext}} \right].
\]  
(7B.39)

This is the additional energy of a domain wall. At the critical field strength \( H^{\text{ext}} = H_c = 1/\sqrt{2} \) pointing in the \( y \)-direction it becomes
\[
\hat{E}_H - \hat{E}_c = \frac{1}{2} \int d^3 x \left[ -\frac{\rho^4}{2} + \left( H - \frac{1}{\sqrt{2}} \right)^2 \right].
\]  
(7B.40)

Inserting \( \kappa = 1/\sqrt{2} \) into Eq. (7B.33) we indeed obtain zero. For \( \kappa = 1/\sqrt{2} \), a domain wall costs no energy.
Assuming that the wall energy is a monotonic function of $\kappa$, we expect the regions $\kappa > 1/\sqrt{2}$ and $\kappa < 1/\sqrt{2}$ to have wall energies of the opposite sign. Indeed, a numerical discussion of the different equations confirms this expectation. The solutions to the field equations are shown in Fig. 7.6. Inserting them into (7B.40) shows that the energy $\hat{F}_H - \hat{F}_c$ is positive for $\kappa < 1/\sqrt{2}$ and negative for $\lambda < 1/\sqrt{2}$. Hence we can conclude that type-I superconductors prefer a uniform state, type-II superconductors a mixed state.

![Figure 7.6](image)

**Figure 7.6** Spatial variation of order parameter $\rho$ and magnetic field $H$ in the neighborhood of a planar domain wall between normal and superconducting phases $N$ and $S$, respectively. The magnetic field points parallel to the wall.

Actually, the planar domain walls calculated above are not the most energetically favorable way of forming a mixed state. A much better configuration is given by a bundle of magnetic vortex lines. In order to see this, let us study the properties of a solution corresponding to a single vortex line.

### 7B.3 Single Vortex Line and Critical Field $H_{c_1}$

In a type-II superconductor, the mixed state begins to form for much lower fields than the critical magnetic field $H_c = 1/\sqrt{2}$. The reason lies in the fact that there exists a solution in which only a very small magnetic flux invades into the superconductor, namely, the flux

$$\Phi_0 = \frac{ch}{q} = \frac{\pi c}{\bar{h}} \approx 2 \times 10^{-7} \text{ gauss} \cdot \text{cm}^2. \quad (7B.41)$$

This solution has the form of a vortex line. Such a vortex line may be considered as a line-like defect in this uniform superconducting state. In this respect, it is a relative of a vortex line in superfluid $^4\text{He}$. The two are, however, objects with quite different physical properties, as we shall now see.

Suppose the system is in the superconducting state without an external voltage so that there is no current $j$. Let us introduce a vortex line along the $z$-axis. Then we can use the current formula (7B.22) to find from it the vector potential,

$$\mathbf{A} = -\frac{j_z}{|\psi|^2} + \frac{1}{2i} \frac{1}{|\psi|^2} \psi^\dagger \vec{\nabla} \psi. \quad (7B.42)$$
Far away from the vortex line, the state is undisturbed, i.e., $j_s$ vanishes, and we have the relation

$$A = \frac{1}{2i|\psi|^2} \psi^\dagger \nabla \psi.$$  \hfill (7B.43)

In a polar decomposition, $\psi(x) = \rho(x)e^{i\theta(x)}$, the derivative of $\rho(x)$ cancels and $A_i(x)$ depends only on the phase of the order parameter,

$$A(x) = \nabla \theta(x).$$  \hfill (7B.44)

Here we can establish contact with the discussion in superfluid $^4$He. There the superflow velocity was proportional to the gradient of a phase angle variable $\theta$. The periodicity of $\theta$ led to the quantization rule that, when taking the integral over $d\theta(x)$ along a closed circuit around the vortex line it had to be an integer multiple of $2\pi$. The same rule now applies here:

$$\oint_B d\theta(x) = \oint_B dx \cdot \nabla \theta(x) = 2\pi n.$$  \hfill (7B.45)

By Stokes' theorem, this is equal to the magnetic flux through the area of the circuit [recall Eq. (7B.33)]

$$\Phi = \int_{S_B} dS \cdot H = \int_{S_B} dS \cdot (\nabla \times A) = \kappa \int_B dx \cdot A = 2\pi nk.$$  \hfill (7B.46)

This holds in natural units. The quantization condition in physical units follows by applying the same argument to the original current (5.141) associated with the energy (5.140), which leads to

$$\Phi = n\Phi_0,$$  \hfill (7B.47)

with $\Phi_0$ given by Eq. (7B.41) [1].

Note that when performing the integral along a circle close to the vortex axis, the angular integral $\oint dx \cdot \nabla \theta$ still remains quantized, equal to $2\pi n$. But there the current no longer vanishes, and we find the quantization rule

$$\oint_B dx \cdot \left( A + \frac{j_s}{|\psi|^2} \right) = 2\pi n,$$  \hfill (7B.48)

or

$$\Phi = -\frac{1}{|\psi|^2} \oint_B dx \cdot j_s + 2\pi nk.$$  \hfill (7B.49)

This shows that through the smaller circuit there is less flux, part of the $2\pi nk$ is compensated by the magnetic field of the supercurrent flowing around the vortex line.
Quantitatively, we can deduce the properties of a vortex line by solving the field equations (7B.24), (7B.25) in cylindrical coordinates. Inserting the second into the first equation, we find

$$\frac{-1}{r} \frac{d}{dr} \frac{d\rho}{dr} + \frac{\kappa^2}{\rho^2} \left( \frac{d}{dr} H \right)^2 - (1 - \rho^2)\rho = 0. \tag{7B.50}$$

Forming the curl of the second gives the cylindrical analogue of (7B.30), i.e.,

$$H = \kappa^2 \frac{1}{r} \frac{f}{\rho^2} \frac{d}{dr} H. \tag{7B.51}$$

For $r \to \infty$ we have the boundary condition $\rho = 1$, $H = 0$ (superconducting state with Meissner effect) and $j_s = 0$ (no current). Since for stationary supercurrents, Ampère’s law (1.181) tells us that $j_s \propto \nabla \times H$, the last condition amounts to

$$H'(r) = 0, \quad r \to \infty. \tag{7B.52}$$

In cylindrical coordinates, flux quantization can be written in the form

$$\Phi = 2\pi \int_0^\infty dr r H = 2\pi n\kappa. \tag{7B.53}$$

Inserting Eq. (7B.51) into this result gives

$$\Phi = 2\pi \kappa^2 \left[ \frac{r}{\rho^2} H' \right]_0^\infty = -2\pi \kappa^2 \left[ \frac{r}{\rho^2} H' \right]_{r=0}, \tag{7B.54}$$

so that the quantization condition turns into a boundary condition at the origin:

$$H' \to -\rho^2 \frac{n}{\kappa r}, \quad r \to 0. \tag{7B.55}$$

Inserting this condition into (7B.50) we see that close to the origin $\rho(r)$ satisfies the equation

$$\frac{-1}{r} \frac{d}{dr} \frac{d}{dr} \rho(r) + \frac{n^2}{r^2} \rho - (1 - \rho^2)\rho \sim 0, \tag{7B.56}$$

which amounts to a behavior

$$\rho(r) = c_n \left( \frac{r}{\kappa} \right)^n + O(r^{n+1}). \tag{7B.57}$$

Putting this back into (7B.55) we have

$$H(r) = H(0) - \frac{c_n^2}{2\kappa} \left( \frac{r}{\kappa} \right)^{2n} + O(r^{2n+1}). \tag{7B.58}$$
For large $r$, where $\rho \to 1$, Eq. (7B.51) is solved by the modified Bessel function $K_0$, with some factor $\alpha$, namely

$$H(r) \to \alpha K_0 \left( \frac{r}{\kappa} \right), \quad r \to \infty.$$  (7B.59)

For large $\kappa \gg 1/\sqrt{2}$ (i.e. deep type-II) $\rho$ goes rapidly to 1 as compared to the length scale over which $H$ changes (which is $\kappa$). Therefore, the behavior (7B.59) holds very close to the origin. We can determine $\alpha$ by matching (7B.59) to (7B.55) and find (since $K_0' = K_1 \sim -1/r$)

$$\alpha \approx \frac{n}{\kappa}.$$  (7B.60)

In general, $H(r)$ and $\rho(r)$ have to be found numerically. A typical solution for $n = 1$ is shown in Fig. 7.7 for $\kappa = 10$. The energy of a vortex line can be calculated by using (7B.21). Inserting the equations of motion, and subtracting the condensation energy $\hat{E}_c = -(1/4) \int d^3x$ gives [compare (7B.39)]

$$\hat{E}_v = \hat{E}_H - \hat{E}_c = \frac{1}{2} \int d^3x \left[ \frac{1}{2} \left( 1 - \rho^4 \right) + H^2 \right].$$  (7B.61)

For $\kappa \gg 1/\sqrt{2}$, we may neglect the small radius $r \leq 1$, over which $\rho$ increases quickly from zero to its asymptotic value $\rho = 1$. Beyond $r \geq 1$ but for $r \leq \kappa$, $H$, is given by (7B.59). Inserting this into (7B.50) with (7B.57), we find

$$\rho(r) \sim 1 - \frac{n^2}{2r^2}.$$  (7B.62)

For the region $1 \leq r \leq \kappa$ this gives for the energy per length unit

$$\frac{1}{L} \hat{E}_v = \frac{1}{2} \pi^2 \int_1^\kappa dr \left[ \frac{1}{2} \left( 1 - \rho^4 \right) + H^2 \right] = \pi n^2 \int_1^\kappa dr \left[ \frac{1}{r^2} + \frac{1}{\kappa^2} K_0^2 \left( \frac{r}{\kappa} \right) \right].$$  (7B.63)

---

3For very large $r$, this has the limit $\sqrt{\pi \kappa/2} e^{-r/\kappa}$. 

H. Kleinert, GRAVITY WITH TORSION
Appendix 7B  Properties of Ginzburg-Landau Theory of Superconductivity

For $\kappa \to \infty$, the second integral goes toward a constant [since $\int_0^\infty dx x K_0^2(x) = \frac{1}{2}$]. The first integral, however, has a logarithmic divergence so that we find the energy of a vortex line

$$\frac{1}{L} \hat{E}_v \approx \pi n^2 [\log \kappa + \text{const.}]. \quad (7B.64)$$

A more careful estimate gives $\pi n^2 (\log \kappa + 0.08)$.

Let us now see at which external magnetic field such a vortex line can form. For this we consider again the magnetic enthalpy (7B.39) and subtract from $(1/L)\hat{E}_v$ the magnetic $\hat{E}_v/L$ coupling $HH^\text{ext}$ so that, per length unit,

$$\frac{1}{L} \hat{E}_H = \pi n^2 (\log \kappa + 0.08) - 2\pi \int_0^\infty dr r HH^\text{ext}. \quad (7B.65)$$

But the integral over $H$ is simply the flux quantum (7B.46) associated with the vortex line, i.e.,

$$\frac{1}{L} \hat{E} = \pi n^2 (\log \kappa + 0.08) - 2\pi n\kappa H^\text{ext}. \quad (7B.66)$$

When this drops below zero, a vortex line invades the superconductor along the $z$-axis. The associated critical magnetic field is

$$H_{c1} = \frac{n}{2\kappa} (\log \kappa + 0.08). \quad (7B.67)$$

For large $\kappa$ this field can be quite small. The asymptotic result is compared with

![Figure 7.8 Critical field $H_{c1}/n$ at which a vortex line of strength $n$ forms when it first invades a type-II superconductor as a function of the parameter $\kappa$. The dotted line indicates the asymptotic result $(1/2\kappa) \log \kappa$ of Eq. (7B.67). The magnetic field $H_{c1}$ is measured in units of $\sqrt{2}H_c$ where $H_c$ is the magnetic field at which the magnetic energy equals the condensation energy.](image)

a numerical solution of the differential equation for $n = 1, 2, 3, \ldots$ in Fig. 7.8. For a comparison with experiment one expresses this field in terms of the critical magnetic field $H_c = 1/\sqrt{2}$ and measures the ratio

$$\frac{H_{c1}}{H_c} = \frac{n}{\sqrt{2}\kappa} (\log \kappa + 0.08). \quad (7B.68)$$

As an example, pure lead is a type-I superconductor with $H_{c1} = H_c \approx 550$ gauss. An admixture of 15% Iridium or 30% Thallium brings $H_c$ up to 650 or 430, and $H_{c1}$ down to 250 or 145, respectively (see Table (7.1)).
Table 7.1 Critical magnetic fields in gauss for Pb and Nb with various impurities.

<table>
<thead>
<tr>
<th>material</th>
<th>$H_c$</th>
<th>$H_{c1}$</th>
<th>$H_{c2}$</th>
<th>$T_c/K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb</td>
<td>550</td>
<td>550</td>
<td>550</td>
<td>4.2</td>
</tr>
<tr>
<td>0.850 Pb, 0.150 Ir</td>
<td>650</td>
<td>250</td>
<td>3040</td>
<td>4.2</td>
</tr>
<tr>
<td>0.750 Pb, 0.250 In</td>
<td>570</td>
<td>200</td>
<td>3500</td>
<td>4.2</td>
</tr>
<tr>
<td>0.700 Pb, 0.300 Tl</td>
<td>430</td>
<td>145</td>
<td>2920</td>
<td>4.2</td>
</tr>
<tr>
<td>0.976 Pb, 0.042 Hg</td>
<td>580</td>
<td>340</td>
<td>1460</td>
<td>4.2</td>
</tr>
<tr>
<td>0.916 Pb, 0.088 Bi</td>
<td>675</td>
<td>245</td>
<td>3250</td>
<td>4.2</td>
</tr>
<tr>
<td>Nb</td>
<td>1608</td>
<td>1300</td>
<td>2680</td>
<td>4.2</td>
</tr>
<tr>
<td>0.5 Nb, 0.5 Ta</td>
<td>252</td>
<td>–</td>
<td>1470</td>
<td>5.6</td>
</tr>
</tbody>
</table>

7B.4 Critical Field $H_{c2}$ where Superconductivity is Destroyed

As the field increases above $H_{c1}$, more and more vortex lines invade the superconductor. For $H \sim H_c$, they form a hexagonal array as shown in (7.9). If the field increases even more, the superconducting regions separating the vortex lines become thinner and thinner until, finally, the whole material is filled with the magnetic field, and superconductivity is destroyed. The field where this happens is denoted by $H_{c2}$. Its value can be estimated quite simply following Abrikosov. He noticed that close to $H_{c2}$, the order parameter is so small that the nonlinear terms can be forgotten, and the Ginzburg-Landau equation reads

$$\left[ \left( \frac{1}{i} \nabla - A \right)^2 - 1 \right] \psi(x) = 0.$$  \hspace{1cm} (7B.69)

For $H$ along the $z$ direction one may choose

$$A(x) = \left( 0, \frac{1}{\kappa} Hx, 0 \right)$$ \hspace{1cm} (7B.70)

and the following equation emerges:

$$\left[ -\partial_x^2 - \left( \frac{1}{i} \partial_y - \frac{1}{\kappa} Hx \right)^2 - \partial_z^2 - 1 \right] \psi(x) = 0.$$ \hspace{1cm} (7B.71)
Appendix 7B Properties of Ginzburg-Landau Theory of Superconductivity

Figure 7.10 Temperature behavior of the critical magnetic fields of a type-II superconductor: $H_{c_1}$ (when the first vortex line invades), $H_{c_2}$ (when superconductivity is destroyed in comparison with the field) and $H_c$ (when the magnetic field energy is equal to the condensation energy).

Figure 7.11 Magnetization curve as a function of the external magnetic field $H^{\text{ext}}$. The dashed curve shows how a type-I superconductor would behave.

The lowest nontrivial eigenstate is

$$\psi(x) = \text{const.} \, e^{-\left(1/\kappa\right)H(x-y^2(H^2/2)e^{ipy}}}.$$  \hfill (7B.72)

For this solution to occur, the energy eigenvalue $H/\kappa - 1$ must be negative. This happens for $H < H_{c_2} = \kappa$. In terms of the critical field $H_c$, $H_{c_2}$ is equal to $H_{c_2} = \sqrt{2\kappa H_c}$, i.e., it can be sizeably larger than $H_c$. As an example, pure lead has $H_{c_2} = H_c = 550$ gauss. The admixture of 15% Indium or 30% Thallium which changes $H_c$ to 650 or 430, increases $H_{c_2}$ to 3040 or 2920 gauss, respectively. The typical behavior of the critical fields $H_c$, $H_{c_1}$, $H_{c_2}$ as a function of $T$ is shown in Fig. 7.10.

The invasion of vortex lines becomes apparent from the curve depicted in Fig. 7.11 which shows the behavior of the magnetization curve as a function of $H^{\text{ext}}$

$$-M = H^{\text{ext}} - H$$  \hfill (7B.73)

in a type-II superconductor as compared with a type-I superconductor.
Notes and References

    H. London, Proc. R. Soc. A 155, 102 (1936);


[5] The classical magnetohydrodynamic equations are discussed in
Relativistic Magnetic Monopoles and Electric Charge Confinement

The theory of multivalued fields in magnetism in Chapter 4 can easily be extended to a full relativistic theory of charges and monopoles [3, 4]. For this we go over to four spacetime dimensions, which are assumed to be Euclidean with a fourth spatial component $dx^4 = icdt$, to avoid factors of $i$.

8.1 Monopole Gauge Invariance

The covariant version of the Maxwell equation (4.53) reads

$$\frac{1}{2} \epsilon_{abcd} \partial_a F_{cd} = -\frac{1}{c} \tilde{j}_b,$$

where $\tilde{j}_a$ is the magnetic current density

$$\tilde{j}_a = (c \rho_m, j_m).$$

Equation (8.1) implies that the magnetic current density is conserved:

$$\partial_a \tilde{j}_a = 0.$$  

The zeroth component of (8.1) reproduces Eq. (4.53) for the monopole charge density [recall the identification of the field components (1.167)], and the spatial components yield the modified Faraday law (1.18):

$$\nabla \times E + \frac{1}{c} \frac{\partial B}{\partial t} = -\frac{1}{c} j_m \quad \text{(modified Faraday law)}. $$

For a single monopole of strength $g$ moving along a world line $q_a(\sigma)$, the magnetic current density $\tilde{j}_a$ can be expressed in terms of a $\delta$-function on the world line,

$$\delta_a(x; L) \equiv \int d\sigma \frac{d\bar{x}_a(\sigma)}{d\sigma} \delta^{(4)}(x - \bar{x}(\sigma)).$$
as follows
\[ \tilde{j}_a = g \, c \, \delta_a(x; L). \]  
(8.6)

This satisfies the conservation law (8.3) as a consequence of the identity
\[ \partial_a \delta_a(x; L) = 0, \]  
(8.7)

the four-dimensional version of (4.10) applied to closed worldlines. The spacetime components of the magnetic current density are [compare (1.205) and (1.206)]
\[ c \rho_m(x, t) = g \, c \int_{-\infty}^{\infty} d\tau \gamma c \, \delta^{(4)}(x - \bar{x}(\tau)), \]  
(8.8)
\[ \mathbf{j}_m(x, t) = g \, c \int_{-\infty}^{\infty} d\tau \gamma \mathbf{v} \, \delta^{(4)}(x - \bar{x}(\tau)). \]  
(8.9)

Note that with this notation, the electric current density (1.207) of a particle on the worldline \( L \) reads
\[ j_a = e \, c \, \delta_a(x; L), \]  
(8.10)

with the same conservation law \( \partial_a j = 0. \)

Equation (8.1) shows that \( F_{ab} \) cannot be represented as a curl of a single-valued vector potential \( A_a \), since left-hand side is equal to \( \epsilon_{abcd} (\partial_a \partial_c - \partial_c \partial_a) A_d)/2 \) implying a violation of Schwarz’s integrability condition. As in the magnetostatic discussion in Section 4.6 the simplest way to incorporate the monopole worldline into the electromagnetic field theory is via an extra monopole gauge field. In four spacetime dimensions, this is defined by
\[ F_{Mab} \equiv g \, \tilde{\delta}_{ab}(x; S), \]  
(8.11)

where \( \tilde{\delta}_{ab}(x; S) \) is the dual tensor
\[ \tilde{\delta}_{ab}(x; S) \equiv \frac{1}{2} \epsilon_{abcd} \delta_{cd}(x; S), \]  
(8.12)
of the \( \delta \)-function \( \delta_{cd}(x; S) \) which is singular on the world surface \( S \):
\[ \delta_{ab}(x; S) \equiv \int d\sigma d\tau \left[ \frac{\partial \bar{x}_a(\sigma, \tau)}{\partial \sigma} \frac{d\bar{x}_b(\sigma, \tau)}{\partial \tau} - (a \leftrightarrow b) \right] \delta^{(4)}(x - \bar{x}(\sigma, \tau)). \]  
(8.13)

This \( \delta \)-function has the obvious property
\[ \partial_a \delta_{ab}(x; S) = a \delta_b(x; L), \]  
(8.14)

where \( L \) is the boundary line of the surface. This follows directly from the simple calculation:
\[ \partial_a \delta_{ab}(x; S) = \int d\tau \left[ \frac{d\bar{x}_b(\sigma_b, \tau)}{\partial \tau} \delta^{(4)}(x - \bar{x}(\sigma_b, \tau)) - \frac{d\bar{x}_b(\sigma_a, \tau)}{\partial \tau} \delta^{(4)}(x - \bar{x}(\sigma_a, \tau)) \right] \]
\[ - \int d\sigma \left[ \frac{d\bar{x}_b(\sigma_b, \tau)}{\partial \tau} \delta^{(4)}(x - \bar{x}(\sigma_b, \tau)) + \frac{d\bar{x}_b(\sigma_a, \tau)}{\partial \tau} \delta^{(4)}(x - \bar{x}(\sigma_a, \tau)) \right]. \]
where \( \sigma_{a,b} \) and \( \tau_{a,b} \) are the lower and upper values of the surface parameters, respectively, so that \( x(\sigma_a, \tau), \bar{x}(\sigma, \tau_a), \bar{x}(\sigma_b, \tau), \bar{x}(\sigma, \tau_b) \), run along the boundary line of the surface. The dual \( \delta \)-function (8.12) satisfies

\[
\frac{1}{2} \epsilon_{abcd} \partial_b \tilde{\delta}_{cd}(x; S) = \delta_a(x; L),
\]

due to identity (1A.24). Equation (8.15) is the four-dimensional version of the local formulation (4.23) of Stokes’ theorem.

For the monopole gauge field (8.11) this implies

\[
\frac{1}{2} \epsilon_{abcd} \partial_a F^M_{cd} = \frac{1}{c} \tilde{j}_b.
\]

The surface \( S \) is the worldsheet of the Dirac string. For any given line \( L \), there are many possible surfaces \( S \). We can go over from one \( S \) to another, say \( S' \), at fixed boundary \( L \) as follows

\[
\tilde{\delta}_{cd}(x; S) \rightarrow \tilde{\delta}_{cd}(x; S') = \tilde{\delta}_{cd}(x; S) + \partial_a \tilde{\delta}_b(x; V) - \partial_b \tilde{\delta}_a(x; V),
\]

where \( \tilde{\delta}_a(x; V) \) is the \( \delta \)-function (6.13) which is singular on the three-dimensional volume \( V \) in four-space swept out when the surface \( S \) moves through four-space. The transformation law (8.17) is the obvious generalization of (4.28).

Many monopoles are, of course, represented by a gauge field (8.11) with a superposition of many different surfaces \( S \).

We are now ready to set up the electromagnetic action in the presence of an arbitrary number of monopoles. By analogy with Eq. (4.85) and (5.22) it depends only on the difference between the total field strength \( F_{ab} = \partial_a A_b - \partial_b A_a \) of the integrable vector potential \( A_a \) and the monopole gauge field \( F^M_{ab} \) of (8.11), i.e., it is given by [5, 6, 7, 8]

\[
A_0 + A_{mg} \equiv A_{0,mg} = \int d^4x \frac{1}{4c} \left( F_{ab} - F^M_{ab} \right)^2.
\]

The subtraction of \( F^M_{ab} \) is essential in avoiding an infinite energy density in the Maxwell action

\[
A_0 \equiv \int d^4x \frac{1}{4c} F^2_{ab},
\]

that would arise from the flux tube in \( F_{ab} \) inside the Dirac string. The difference

\[
F^\text{obs}_{ab} \equiv F_{ab} - F^M_{ab}
\]

is the nonsingular observable field strength. Since only fields with finite action are physical, the action contains no contributions from squares of \( \delta \)-functions as it might initially appear.

The action (8.18) exhibits two types of gauge invariances. First, the original electromagnetic one under [compare (2.103)]

\[
A_a(x) \rightarrow A'_a(x) = A_a(x) + \partial_a \Lambda(x),
\]
where $\Lambda(x)$ is any smooth field which satisfies the integrability condition

$$\left( \partial_a \partial_b - \partial_b \partial_a \right) \Lambda(x) = 0,$$

(8.22)

under which $F^M_{ab}$ is trivially invariant. Second, there is gauge invariance under monopole gauge transformations

$$F^M_{ab} \to F^M_{ab} + \partial_a \Lambda^M_b - \partial_b \Lambda^M_a,$$

(8.23)

with integrable vector functions $\Lambda^M_a(x)$, which by (8.17) have the general form

$$\Lambda^M_a(x) = g \delta_a(x; V),$$

(8.24)

with arbitrary choices of three-volumes $V$. If the monopole gauge field (8.11) contains many jumping surfaces $S$, the function $\Lambda^M_a(x)$ will contain a superposition many volumes $V$.

To have invariance of the action (8.18), the transformation (8.23) must be accompanied by a shift in the electromagnetic gauge field [5, 6, 7, 8]

$$A_a \to A_a + \Lambda^M_a.$$  

(8.25)

From Eqs. (8.11), (8.15), and (8.17) we see that the physical significance of the part (8.23) of the monopole gauge transformation is to change the Dirac world surface without changing its boundary, the monopole world line. An exception are vortex gauge transformations (8.25) of the gradient type, in which $\Lambda^M_a$ is $g$ times the gradient $\partial_a \delta$-function on the four-volume $V_4$:

$$\delta(x; V_4) \equiv \epsilon_{abcd} \int d\sigma d\tau d\lambda d\kappa \frac{\partial \bar{x}_a}{\partial \sigma} \frac{\partial \bar{x}_b}{\partial \tau} \frac{\partial \bar{x}_c}{\partial \lambda} \frac{\partial \bar{x}_d}{\partial \kappa} \delta^{(4)}(x - \bar{x}(\sigma, \tau, \lambda, \kappa)),$$

(8.26)

i.e.,

$$A_a \to A_a + g \partial_a \delta(x; V_4).$$

(8.27)

These do not give any change in $F^M_{ab}$ since they are particular forms of the original electromagnetic gauge transformations (8.21).

The field strength $F_{ab}$ is, of course, changed by moving the Dirac string through space, only the observable field strength $F^{\text{obs}}_{ab} = F_{ab} - F^M_{ab}$ remains invariant.

The part (8.25) of the monopole gauge transformations expresses the fact that in the presence of monopoles the gauge field $A_a$ is necessarily a cyclic variable for which $A_a(x)$ and $A_a(x) + gn$ are identical at each point $x$ for any integer $n$.

The partition function of magnetic monopoles and their electromagnetic interactions is given by the functional integral

$$Z = \int \mathcal{D}A_a^T \int \mathcal{D}F^M_{ab} e^{-A_0 - mg}.$$  

(8.28)

Here we have used the same short notation for the measure as in (5.265), indicating the gauge-fixed sum $\sum_{\{S\}} \Phi[F^M_{ab}]$ over fluctuating jumping surfaces $S$, the world sheet of the Dirac strings, by the symbol $\int \mathcal{D}F^M_{ab}$, and indicate the gauge fixed functional integral over $A_a$ in the Lorentz gauge $\partial_a A_a = 0$ by the symbol $\int \mathcal{D}A_a^T$. 

H. Kleinert, GRAVITY WITH TORSION.
8.2 Charge Quantization

Let us now introduce electrically charged particles into the action (8.18). This is done via the current interaction

\[ A_{el} = \frac{i}{c^2} \int d^4x j_a(x)A_a(x), \]  

(8.29)

where \( j_a(x) \) is the electric current of the world line of a charged particle

\[ j_a = e \delta_a(x; L). \]  

(8.30)

This is the Euclidean version of the current interaction (2.83) in Minkowski space-time.

Due to ordinary current conservation

\[ \partial_a j_a = 0, \]  

(8.31)

the action (8.29) is trivially invariant under electromagnetic gauge transformations (8.21). In contrast, it can remain invariant under monopole gauge transformations (8.23), (8.25) only if the monopole charge satisfies the famous Dirac quantization condition derived before in Eq. (4.108). Let us see how this comes about in the present four-dimensional theory.

Under the monopole gauge transformation (8.25), only the part \( A_0 + A_{mg} \) of the total action

\[ A_{tot} \equiv A_0 + A_{mg} + A_{el} = \int d^4x \frac{1}{4c} \left( F_{ab} - F_{ab}^M \right)^2 + \frac{i}{c^2} \int d^4x j_a(x)A_a(x) \]  

(8.32)

is manifestly invariant. The electric part \( A_{el} \), and thus the total action, changes by

\[ \Delta A_{tot} = \Delta A_{el} = i \frac{eg}{c} I, \]  

(8.33)

where \( I \) denotes the integral

\[ I \equiv \int d^4x \delta_a(L)\delta_a(V). \]  

(8.34)

This is an integer number if \( L \) passes through \( V \) and zero if it misses \( V \). In the former case, the string in the operation (8.17) sweeps across \( L \), in the other case it does not. To prove this we let \( L \) run along the first axis and let \( V \) be the entire volume in 234-subspace. Then \( \delta_a(x; L) \) and \( \delta_a(x; V) \) have nonzero components only in the 1-direction; \( \delta_1(x; L) = \delta(x_2)\delta(x_3)\delta(x_4) \), and \( \delta_1(x; V) = \delta(x_1) \). Inserting these into the integral (8.17) yields \( I = 1 \).

The Dirac charge quantization condition follows now from the rules of quantum mechanics that all amplitudes are found from a functional integral over \( e^{i\mathcal{A}/\hbar} \) [recall (4.109)]. Hence physics is invariant under jumps of the action by \( 2\pi\hbar \times \) integer since
these do not contribute to any quantum-mechanical amplitude. From (8.49) this implies that
\[ \frac{eg}{\hbar c} = 2\pi \times \text{integer}. \] (8.35)

The result may be stated in a dimensionless way by expressing \( e \) in terms of the fine-structure constant \( \alpha \approx \frac{1}{137.0359} \ldots \) of Eq. (1.189) as \( e^2 = 4\pi \hbar c \alpha \), so that the charge quantization condition becomes\(^1\)
\[ \frac{g}{e} = \frac{1}{2\alpha}. \] (8.36)

It must be emphasized that the above derivation of (8.36) requires much less quantum mechanical input than most derivations found in the literature which involve the wave functions for the charged particle in a monopole field \([3, 4, 9]\). In the above derivation, however, the particle orbits remain fixed, and only the worldsheets of the Dirac strings are moved around by monopole gauge transformation and the quantization follows from the requirement of invariance under these transformations.

Observe that after the quantization of the charge, the total action (8.32) is double-gauge invariant—it is invariant under the ordinary electromagnetic gauge transformations (8.21) and the monopole gauge transformations (8.23).

### 8.3 Electric and Magnetic Current-Current Interactions

If we integrate out the \( A_a \)-field in the partition function associated with the action (8.63) we obtain the interaction
\[
A_{\text{int}} = \int d^4x \left\{ \frac{1}{4e} \left[ (F^{M}_{ab})^2 + 2\partial_a F^{M}_{ab} (-\partial^{2})^{-1} \partial_c F^{M}_{cb} \right] \right. \\
+ \left. \frac{1}{2e} j_a (-\partial^{2})^{-1} j_a + \frac{i}{2e^2} \partial_a F^{M}_{ab} (-\partial^{2})^{-1} j_b \right\}. \] (8.37)

The second term
\[ A_{jj} = \frac{1}{2e^3} \int d^4x j_a (-\partial^{2})^{-1} j_a \] (8.38)
is the usual electric current-current interaction, where \((-\partial^{2})^{-1}\) denotes the Euclidean version of \textit{retarded Green function} of the vector potential \( A^a(x) \)
\[ (-\partial^{2})^{-1}(x, x') = G(x - x') \equiv \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik(x-x')}}{k^2}. \] (8.39)

Indeed, inserting the components of the four-component current density \( j^a = (c\rho, j) \) [recall (1.191)], the interaction (8.38) reads
\[ A_{jj} = \frac{1}{2c} \int d^4x d^4x' \rho(x) G(x, x') \rho(x') + \frac{1}{2e^3} \int d^4x d^4x' j(x) G(x, x') j(x'). \] (8.40)

\(^1\)In many textbooks, the action (8.18) has a prefactor \( 1/4\pi \), leading to Dirac’s charge quantization condition in the form \( 2eg/\hbar c = \text{integer} \). In these conventions, \( e^2 = \hbar c \alpha \), so that the condition (8.36) is the same.
For the static charges and currents in Minkowski spacetime, this becomes [compare (4.97)]

\[ A_{jj} = \frac{1}{2} \int dt d^3 x' \rho(t, x') \frac{1}{|x - x'|} \rho(t, x) - \frac{1}{2c^2} \int dt d^3 x' j(t, x') \frac{1}{|x - x'|} j(t, x'). \]  
(8.41)

The first term is the **Coulomb interaction**, the second the **Biot-Savart interaction** of an arbitrary current distribution.

The first two terms in the interaction (8.37) reduce to the magnetic current-current interaction [compare (4.99)]

\[ \tilde{A}_{jj} = \frac{1}{2c^3} \int d^4 x \tilde{j}_a (-\partial^2)^{-1} \tilde{j}_a. \]  
(8.42)

This follows from (8.1) and the simple calculation with the help of the tensor identity (1A.23):

\[ \tilde{j}^2 = \left( \frac{c^2}{2} \epsilon_{abcd} \partial_b F^M_{cd} \right)^2 = c^2 \left[ \partial^2 (F^M_{cd})^2 - 2(\partial_a F^M_{ab})^2 \right]. \]  
(8.43)

The magnetic interaction (8.42) can be decomposed into time- and space-like components in the same way as in Eq. (8.40), but with magnetic and electric charge and current densities.

The last term in (8.37)

\[ A_{ji} = \int d^4 x \frac{i}{2c^2} \partial_a F^M_{ab} (-\partial^2)^{-1} j_b \]  
(8.44)

specifies the interaction between electric and magnetic currents. It is the relativistic version of the interaction (4.101).

All three interactions are invariant under monopole gauge transformations (8.23). For the electric and magnetic current-current interactions (8.38) and (8.42) this is immediately obvious since they depend only on the world lines of electric and magnetic charges. Only for the mixed interaction (8.44) is the invariance not obvious. In fact, performing a monopole gauge transformation (8.23), and using the world line representation (8.10) of the electric four-vector current, this interaction is changed by

\[ \Delta A_{jj} = i \int d^4 x \frac{g}{c^2} \partial^2 \delta_b(x; V) (-\partial^2)^{-1} j_b = i \frac{ge}{c} \int d^4 x \delta_b(x; V) \delta(x; L) = \frac{ge}{c} I. \]  
(8.45)

This is nonzero, but the theory is still invariant since \((ge/c)I\) is equal to \(2\pi\hbar\) times a number which is integer due to Dirac’s charge quantization condition (8.35). Thus \(e^{-\Delta A_{jj}/\hbar}\) is equal to one and the theory invariant. The reader will recognize the analogy with the three-dimensional situation in the mixed interaction Eq. (4.101).
8.4 Dual Gauge Field Representation

It is instructive to subject the total action (8.32) to a similar duality transformation as to the Hamiltonian (4.85) by which we derived the dual gauge formulation (4.90). Thus we introduce an independent fluctuating field $f_{ab}$ and replace the action (8.18) by the equivalent one [the four-dimensional analog of (4.86)]

$$\tilde{A}_{0,mg} = \int d^4x \left[ \frac{1}{4c} f_{ab}^2 + \frac{i}{2c} f_{ab} \left( F_{ab} - F^M_{ab} \right) \right],$$

(8.46)

with the two independent fields $A_a$ and $f_{ab}$. Inserting $F_{ab} \equiv \partial_a A_b - \partial_b A_a$, the partition function (8.28) becomes

$$Z = \int D A^a_T \int D f_{ab} \int D F^M_{ab} e^{-\tilde{A}_{0,mg}}.$$  

(8.47)

Here we may integrate out the vector potential $A_a$ to obtain the constraint

$$\partial_b f_{ab} = 0.$$  

(8.48)

This can be satisfied identically (as a Bianchi identity) by introducing a dual magnetoelectric vector potential $\tilde{A}_a$ and writing [compare (4.88)]

$$f_{ab} = \epsilon_{abcd} \partial_c \tilde{A}_d.$$  

(8.49)

If we also introduce a dual field tensor

$$\tilde{F}_{ab} \equiv \partial_a \tilde{A}_b - \partial_b \tilde{A}_a,$$  

(8.50)

the action (8.46) takes the dual form

$$\tilde{A}_{0,mg} \equiv \tilde{A}_0 + \tilde{A}_{mg} = \int d^4x \left( \frac{1}{4c} \tilde{F}_{ab}^2 + \frac{i}{c} \tilde{A}_{ab} \tilde{j}_a \right),$$  

(8.51)

with the magnetoelectric source

$$\tilde{j}_a \equiv \frac{c}{2} \epsilon_{abcd} \partial_b F^M_{cd}.$$  

(8.52)

By inserting (8.11) and (8.15), we see that $\tilde{j}_a$ is the magnetic current density (8.6). Since (8.52) satisfied trivially the current conservation law $\partial_a \tilde{j}_a = 0$ [recall (8.3)], the action (8.51) allows for an additional set of gauge transformations which are the magnetoelectric ones

$$\tilde{A}_a \to \tilde{A}_a + \partial_a \tilde{\Lambda},$$  

(8.53)

with arbitrary integrable functions $\tilde{\Lambda}$,

$$(\partial_a \partial_b - \partial_b \partial_a) \tilde{\Lambda} = 0.$$  

(8.54)
If we include the electric current (8.30) into the dual form of the action (8.51) it becomes
\[ \tilde{A}_{\text{tot}} = \int d^4x \left[ \frac{1}{4c} f^{ab}_2 + i \frac{e}{2c} j_a (F_{ab} - F_{ab}^M) + i \frac{e}{c^2} j_a A_a \right]. \] (8.55)
Extremizing this with respect to the field \( A_a \) gives now
\[ \partial_a f_{ab} = -\frac{1}{c} j_b, \] (8.56)
rather than (8.48). The solution of this requires the introduction of a gauge field analog to (8.11), the \textit{charge} gauge field
\[ \tilde{F}^E_{ab} = e \tilde{\delta}_{ab}(x; S'). \] (8.57)
Then (8.56) is solved by
\[ f_{ab} \equiv \frac{1}{2} \epsilon_{abcd} (\tilde{F}_{ab} - \tilde{F}^E_{ab}). \] (8.58)
The identity (8.15) ensures (8.56).
When inserting (8.58) into (8.55), we find
\[ \tilde{A}_{\text{tot}} = \int d^4x \left[ \frac{1}{4c} (\tilde{F}_{ab} - \tilde{F}^E_{ab})^2 - i \frac{e}{4c} \tilde{F}_{ab} \epsilon_{abcd} \tilde{F}^M_{cd} + i \frac{e}{4c} \tilde{F}^E_{ab} \epsilon_{abcd} \tilde{F}^M_{cd} \right]. \] (8.59)
Integrating the second term by parts and using Eq. (8.16) we obtain to
\[ \int d^4x \left[ \frac{1}{4c} (\tilde{F}_{ab} - \tilde{F}^E_{ab})^2 + i \frac{e}{c^2} \tilde{A}_a j_a \right] + \Delta A, \] (8.60)
where
\[ \Delta A = i \frac{e}{4c} \int d^4x \tilde{F}^E_{ab} \epsilon_{abcd} \tilde{F}^M_{cd}. \] (8.61)
Remembering Eqs. (8.11) and (8.57), this can be shown to be an integer multiple of \( eg/c \):
\[ \Delta A = eg \frac{i}{4c} \int d^4x \tilde{\delta}_{ab}(x; S) \epsilon_{abcd} \tilde{\delta}_{cd}(x; S') = i \frac{eg}{c} n, \quad n = \text{integer}. \] (8.62)
To see this we simply choose the surface \( S \) to be the 12-plane and \( S' \) to be the 34-plane. Then \( \tilde{\delta}_{12}(x; S) = -\tilde{\delta}_{21}(x; S) = \delta(x_1) \delta(x_2) \) and \( \tilde{\delta}_{34}(x; S') = -\tilde{\delta}_{43}(x; S') = \delta(x_3) \delta(x_4) \), and all other components vanish, so that \( \int d^4x \epsilon_{abcd} \tilde{\delta}_{ab}(x; S) \delta_{cd}(x; S') = 4 \int d^4x \delta(x_1) \delta(x_2) \delta(x_3) \delta(x_4) = 4 \). This proof can easily be generalized to arbitrary \( S, S' \) configurations.
We now impose Dirac’s quantization condition (8.35), which was required in (8.35) to guarantee the invariance under monopole gauge transformations (8.23) ensuring the invariance under the string deformations (8.17). This makes the phase factor \( e^{-\Delta A/h} \) equal to unity, so that it has no influence upon any quantum process.
The dual version of the total action (8.32) of monopoles and charges is therefore
\[ \tilde{A}_{\text{tot}} \equiv \tilde{A}_0 + \tilde{A}_c + \tilde{A}_m = \int d^4x \left[ \frac{1}{4c} (\tilde{F}_{ab} - \tilde{F}^E_{ab})^2 + i \frac{e}{c^2} \tilde{A}_a j_a \right]. \] (8.63)
It describes the same physics as the action (8.32). Here the magnetic monopole is coupled locally whereas the world line of the charged particle is represented by the charge gauge field (8.57). With the predominance of electric charges in nature, however, this dual action is only of academic interest.

Note that just as before the electromagnetic action with monopoles (8.32), also the dual magnetoelectric action (8.63) is double-gauge invariant after the Dirac quantization of the charge—it is invariant under the magnetoelectric gauge transformations (8.53) and the deformations of the surface $S$ monopole gauge transformations (8.23).

### 8.5 Monopole Gauge Fixing

First we should eliminate the superfluous monopole gauge transformation (8.27) with the special gauge functions $\Lambda_a^M = g \partial_a \Sigma V_4 \delta(x;V_4)$ which do not give any change in $F_{ab}^M$. They may be removed from $\Lambda_a^M$ by a gauge-fixing condition such as

$$n_a \Lambda_a^M \equiv 0,$$  \hspace{1cm} (8.64)

where $n_a$ is an arbitrary fixed unit vector.

The remaining monopole gauge freedom can be used to bring all Dirac strings to a standard shape so that $F_{ab}^M(x)$ becomes a function of only the boundary lines $L$. In fact, for any choice of the above unit vector $n_a$, we may always go to the axial monopole gauge defined by

$$n_a F_{ab}^M = 0.$$  \hspace{1cm} (8.65)

To see this we take $n_a$ along the 4-axis and consider the gauge fixing equations

$$F_4^i + \partial_4 \Lambda_4^M - \partial_i \Lambda_4^M = 0, \hspace{1cm} i = 1, 2, 3.$$  \hspace{1cm} (8.66)

With (8.64) we have $\Lambda_4^M \equiv 0$ and $\Lambda_a^M$ could certainly all be determined if they were ordinary real functions. It is nontrivial to show that the gauge (8.65) can be reached using only the restricted class of gauge functions of the form (8.24). This is seen most easily by approximating the four-space by a fine-grained hypercubic lattice of spacing $\epsilon$ and imagining $F_{ab}^M$ to be functions defined on the plaquettes. Then the above defined $\delta$-functions (8.26), (8.24), (8.13), and (8.5) correspond to integer-valued functions on sites $\delta(x;V_4) = N(x)$, on links $\delta_a(x;V) = N_a/\epsilon$, on plaquettes $\delta_{ab}(x;S) = N_{ab}/\epsilon^2$, and on links $\delta_a(x;L) = N_a/\epsilon^3$, respectively, and the derivatives $\partial_a$ correspond to $1/\epsilon$ times lattice differences $\nabla_a$ across links. Thus $F_{ab}^M$ can be written as $g N_{ab}(x)/\epsilon^2$ with integer $N_{ab}(x)$. The gauge fixing in (8.66) with the restricted gauge functions amounts then to solving a set of integer-valued equations of the type

$$N_4^i + \nabla_4 N_4^M - \nabla_i N_4 = 0, \hspace{1cm} i = 1, 2, 3,$$  \hspace{1cm} (8.67)

with $N_4 \equiv 0$. This is always possible as has been shown with similar equations in Ref. [11].
Having fixed the gauge in this way we can solve Eq. (8.1) uniquely by the monopole gauge field
\[ F_{ab}^M = -2\epsilon_{abcd}n_c(n\partial)^{-1}\tilde{j}_d. \] (8.68)
With this, the interaction between electric and magnetic currents in the last term of (8.37) becomes
\[ \mathcal{A}_{\tilde{j}j} = \epsilon_{abcd}\int d^4x j^a (n\partial \partial^0)^{-1}n_b\partial_c j^d. \] (8.69)
This interaction can be found in textbooks [12].

### 8.6 Quantum Field Theory of Spinless Electric Charges

The full Euclidean quantum field theory of electrically and magnetically charged particles is obtained from the functional integral over the Boltzmann factors \( e^{-A_{\text{tot}}/\hbar} \) with the action (8.32). The functional integral has to be performed over all electromagnetic fields \( A_a \) and over all electric and magnetic world line configurations \( L \) and \( L' \). These, in turn, can be replaced by fluctuating disorder fields which account for grand-canonical ensembles of world lines [13]. This replacement, the Euclidean analog of second quantization, in many-body quantum mechanics, was explained in the last chapter.

Let us assume that only a few fixed worldlines \( L \) of monopoles are present. The electric world lines, on the other hand will be assumed to consist of a few fixed world lines \( L' \) plus a fluctuating grand-canonical ensembles of closed world lines \( L'' \). The latter are converted into a single complex field \( \psi_e \) whose Feynman diagrams are pictures of the lines \( L'' \) [14]. The technique was explained in Subsection 5.1.10 and corresponds to the second quantization of many-body quantum mechanics. In other words, we shall study the following partition function
\[ Z = \int DA_T e^{-A_{\text{tot}}} \int D\psi_e \int D\psi_e^* e^{-A_{\psi_e}}, \] (8.70)
with the field action of the fluctuating electric orbits
\[ A_{\psi_e} = \int d^4x \frac{1}{2} \left[ |D\psi_e|^2 + m^2 |\psi_e|^2 + \lambda |\psi_e|^4 \right], \] (8.71)
where \( D_a \) denotes the covariant derivative involving the gauge field \( A_a \):
\[ D_a \equiv \partial_a - ie^c A_a. \] (8.72)
When performing a perturbation expansion of this functional integral in powers of the coupling constant \( e \), the Feynman loop diagrams of the \( \psi_e \) field provide direct pictures of the different ways in which the fluctuating closed charged worldlines interact in the ensemble.
8.7 Theory of Magnetic Charge Confinement

The field action of fluctuating electric charges is the four-dimensional version of the Ginzburg-Landau Hamiltonian (5.140). We have learned in the previous chapter that this Hamiltonian allows for a phase transition as a function of the mass parameter $m^2$ in (8.71). There exists a critical value of $m^2$ where the system changes from ordered to disordered. At the mean-field level, the critical value is zero. For $m^2 > 0$, only a few vortex loops are excited. In this phase, the field has a vanishing expectation value $\langle \psi_e \rangle$. For $e < e_c$, on the other hand, the configurational entropy wins over the Boltzmann suppression and infinitely long vortex loops too small the $m^2$ is negative and the disorder field $\psi_e$ develops a nonzero expectation $\langle \psi_e \rangle$ whose absolute value is equal to $\sqrt{|m^2|/2\lambda}$. This is a condensed phase where the charge worldlines are infinitely long and prolific. The passage of $e$ through $e_c$ is a phase transition. From the derivative term $|D\psi_e|^2$, the vector field $A_a$ receives a mass term $(m^2 A^2/c) A_a^2$ with $m_A$ equal to $q^2 |m^2|/\lambda$. For very small $e \ll e_c$, the penetration depth $1/m_A$ of the vector potential is much larger than the coherence length $1/m$ of the disorder field and the system behaves like a superconductor of type II.

Between magnetic monopoles of opposite sign, the magnetic field lines are squeezed into the four-dimensional analogs of the Abrikosov flux tubes. Within the present functional integral, the initially irrelevant surfaces $S$ enclosed by the charge worldlines $L$ acquire, via the phase transition, an energy proportional to their area which removes the charge gauge invariance of the action. They become physical fluctuating objects and generate the linearly rising static potential between the charges, thus causing magnetic charge confinement.

The confinement mechanism is particularly simple to describe in the hydrodynamic or London limit. In this limit, the magnitude $|\psi_e|$ of the field is frozen so it can be replaced by a constant $|\psi_e|$ multiplied by a spacetime-dependent phase factor $e^{i\theta(x)}$, and the functional integral over $\psi_e$ and $\psi_e^*$ in (8.70) reduces to

$$\sum_{\{V\}} \int_{-\infty}^{\infty} D\theta \exp \left\{-\frac{m^2 c}{2q^2} \int d^4x \left[ \partial_a \theta - \theta_a^\nu(x) - \frac{e}{c} A_a \right]^2 \right\},$$

(8.73)

where $\theta_a^\nu(x)$ is the four-dimensional vortex gauge field

$$\theta_a^\nu(x) \equiv 2\pi \delta_a(x; V).$$

(8.74)

This may be chosen in a specific gauge, for instance in the axial gauge with $\delta_4(x; V) = 0$, so that $V$ is uniquely fixed by its surface $S$, the worldsheets of a vortex line in the $\phi_e$-field. Thus the action in (8.70) reads, in the hydrodynamic limit,

$$A_{hy} = \int d^4x \left\{ \frac{1}{4c} (F_{ab} - F^M_{ab})^2 + i \frac{e}{c} \int d^4x \ j_a(x) A_a(x) + \frac{m_A^2 c}{2q^2} \left[ \partial_a \theta - \theta_a^\nu(x) - \frac{e}{c} A_a \right]^2 \right\},$$

(8.75)

H. Kleinert, GRAVITY WITH TORSION
If we ignore the vortices and eliminate the \( \theta \)-fluctuations from the functional integral, we generate a transverse mass term

\[
\frac{m_A^2}{2c} A_T^2
\]  

(8.76)

where \( A_T^a \equiv (g_{ab} - \partial_a \partial_b / \partial^2) A_b \). This causes the celebrated Meissner effect in the superconductor. In the hydrodynamic limit the action becomes, therefore,

\[
A_{\text{hy}} = \int d^4 x \left[ \frac{1}{4c} (F_{ab} - F_{ab}^M)^2 + \frac{m_A^2}{2c} A_T^2 \right].
\]  

(8.77)

If we now integrate out the \( A_a \)-fields in the partition function (8.70), we obtain the interaction between the worldlines of electric charges \( L \) and the surfaces \( S \) whose boundaries are the monopole worldlines:

\[
A_{\text{hy}}^{\text{int}} = \int d^4 x \left\{ \frac{1}{16\pi} F_{ab}^M(x) G_{m_A}^M(x - x') F_{ab}^M(x') + \frac{1}{2c^2} \partial_a F_{ab}^M(\partial^2 + m_A^2)^{-1} j_b + \frac{1}{2c^2} j_a(\partial^2 + m_A^2)^{-1} j_a \right\}.
\]  

(8.78)

This is a generalization of the previous current-current interaction (8.37), to which it reduces for \( m_A = 0 \). Using the relation (8.43), this becomes

\[
A_{\text{hy}}^{\text{int}} = \int d^4 x \int d^4 x' \left[ \frac{m_A^2}{16\pi} F_{ab}^M(x) G_{m_A}^M(x - x') F_{ab}^M(x') + \frac{1}{2} \partial_a F_{ab}^M(x - x') j_a(x - x') + \frac{4\pi}{2} j_a(x) G_{m_A}^M(x - x') j_a(x') \right]
\]  

(8.79)

with the massive correlation function

\[
(-\partial^2)^{-1}(x, x') = G_{m_A}(x - x') \equiv \int \frac{d^4 k}{(2\pi)^4} e^{-ik(x-x')} \frac{1}{k^2 + m_A^2}.
\]  

(8.80)

Due to the mass \( m_A \), the interactions have changed. The last term in (8.79) is now a short-range Yukawa-type interaction between the electric charges.

The second term is a short-range interaction between the surfaces and the boundary lines.

The first term is most interesting. It gives an energy to the previously irrelevant surfaces \( S \) enclosed by the magnetic worldlines \( L \). The energy covers \( S \) and a neighborhood of it up to a distance \( 1/m_A \). It converts \( S \) into a thick worldsheet. This is the world surface of a thick flux tube of thickness \( 1/m_A \) between connecting the magnetic charges. To lowest order in the thickness, this causes a surface tension, giving rise to a linearly rising potential between magnetic charges, and thus to confinement. To next order, it causes a curvature stiffness [16].

The fact that the energy of the surface \( S \) enclosed by a monopole world line causes confinement can be phased as a criterion for confinement due to Wilson. In
the duality transformation of the monopole part of the action (8.32) to (8.51) we
have observed that a surface \( S \) in the monopole gauge field \( F_M^{ab} \) corresponds to a
local coupling \((i/c) \int d^4x \tilde{A}_a j_a\) in the dual action. This implies that the expectation
value of the exponential
\[
\langle \exp \left( \frac{i}{c} \oint L d^4x \tilde{A}_a \right) \rangle
\]
(8.81)
falls off like \( \exp (- \text{area enclosed by } L) \) in the confined phase where the interaction
is given by (8.79), but only like \( \exp (- \text{length of } L) \) in the unconfined phase where
the interaction is given by (8.37).

If the charged particles are electrons, the field \( \psi_e(x) \) must consist of four anti-
commuting Grassmann components and the action must be of the Dirac type which
has the form in Minkowski spacetime:
\[
A_{\text{Dirac}} = \int d^4x \left\{ \bar{\psi}_e(x) \left[ \gamma^a \left( i\hbar \partial_a - \frac{e}{c} A_a \right) \psi_e(x) - m_e c^2 \bar{\psi}_e(x) \psi_e(x) \right] \right\},
\]
(8.82)
where \( m_e \) is the mass of the electron and \( \psi_e(x) \) are the standard Dirac fields of the
electron. Fermi fields cannot form a condensate, so that there is no confinement of
monopoles, and the second quantization leads to the standard quantum field theory
of electromagnetism (QED) with the minimal electromagnetic interaction:
\[
A_{\text{el}} = i \frac{e^2}{c^2} \int d^4x A_a j_a, \quad j_a = e \bar{\psi} \gamma_a \psi.
\]
(8.83)

### 8.8 Second Quantization of the Monopole Field

For monopoles described by the action (8.71), second quantization seems at first
impossible since the partition function contains sum over a grand-canonical ensemble
of surfaces \( S \) rather than lines. Up to now, there exists no satisfactory second-
quantized field theory which could replace such a sum. According to present belief,
the vacuum fluctuations of some non-abelian gauge theory is able to do this, but a
convincing theoretical formulation is still missing.

Fortunately, the monopole gauge invariance of the action (8.63) under (8.23)
makes most configurations of the surfaces \( S' \) physically irrelevant and allows us to
return to a worldline description of the monopoles after all. We simply fix the gauge
as described above, which makes the monopole gauge field unique. It is given by
Eq. (8.68) and thus completely specified by the worldlines \( L' \) of the monopoles. With
this we can rewrite the action as [15]
\[
A = A_1 + A_{\text{el}} + A_{\text{c}1} + A_{\text{c}2}
= \int d^4x \left\{ \left[ \frac{1}{4c}(F_{ab} - f_{ab}^M)^2 \right] + i \frac{e}{c^2} A_a j_a \right\} + \frac{i}{c^2} \lambda_{ab} \left( n_\sigma \partial_\sigma f_{ab}^M + 2 \epsilon_{abcd} n_c j_d \right),
\]
(8.84)
where \( f_{ab}^M \) and \( \lambda_{ab} \) are now two arbitrary fluctuating fields, i.e., \( f_{ab}^M \) is no longer of
the restricted \( \delta \)-function type implied by (8.11). This form restriction is enforced by
the fluctuating $\lambda_{ab}$-field. The two terms in the action in which $\lambda_{ab}$ appears have been denoted by $A_{\lambda_{ab}1}$ and $A_{\lambda_{ab}2}$. The monopole worldline appears only in the magnetic current coupling

$$A_{mg} \equiv A_{\lambda_{ab}2} = \frac{i}{c^2} \int d^4x \tilde{A}_d^a \tilde{j}_d, \quad (8.85)$$

where $\tilde{A}^a_d$ is short for the vector field

$$\tilde{A}^a_d \equiv 2\lambda_{ab} \epsilon^{abcd} n_c. \quad (8.86)$$

In the partition function associated with this action we may now sum over a grand-canonical ensemble of monopole worldlines $L$ by converting it into a functional integral over a single fluctuating monopole field $\phi_g$ as in the derivation of (8.70). If monopoles carry no spin, this obviously replaces the sum over all fluctuating monopole world lines with the magnetic interaction (8.85) by a functional integral

$$\int \mathcal{D}\phi_g \mathcal{D}\phi^*_g e^{-A_g^n}, \quad (8.87)$$

where $A_g$ is the action of the complex monopole field

$$A_g^n = \int d^4x \frac{1}{2} \left[ |D_a^n \phi_g|^2 + m_g^2 |\phi_g|^2 + \lambda |\phi_g|^4 \right], \quad (8.88)$$

and $D_a^n$ the covariant derivative

$$D_a^n \equiv \partial_a - \frac{g}{c} \tilde{A}_a^n. \quad (8.89)$$

By allowing all fields $\psi_e, \phi_g, A_a, f_{ab}^M$, and $\lambda_{ab}$, to fluctuate with a Euclidean amplitude $e^{-A/\hbar}$ we obtain the desired quantum field theory of electric charges and Dirac monopoles [17]. The total field action of charged spin-1/2 particles and spin-zero monopoles is therefore

$$\mathcal{A} = \int d^4x \left[ \frac{1}{4e} (F_{ab} - f_{ab}^M)^2 + \frac{i}{c^2} \lambda_{ab} n_\sigma \partial_\sigma f_{ab}^M \right] + \mathcal{A}_{\text{Dirac}} + A_g^n. \quad (8.90)$$

Note that the effect of monopole gauge invariance is much more dramatic than that of the ordinary gauge invariance in pure QED. The electromagnetic gauge transformation $A_a \to A_a + \partial_a \Lambda$ eliminated only the longitudinal polarization of the photons. The monopole gauge transformations, in contrast, (8.23) reduce the dimensionality of the fluctuations from surfaces $S$ to lines $L$, which is crucial for setting up the disorder field theory (8.87).

It is obvious that there exists a dual formulation of this theory with the action

$$\mathcal{A} = \int d^4x \left[ \frac{1}{4e} (\tilde{F}_{ab} - \tilde{f}_{ab}^E)^2 + \frac{i}{c^2} \lambda_{ab} n_\sigma \partial_\sigma \tilde{f}_{ab}^E \right] + \mathcal{A}_g + \mathcal{A}_{\text{Dirac}}^n, \quad (8.91)$$
where $A_g$ is the action (8.87) with the covariant derivative

$$D_a = \partial_a - \frac{g}{c} \tilde{A}_a,$$

(8.92)

and $A_{e\text{Dirac}}$ is the Dirac action coupled minimally to the vector potential (8.86):

$$A_{e\text{Dirac}} = \int d^4x \left\{ \bar{\psi}_e(x) \left[ \gamma^a \left( i\hbar \partial_a - \frac{e}{c} A^a_e \right) \psi_e(x) - m_e c^2 \bar{\psi}_e(x) \psi_e(x) \right] \right\}.$$

(8.93)

### 8.9 Quantum Field Theory of Electric Charge Confinement

It has long been known that quantum electrodynamics on a lattice with a cyclic vector potential (called compact QED) shows quark confinement for a sufficiently strong electric charge $e$. The system contains a grand-canonical ensemble of magnetic monopoles which condense at some critical value $e_c$. The condensate squeezes the electric field lines emerging from any charge into a thin tube giving rise to a confining potential [18, 19, 20]. It is possible to transform the partition function to the dual version of a standard Higgs model coupled minimally to the dual vector potential $\tilde{A}_a$ [21]. The Higgs field is the disorder field [11] of the magnetic monopoles, i.e., its Feynman graphs are the direct pictures of the monopole worldlines in the ensemble. Two electric charges in this model are connected by Abrikosov vortices producing the linearly rising potential between the charges and thus confinement. The system is a perfect dielectric. While there is no problem in taking the dual Higgs field description of quark confinement to the continuum limit [21], the same thing has apparently never been done in the original formulation in terms of the gauge field $A_a$. The reason was a lack of an adequate continuum description of the integer-valued jumps in the electromagnetic gauge field $A_a$ across the worldsheets spanned by the worldlines of the magnetic monopoles. After the development of the previous sections we can easily construct a simple quantum field theory which exhibits electric charge confinement. It is based on a slight modification of the dual magnetoelectric action (8.63). The modification will lead to the formation of thin electric flux tubes between opposite electric charges.

For a fixed set of electric and magnetic charges, the Euclidean action reads

$$A = \frac{1}{16\pi} \int d^4x \left[ F_{ab}(x) - F_{ab}^M(x) \right]^2 + i \int d^4x j_a(x) A_a(x),$$

(8.94)

where $F_{ab} = \partial_a A_b - \partial_b A_a$ is the usual field tensor,

$$j_a(x) = e \delta_a(x; L)$$

(8.95)

is the charge distribution along closed worldlines $L$ of the electric charges with $\delta_a(x; L)$ being $\delta$-functions singular on the lines $L$

$$\delta_a(x; L) = \int d\tau \frac{d\bar{x}_a}{d\tau} \delta^{(4)}(x - \bar{x}(\tau)),$$

(8.96)
8.9 Quantum Field Theory of Electric Charge Confinement

while \( F_{ab}^M(x) \) is the \textit{gauge field of monopoles}. It is defined as follows: Let \( \tilde{L} \) be the worldline of a monopole and \( \tilde{S} \) an arbitrary surface enclosed by \( \tilde{L} \), then we take the \( \delta \)-function on this surface

\[
\delta_{ab}(x; \tilde{S}) = \int d\sigma d\tau \frac{\partial \bar{x}_a}{\partial \sigma} \frac{\partial \bar{x}_b}{\partial \tau} - (a \leftrightarrow b) \delta^{(4)}(x - q(\sigma, \tau))
\]

and define \( F_{ab}^M \) in terms of the dual of this

\[
F_{ab}^M(x) \equiv 4\pi g \frac{1}{2} \epsilon_{abcd} \delta_{cd}(x; \tilde{S}).
\]

This field has the property that its curl is singular on the boundary line \( \tilde{L} \)

\[
\frac{1}{2} \epsilon_{abcd} \partial_b F_{cd}^M(x) = 4\pi g \delta_a(x; \tilde{L}) = 4\pi \tilde{j}_a(x),
\]

this being a reformulation of Stokes’ integral theorem in terms of distributions. The constant \( g \) is the magnetic charge of the monopoles which is assumed to satisfy Dirac’s charge quantization condition (8.35). The Euclidean quantum partition function of the system is found by summing, in a functional integral, the Boltzmann factor \( e^{-A} \) over all field configurations \( A_a \), all line configurations \( L \) in \( j^n \), and all surface configurations \( \tilde{S} \) in \( F_{ab}^M \). It was pointed out in [6, 7, 8] that the action (8.94) is invariant under two types of gauge transformations, the ordinary \textit{electromagnetic gauge transformations}

\[
A_a \rightarrow A_a + \partial_a \Lambda
\]

and the completely independent \textit{monopole gauge transformations}

\[
A_a \rightarrow A_a + \Lambda_a^M
\]

\[
F_{ab}^M \rightarrow F_{ab}^M + \partial_a \Lambda_a^M - \partial_b \Lambda_a^M
\]

which involve an arbitrary superposition of \( \delta \)-functions on three-volumes \( \tilde{V} \)

\[
\Lambda_a^M(x) = 4\pi g \sum \delta_a(x; \tilde{V}),
\]

with

\[
\delta_a(x; \tilde{V}) \equiv \epsilon_{abcd} \int d\sigma d\tau d\lambda \frac{\partial \bar{x}_a}{\partial \sigma} \frac{\partial \bar{x}_c}{\partial \tau} \frac{\partial \bar{x}_d}{\partial \lambda} \delta^{(4)}(x - \bar{x}(\sigma, \tau, c)).
\]

The invariance of the gradient term in the action (8.94) is obvious. The current term, on the other hand, changes under the two gauge transformations by \( i \oint d^4 x j_a \partial_a \Lambda \) and by \( i \oint d^4 x j_a \Lambda_a^M \), respectively. The first change vanishes after a partial integration since for closed worldlines \( \partial_a \delta_a(x; L) = 0 \) ensuring the electric current conservation law \( \partial_a j_a = 0 \). The second change is irrelevant since \( \oint d^4 x \delta_a(x; L) \delta_a(x; \tilde{V}) \) is an integer, counting the number of times by which the line \( L \) pierces the volume \( \tilde{V} \).
The exponential $e^{-A}$ governing the fluctuations in the functional integral changes by $\exp(-i4\pi egn)$ which is a trivial unit factor due to (8.35). Certainly, the functional integrals over $A_a$ and the surfaces $\tilde{S}$ require gauge fixing to remove infinite degeneracies. The options for gauge fixing $A_a$ are well known; for $\tilde{S}$ one may fix the surface shapes in such a way that they are uniquely determined by their boundary lines $\tilde{L}$. This was the key for constructing a field theory of magnetic monopoles in [8]. It was also shown in [6, 7, 8] that a duality transformation brings $\tilde{A}$ to the completely equivalent form.

$$\tilde{A} = \frac{1}{16\pi} \int d^4x \left[ \tilde{F}_{ab}(x) - \tilde{F}_{ab}^E(x) \right]^2 + i \int d^4x \tilde{j}_a(x) \tilde{A}_a(x) \quad (8.105)$$

where $\tilde{F}_{ab} = \partial_a \tilde{A}_b - \partial_b \tilde{A}_a$ is the dual field tensor $\tilde{F}_{ab} = (1/2)\epsilon_{abcd}F_{ab}$ and $\tilde{j}_a(x) = g\delta_a(x; \tilde{L})$ the dual current density singular on the magnetic monopole worldlines $\tilde{L}$. Now the electric charges are described by a charge gauge field $\tilde{F}_{ab}^E$ which is singular on arbitrary worldsheets $S$ enclosed by the electric worldlines $L$: 

$$\tilde{F}_{ab}^E(x) \equiv 4\pi e \frac{1}{2}\epsilon_{abcd}\delta_{cd}(x; S). \quad (8.106)$$

This action is, of course, invariant under the magnetoelectric gauge transformations

$$\tilde{A}_a \rightarrow \tilde{A}_a + \partial_a \tilde{\Lambda} \quad (8.107)$$

and under the discrete-valued charge gauge transformations

$$\tilde{A}_a \rightarrow \tilde{A}_a + \tilde{\Lambda}_a^E, \quad \tilde{F}_{ab}^E \rightarrow \tilde{F}_{ab}^E + \partial_a \tilde{\Lambda}_b^E - \partial_b \tilde{\Lambda}_a^E. \quad (8.108)$$

We proceed as in the derivation of the disorder theory (8.70), but apply the transformation to the worldlines of the magnetic monopoles in the dual action (8.63). The resulting second-quantized action is

$$\tilde{A}_{\text{tot}} = \int d^4x \frac{1}{4c} (\tilde{F}_{ab} - \tilde{F}_{ab}^E)^2 + \int d^4x \left[ \left| \tilde{D}\phi_g \right|^2 + m^2 \left| \phi_g \right|^2 + \lambda \left| \phi_g \right|^4 \right], \quad (8.109)$$

where

$$\tilde{D}_a \equiv \partial_a - \frac{g}{c} \tilde{A}_a. \quad (8.110)$$

If we choose the mass parameter $m^2$ of the monopole field $\phi_g$ to be negative, then $\phi_g$ acquires a nonzero expectation value $\sqrt{-m^2/2\lambda}$, which generates a Meissner mass $m^2_{\tilde{A}}$ for the dual vector potential $\tilde{A}_a$. The energy has again the form (8.79), but with electric and magnetic sources exchanged:

$$\mathcal{A}_{\text{int}}^\text{by} = \int d^4x \int d^4x' \frac{m^2_{\tilde{A}}}{16\pi} F_{ab}(x)G_{m,\tilde{A}}(x-x')F_{ab}^E(x')$$

$$+ \frac{1}{2} \partial_a F_{ab}^E G_{m,\tilde{A}}(x-x')j_a(x-x') + 4\pi j_a(x)G_{m,\tilde{A}}(x-x')j_a(x') \quad (8.111)$$

H. Kleinert, GRAVITY WITH TORSION
This gives the surfaces $S'$ enclosed by the electric world lines an energy with the properties discussed above, causing now the confinement of electric charges. The thick energetic surface has tension and curvature stiffness as proposed independently by the author [22] and Polyakov [23] for world sheets of hadronic strings.

In the case of electric charge confinement, the expectation value of the dual of the Wilson loop (8.81)
\[
\langle \exp \left( \frac{i}{\hbar} \oint_L d^4x A_{\alpha} \tilde{\gamma}_\alpha \right) \rangle
\]
behaves like $\exp (-\text{area enclosed by } L)$.

It goes without saying that in order to apply the model to quarks, the action (8.82) has to be replaced by a Dirac action with three colors and six flavors in a gauge-invariant coupling
\[
A_D = \int d^4x \bar{\psi} (\not{D} - M) \psi,
\]
where $D_\mu + iG_\mu$ is a covariant derivative in color space, and $G_\mu$ a traceless $3 \times 3$-matrix color-electric gauge field with the field action
\[
A_{G_\mu} = -\frac{1}{4} \int d^4x \text{tr} \left( \partial_\mu G_\nu - \partial_\nu G_\mu - [G_\mu, G_\nu] \right)^2.
\]
The symbol $\mathcal{M}$ denotes a mass matrix in the six-dimensional flavor space of $u, d, c, s, t, b$.

If one applies the above model interaction (8.79) to quarks, one may study low-energy phenomena by approximating it roughly by a four-Fermi interaction. This can be converted into a chirally invariant effective action for pseudoscalar, scalar, vector, and axial-vector mesons by functional integral technique (hadronization) [24]. The effective action reproduces qualitatively many of the low-energy properties of these particles, in particular their chiral symmetry, its spontaneous breakdown, and the difference between the observed quark masses and the masses in the action (8.113) (current quark masses). It also explains why the quarks $u, d$ in a nucleon have approximately a third of a nucleon mass while their masses $M_u, M_d$ in the action (8.113) are very small.\(^2\)

The technique of hadronization developed in [24] has been generalized in various ways, in particular by including the color degree of freedom [25]. It has also been used to describe the low-lying baryons and the restoration of chiral symmetry by thermal effects [26].

An interesting aspect of (8.79) is that the local part of the four-Fermi interaction, which is proportional to $1/\tilde{m}_A^2$, arises by the same mechanism as the confining potential, whose tension is proportional to $\tilde{m}_A^2 \log(\Lambda^2/\tilde{m}_A^2)$, with $\Lambda$ being some ultraviolet cutoff parameter. One would therefore predict that at an increased temperature of the order of $\tilde{m}_A$ the spontaneous symmetry breakdown, which is caused by the

\(^2\)The quark masses in the action (8.113) are $M_u \approx 4 \text{ MeV}, M_d \approx 8 \text{ MeV}, M_c \approx 1.5 \text{ GeV}, M_s \approx 0.15 \text{ GeV}, M_t \approx 176 \text{ GeV}, M_b \approx 4.7 \text{ GeV}$. 
four-Fermi interaction, takes place at the same temperature at which the potential looses its deconfinement properties. This initially surprising coincidence has long been observed in Monte Carlo simulations of lattice gauge theories.

It is an important open problem to generalize the above hydrodynamic discussion to the case of colored gluons. In particular, the existence of three- and four-string vertices must be accounted for in a simple way. A promising intermediate solution is suggested by ’t Hooft’s [27] hypothesis of dominance of abelian monopoles [28].

Notes and References

[1] As a cross check, we calculate the value of the flux one more by transforming the twice reduced field variables to physical ones. Thus we insert into the physical flux \( \Phi^{\text{phys}} = \oint d^4x \cdot A^{\text{phys}} = \sqrt{k_B T_c / \xi_0} \sqrt{-\tau / g \kappa A} \) and \( x^{\text{phys}} / x = \xi_0 / \sqrt{-\tau} \) to find \( \Phi^{\text{phys}} = \sqrt{\xi_0 k_B T_c (1/q)} \oint d^4x \cdot A = \sqrt{\xi_0 k_B T_c} 2\pi n / q = n(ch/2e) \), where we have used that by Eq. (7A.136), \( q = (2e/\pi e) \sqrt{k_B T_c \xi_0} \).


[9] See, most notably, the textbook by Jackson in Ref. [3], p. 258, where it is stated that “a choice of different string positions is equivalent to different choices of (electromagnetic) gauge”; also his Eq. (6.162) and the lines below it. In the unnumbered equation on p. 258 Jackson observes that the physical monopole field is \( F_{\text{monop}}^{ab} = F_{ab} - F_{M}^{ab} \) but the independent gauge properties of \( F_{M}^{ab} \) and the need to use the action (8.18) rather than (8.19) are not noticed.

[10] Compare the review article by P. Goddard and D. Olive, Progress in Physics 41, 1357 (1978) who use in their review article “general gauge transformations” [their Eq. (2.46)]. They observe that the field tensor \( F_{\text{obs}}^{ab} = F_{ab} + (1/a^2 a) \Phi \cdot (\partial_a \Phi \times \partial_b \Phi) \) introduced by ’t Hooft into his SU(2) gauge theory with Higgs fields \( \Phi \) to describe magnetic monopoles can be brought to the form \( F_{\text{obs}}^{ab} = F_{ab} - F_{M}^{ab} \) by a gauge transformation within the SU(2) gauge group which moves magnetic fields from \( F_{ab} \) to \( F_{M}^{ab} \) without changing \( F_{\text{obs}}^{ab} \) [see their Eq. (4.30) and the last two equations in their Section 4.5]. Note that these are not permissible gauge transformation of the electromagnetic type. They are similar to our monopole gauge transformations of the form (8.25), although with more general transformation functions than (8.24) due to the more general SU(2) symmetry.


[16] Surface tension and stiffness coming from interactions like the first term in (8.79) have been calculated for biomembranes in H. Kleinert, *Dynamical Generation of String Tension and Stiffness in Strings and Membranes*, Phys. Lett. B 211, 151 (1988) (k1/177).


H. Reinhardt, Phys. Lett. B 244, 316 (1990);


F. Brandstaeter, G. Schierholz, and U.-J. Wiese, DESY preprint 91-040 (1991);
V.G. Bornyakov, E.M. Ilgenfritz, M.L. Laursen, V.K. Mitrijushkin, M. Müller-
J. Greensite and J. Winchester, Phys. Rev. D 40, 4167 (1989);
A. Di Giacomo, M. Maggiore, and Š. Olejn´ık, Nucl. Phys. B 347, 441 (1990);
9

Multivalued Mapping from Ideal Crystals to Crystals with Defects

In the last chapter we have learned how multivalued gauge transformations allow us to transform theories in field-free space into theories coupled to electromagnetism. By analogy, we expect that multivalued coordinate transformations can be used to map theories in flat space into theories in spaces with curvature and torsion. This is indeed possible. The mathematical methods have been developed in the theory of line-like defects in crystals [1, 2, 3]. Let us briefly review those parts of the theory which will be needed for our purposes.

9.1 Defects

No crystal produced in the laboratory is perfect. It always contains a great number of defects. These may be chemical, electrical, or structural in character, involving foreign atoms. They may be classified according to their space dimensionality. The simplest type of defect is the point defect. It is characterized by the fact that within a certain finite neighborhood only one cell shows a drastic deviation from the perfect crystal symmetry. The most frequent origin of such point defects is irradiation or an isotropic mechanical deformation under strong shear stresses. There are two types of intrinsic point defects. Either an atom may be missing from its regular lattice site (vacancy) or there may be an excess atom (interstitial) (see Fig. 9.1). Vacancies and interstitials are mobile defects. A vacancy can move if a neighboring atom moves into its place, leaving a vacancy at its own former position. An interstitial atom can move in two ways. It may hop directly from one interstitial site to another. This happens in strongly anisotropic materials such as graphite but also in some cubic materials like Si or Ge. Or it may move in a way more similar to the vacancies by replacing atoms, i.e., by pushing a regular atom out of its place into an interstitial position which, in turn affects the same change on its neighbor, etc.

Intrinsic point defects have the property that if a number of them move close together, the total energy becomes smaller than the sum of the individual energies.
The reason for this is easily seen. If two vacancies in a simple cubic lattice come to lie side by side, there are only 10 broken valencies compared to 12 when they are separated. If a larger set of vacancies comes to lie side by side forming an entire disc of missing atoms, the crystal planes can move together and make the disc disappear (see Fig. 9.2). In this way, the crystal structure is repaired. Only close to the boundary line, such a repair is impossible. The boundary line forms a line-like defect.

Certainly, line-like defects can arise also in an opposite process of clustering of interstitial atoms. If they accumulate side by side forming an interstitial disc, the crystal planes move apart and accommodate the additional atoms in a regular atomic array, again with the exception of the boundary line. Line-like defects of this type are called dislocation lines.

It is obvious that a dislocation line need not only consist of a single disc of missing or excessive atoms. There can be several discs stacked on top of each other. Their boundary forms a dislocation line of higher strength. The energy of such a higher

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure91.png}
\caption{Intrinsic point defects in a crystal. An atom may become interstitial, leaving behind a vacancy. It may perform random motion via interstitial places until it reaches another vacancy where it recombines. The exterior of the crystal may be seen as a reservoir of vacancies.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure92.png}
\caption{Formation of a dislocation line (of the edge type) from a disc of missing atoms. The atoms above and below the missing ones have moved together and repaired the defect, except at the boundary.}
\end{figure}
dislocation line increases roughly with the square of the strength. Dislocations are created and set into motion if stresses exceed certain critical values. This is why they were first seen in plastic deformation experiments of the nineteenth century in the form of slip bands. The grounds for their theoretical understanding were laid much later by Frenkel who postulated the existence of crystalline defects in order to understand why materials yield to plastic shear about a thousand times more easily than one might expect on the basis of a naive estimate (see Fig. 9.3).

The large discrepancy was explained by Frenkel who noted that the plastic slip would not proceed by the two halves moving against each other as a whole but stepwise, by means of defects. In 1934, Orowan, Polany, and Taylor identified these defects as dislocation lines. The presence of a single moving edge dislocation allows for a plastic shear movement of the one crystal half against the other. The movement proceeds in the same way as that of a caterpillar. This is pictured in Fig. 9.4. One leg is always in the air breaking translational invariance and this is exchanged against the one in front of it, etc. In the crystal shown in the lower part of Fig. 9.4, the single leg corresponds to the lattice plane of excess atoms. Under stress along the arrows, this moves to the right. After a complete sweep across the crystal, the upper half is shifted against the lower by precisely one lattice spacing.

If many discs of missing or excess atoms come to lie close together there exists a further cooperative phenomenon. This is illustrated in Fig. 9.5. On the left-hand side, an infinite number of atomic half planes (discs of semi-infinite size) has been removed from an ideal crystal. If the half planes themselves form a regular crystalline array, they can fit smoothly into the original crystal. Only at the origin is there a breakdown of crystal symmetry. Everywhere else, the crystal is only slightly

![Figure 9.3](image_url)

**Figure 9.3** Naive estimate of maximal stress supported by a crystal under shear stress as indicated by the arrows. The two halves tend to slip against each other. Assuming a periodic behavior $\sigma = \sigma_{\text{max}} \sin(2\pi x/a)$, this reduces to $\sigma \sim \sigma_{\text{max}} 2\pi (x/a) \sim \mu(x/a)$. Hence $\sigma_{\text{max}} = \mu/2\pi$. Experimentally, however, $\sigma_{\text{max}} \sim 10^{-3}\mu$ to $10^{-4}\mu$. 
9.2 Dislocation Lines and Burgers Vector

Distorted. What has been formed is again a line-like defect called a disclination. Dislocations and disclinations will play a central role in our further discussion.

Before coming to this let us complete the dimensional classification of two-dimensional defects. They are of three types. There are grain boundaries where two regular lattice parts meet, with the lattice orientations being different on both sides of the interface (see Fig. 9.6).

They may be considered as arrays of dislocation lines in which half planes of point defects are stacked on top of each other with some spacing, having completely regular lattice planes between them. The second type of planar defects are stacking faults. They contain again completely regular crystal pieces on both sides of the plane, but instead of being oriented differently they are shifted one with respect to the other (see Fig. 9.7). The third unavoidable type is the surface of the crystal.

From now on we shall focus attention upon the line-like defects.
Figure 9.5 Formation of a disclination from a stack of layers of missing atoms (cf. Fig. 9.2). Equivalently, one may cut out an entire section of the crystal. In a real crystal, the section has to conform with the symmetry angles. In the continuum approximation, the angle $\Omega$ is meant to be very small.

Figure 9.6 Grain boundary where two crystal pieces meet with different orientations in such a way not every atomic layer matches (here only every other one does).

9.2 Dislocation Lines and Burgers Vector

Let us first see how a dislocation line can be characterized mathematically. For this we look at Fig. 9.8 in which a closed circuit in the ideal crystal is mapped into the disturbed crystal. The orientation is chosen arbitrarily to be anticlockwise. The prescription for the mapping is that for each step along a lattice direction, a corresponding step is made in the disturbed crystal. If the original lattice sites are denoted by $x_n$, the image points are given by $x_n + u(x_n)$, where $u(x_n)$ is the displacement amount field: At each step, the image point moves in a slightly different original point. After the original point has completed a closed circuit, call it $B_0$, the image point will not have arrived at the point of departure. The image of the closed contour $B_0$ is no longer closed. This failure to close is given precisely by a
lattice vector \( \mathbf{b}(x) \) called a local Burgers vector, which points from the beginning to the end of the circuit.\(^1\) Thus, the dislocation line is characterized by the following equation,

\[
\sum_{B_0} \Delta u_i(x) = b_i, \tag{9.1}
\]

where \( \Delta u_i(x_n) \) are the increments of the displacement vector from step to step. Equivalently, we can consider a closed circuit in the disturbed crystal, call it \( B \), and find that its counter image in the ideal crystal does not close by a vector \( \mathbf{b} \) called the true Burgers vector which now points from the end to the beginning of the circuit.

If we consider the same process in the continuum limit, we can write

\[
\int_{B_0} du_i(x_a) = b_i, \quad \int_{\mathcal{B}} du_i(x_a') = b_i. \tag{9.2}
\]

The closed circuit \( \mathcal{B} \) is called Burgers circuit. The two Burgers vectors are the same if both circuits are so large that they lie deep in the ideal crystal. Otherwise they differ by an elastic distortion.

A few remarks are necessary concerning the convention employed in defining the Burgers vector. The singular line \( L \) is in principle without orientation. We may arbitrarily assign a direction to it. The Burgers circuit is then taken to encircle this chosen direction in the right-handed way. If we choose the opposite direction, Burgers vector changes sign. However, the products \( b_i, dx_j \), where \( dx_j \) is the infinitesimal tangent vector to \( L \), are invariant under this change. Note that this is similar to the magnetic case discussed in Part II. There one defined the direction of the current by the flow of positive charge. The Burgers circuit gives \( \oint du = I \). One

---

\(^1\)Our sign convention is the opposite of Bilby et al. and the same as Read’s (see Notes and References). Note that in contrast to the local Burgers vector, the true Burgers vector is defined on a perfect lattice.
could, however, also reverse this convention referring to the negative charge. Then \( \oint du \) would give \(-I\). Again, \( I \cdot dx_i \) is an invariant. Only these products can appear in physical observables such as the biot-Savart law.

The invariance of \( b_i dx_j \) under reversal of the orientation has a simple physical meaning. In order to see this, consider once more the above dislocation line which was created by removing a layer of atoms. We can see in Fig. 9.8 that in this case \( b \times dx \) points \textit{inwards}, namely, towards the vacancies. Consider now the opposite case in which a layer of new atoms is inserted between the crystal planes forcing the planes apart to relax the local stress. If we now calculate \( \oint_P du_i(x) = b_i \), we find that \( x \times dx \) points \textit{outwards}, i.e. away from the inserted atoms. This is again the direction in which there are fewer atoms. Both statements are independent of the choice of the orientation of the Burgers circuit. Since the second case has extra atoms inside the circle, where the previous one had vacancies, the two can be considered as antidefects of one another. If the boundary lines happen to fall on top of each other, they can annihilate each other and re-establish a perfect crystal. This can happen only piece-wise in which case the parts where the lines differ remain a
dislocation lines. In both examples, the Burgers vectors are everywhere orthogonal to the dislocation line and one speaks of a pure edge dislocation (see Fig. 9.2).

There is no difficulty in constructing another type of dislocation by cutting a crystal along a lattice half-plane up to some straight line $L$, and translating one of the lips against the other along the direction of $L$. In this way one arrives at the so-called screw dislocation shown in Fig. 9.9 in which the Burgers vector points parallel to the line $L$.

When drawing crystals out of a melt, it always contains a certain fraction of dislocations. Even in clean samples, at least one in $10^6$ atoms is dislocated. Their boundaries run in all directions through space. We shall see very soon that their Burgers vector is a topological invariant for any closed dislocation loop. Therefore, the character “edge” versus “screw” of a dislocation line is not an invariant. It changes according to the direction of the line with respect to the invariant Burgers vector $b_i$. It is obvious from the Figs. 9.2 to 9.9 that a dislocation line destroys the translational invariance of the crystal by multiples of the lattice vectors. If there are only a few lines this destruction is not very drastic. Locally, i.e., in any small subspecimen which does not lie too close to the dislocation line, the crystal can still be described by a periodic array of atoms whose order is disturbed only slightly by a smooth displacement field $u_i(x)$.

9.3 Disclination Lines and Frank Vector

Since the crystal is not only invariant under discrete translations but also under certain discrete rotations we expect the existence of another type of defect which is capable of destroying the global rotational order, while maintaining it locally. These are the disclination lines of which one example was given in Fig. 9.5. It arose as a superposition of stacks of layers of missing atoms. In the present context, it is useful to construct it by means of the following Gedanken experiment. Take a
regular crystal in the form of cheese and remove a section subtending an angle Ω (see Fig. 9.10). The free surfaces can be forced together. For large Ω this requires considerable energy. Still, if the atomic layers on the free surfaces match together perfectly, the crystal can re-establish locally its periodic structure. This happens for all symmetries of the crystal. In a simple cubic crystal, Ω can be 90°, 180°, 270°. The 90° case is displayed in Fig. 9.11.

![Figure 9.10 Volterra cutting and welding process leading to a wedge disclination.](image)

**Figure 9.10** Volterra cutting and welding process leading to a wedge disclination.

In Fig. 9.11 we can imagine also the opposite procedure going from the right in Fig. 9.10 to the left. We may cut the crystal, force the lips open by Ω and insert new undistorted crystalline matter to match the atoms in the free surfaces. These are the disclinations of negative angles. The case for Ω = −90° is shown in Fig. 9.11.

The local crystal structure is destroyed only along the singular line along the axis of the cheese. The rotation which has to be imposed upon the free surfaces in order to force them together may be represented by a rotation vector Ω which, in the present example, points parallel to \( L \) and to the cut. This is called a *wedge disclination*. It is not difficult to construct other rotational defects. The three possibilities are shown in Figs. 9.12. Each case is characterized by a vector. In the first case, Ω pointed parallel to the line \( L \) and the cut. Now, in the second case, it is orthogonal to the line \( L \) and Ω points parallel to the cut. This is a *splay disclination*. In the third case, Ω points orthogonal to the line and cut. This is a *twist disclination*.
9.4 Interdependence of Dislocation and Disclinations

The vector Ω is referred to as the Frank vector of the disclination. Just as in the construction of dislocations, the interface at which the material is joined together does not have any physical reality. For example, in Fig. 9.12a we could have cut out the piece along any other direction which is merely rotated with respect to the first around L by a discrete symmetry irregular piece as long as the faces fit together smoothly (recall Fig. 9.10). Only the singular line is a physical object.

The Gedanken experiments of cutting a crystal, removing or inserting slices or sections, and joining the free faces smoothly together were first performed by Volterra in 1907. For this reason one speaks of the creation of a defect line as a Volterra process and calls the cutting surfaces, where the free faces are joined together, Volterra surfaces.

Figure 9.12 Three different possibilities of constructing disclinations: (a) wedge, (b) splay, and (c) twist disclinations.

9.4 Interdependence of Dislocation and Disclinations

It must be pointed out that dislocation and disclination lines are not completely independent. We have seen before in Fig. 9.5 that a disclination line was created by removing stacks of atomic layers from a crystal. But each layer can be considered as
a dislocation line running along the boundary. Thus a disclination line is apparently indistinguishable from a stack of dislocation lines, placed with equal spacing on top of each other. Conversely, a dislocation line is very similar to a pair of disclination lines running in opposite directions close to each other. This is illustrated in Fig. 9.13. What we have here is a pair of opposite Volterra processes of disclination lines. We have cut out a section of angle $\Omega$, but instead of removing it completely we have displaced it merely by one lattice spacing $a$. This is equivalent to generating a disclination of the Frank vector $\Omega$ and another one with the opposite Frank vector $-\Omega$ whose rotation axis is displaced by $a$. It is obvious from the figure that the result is a dislocation line with Burgers vector $b$.

Because of this interdependence between dislocations and disclinations, the defect lines occurring in a real crystal will, in general, be of a mixed nature. It must be pointed out that disclinations were first observed and classified by F.C. Frank in 1958 in the context of liquid crystals. Liquid crystals are mesophases. They are liquids consisting of rod-like molecules. Thus, they cannot be described by a displacement field $u_i(x)$ alone but require an additional orientational field $n_i(x)$ for their description. This orientation is independent of the rotational field $\omega_i(x) = \frac{1}{2} \varepsilon_{ijk} \partial_j u_k(x)$. The disclination lines defined by Frank are the rotational defect lines with respect to this independent orientational degree of freedom. Thus, they are a priori unrelated to the disclination lines in the rotation field $\omega_i(x) = \frac{1}{2} \varepsilon_{ijk} \partial_j u_k(x)$. In fact, the liquid is filled with dislocations and $\omega$-disclinations even if the orientation field $n_j(x)$ is completely ordered.

Friedel in his book on dislocations (see the references at the end) calls the $n_j$-disclinations, rotation dislocations. But later the name disclinations became customary (see Kléman’s article cited in the Notes and References). In general, there is little danger of confusion, if one knows what system and phase one is talking about.

![Figure 9.13](image)

**Figure 9.13** Generation of dislocation line from a pair of disclination lines running in opposite directions at a fixed distance $b$. The Volterra process amounts to cutting out a section and reinserting it, but shifted by the amount $b$.

### 9.5 Defect Lines with Infinitesimal Discontinuities

H. Kleinert, GRAVITY WITH TORSION
9.5 Defect Lines with Infinitesimal Discontinuities in Continuous Media

The question arises as to how one can properly describe the wide variety of line-like defects which can exist in a crystal. In general, this is a rather difficult task due to the many possible different crystal symmetries. For the sake of gathering some insight it is useful to restrict oneself to continuous isotropic media. Then defects may be created with arbitrarily small Burgers and Frank vectors. Such infinitesimal defects have the great advantage of being accessible to differential analysis. This is essential for a simple treatment of rotational defects. It permits a characterization of disclinations in a way which is very similar to that of dislocations via a Burgers circuit integral. Consider, for example, the wedge disclination along the line \( L \) (shown in Figs. 9.5, 9.10, 9.11 or 9.12a), and form an integral over a closed circuit \( B \) enclosing \( L \).

Just as in the case of dislocations this measures the thickness of the material section removed in the Volterra process. Unlike the situation for dislocations, this thickness increases with distance from the line. If \( \Omega \) is very small, the displacement field across the cut has a discontinuity which can be calculated from an infinitesimal rotation

\[
\Delta u_i = (\Omega \times x)_i,
\]

where \( x \) is the vector pointing to the place where the integral starts and ends. In order to turn this statement into a circuit integral it is useful to remove the explicit dependence on \( x \) and consider not the displacement field \( u_i(x) \) but the local rotation field accompanying the displacement instead. This is given by the antisymmetric tensor field

\[
\omega_{ij}(x) = \frac{1}{2} \left[ \partial_i u_j(x) - \partial_j u_i(x) \right].
\]

(9.4)

The rotational character of this tensor field is obvious by looking at the change of an infinitesimal distance vector under a distortion

\[
dx'_i - dx_i = (\partial_j u_i) dx_j = u_{ij} dx_j - \omega_{ij} dx_j.
\]

(9.5)

The tensor field \( \omega_{ij} \) is associated with a vector field \( \omega_i \) as follows:

\[
\omega_{ij}(x) = \varepsilon_{ijk} \omega_k(x)
\]

(9.6)

i.e.,

\[
\omega_{ij}(x) = \frac{1}{2} \varepsilon_{ijk} \omega_j(x) = \frac{1}{2} (\nabla \times u)_i.
\]

(9.7)

The right-hand side of (9.6) separates the local distortion into a sum of a local change of shape and a local rotation. Now, when looking at the wedge disclination
in Fig. 9.12a, we see that due to (9.3), the field $\omega_i(x)$ has a constant discontinuity $\Omega$ across the cut. This can be formulated as a circuit integral

$$\Delta \omega_i = \oint_B d\omega_i = \Omega_i.$$  \hspace{1cm} (9.8)

The value of this integral is the same for any choice of the circuit $B$ as long as it encloses the disclination line $L$.

This simple characterization depends essentially on the infinitesimal size of the defect. If $\Omega$ were finite, the differential expression (9.3) would not be a rotation and the discontinuity across the cut could not be given in the form (9.8) without specifying the circuit $B$. The difficulties for finite angles are a consequence of the non-Abelian nature of the rotation group. Only infinitesimal local rotations have additive rotation angles, since the quadratic and higher-order corrections can be neglected.

### 9.6 Multivaluedness of Displacement Field

As soon as a crystal contains a few dislocations, the definition of displacement field is intrinsically non-unique. The displacement field is intrinsically non-unique. The displacement fields is really *multivalued*. In a perfect crystal, in which the atoms deviate little from their equilibrium positions $x$, it is natural to draw the displacement vector from the lattice places $x$ to the nearest atom. In principle, however, the identity of the atoms makes such a specific assignment impossible. Due to thermal fluctuations, the atoms exchange positions from time to time by a process called *self-diffusion*. After a very long time, the displacement vector, even in a regular crystal, will run through the entire lattice. Thus, if we describe a regular crystal initially by very small displacement vectors $u_i(x)$, then, after a very long time, these will have changed to a permutation of lattice vectors, each of them occurring precisely once, plus some small fluctuations around them. Hence the displacement vectors are intrinsically multivalued, with $u_i(x)$ being indistinguishable from $u_i(x) + aN_i(x)$, where $N_i(x)$ are integer numbers and $a$ is the lattice spacing.

It is interesting to realize that this property puts the displacement fields on the same footing with the phase variable $\gamma(x)$ of superfluid $^4$He. There the indistinguishability of $\gamma(x)$ and $\gamma(x) + 2\pi N(x)$ has an entirely different reason: it follows directly from the fact that the physical field is the complex field $\psi(x) = |\psi(x)| e^{i\gamma(x)}$, which is invariant under the exchange $\gamma(x) \to \gamma(x) + 2\pi N(x)$.

Thus, in spite of the different physics described by the variables $\gamma(x), u_i(x)$, they both share the characteristic multivaluedness. It is just as if the rescaled $u_i(x)$ variables $\gamma_i(x) = (2\pi/a)u_i(x)$ were phases of three complex fields

$$\psi_i(x) = |\psi_i(x)| e^{i\gamma_i(x)},$$

which serve to describe the positions of the atoms in a crystal.

In a regular crystal, the multivaluedness of $u_i(x)$ has no important physical consequences. The atoms are strongly localized and the exchange of positions occurs...
very rarely. The exchange is made irrelevant by the identity of the atoms and symmetry of the many-body wave function. This is why the natural assignment of \( u_i(\mathbf{x}) \) to the nearest equilibrium position \( \mathbf{x} \) presents no problems. As soon as defects are present, however, the full ambiguity of the assignment comes up: When removing a layer of atoms, the result is a dislocation line along the boundary of the layer. Across the layer, the positions \( u_i(\mathbf{x}) \) jump by a lattice spacing. This means that the atoms on both sides are interpreted as having moved towards each other. Figure 9.14 shows that the same dislocation line could have been constructed by removing a completely different layer of atoms, say \( S' \), just as long as it has the same boundary line. The jump of the displacement field across the shifted layer \( S' \) corresponds to the neighboring atoms of this layer having moved together and closed the gap. Physically, there is no difference. There is only a difference in the descriptions which amounts to a difference in the assignment of the equilibrium positions from where to count the displacement field \( u_i(\mathbf{x}) \). In contrast to regular crystals there now exists no choice of the nearest equilibrium point. It is this multivaluedness which will form the basis for the geometric description of the defects in solids.

\[
\left( \partial_i \partial_j - \partial_j \partial_i \right) \Omega(\mathbf{x}) = 0.
\] (9.9)
In the crystal, this property will hold away from the cutting surface $S$, where $u_i(x)$ is perfectly smooth and satisfies the corresponding integrability condition

$$\left( \partial_i \partial_j - \partial_j \partial_i \right) u_k(x) = 0. \quad (9.10)$$

Across the surface, $u_i(x)$ is discontinuous. However, the open faces of the crystalline material must fit properly to each other. This implies that the strain as well as its first derivatives should have the same values on both sides of the cutting surface $S$:

$$\Delta u_{ij} = 0, \quad (9.11)$$
$$\Delta \partial_k u_{ij} = 0. \quad (9.12)$$

This severely restricts the discontinuities of $u_i(x)$ across $S$. In order to see this let $x(1), x(2)$ be two different crystal points slightly above and below $S$ and $C^+, C^-$ be two curves connecting the two points (see Fig. 9.15). We can then calculate the difference of the discontinuities as follows:

$$\Delta u_i(1) - \Delta u_i(2) = \left[ u_i(1^-)u_i(1^+) \right] - \left[ u_i(2^-) - u_i(2^+) \right]$$
$$= \int_{C^+}^{2^+} dx_j \partial_j u_i - \int_{C^-}^{1^-} dx_j \partial_j u_i. \quad (9.13)$$

Using the local rotation field $\omega_{ij}(x)$ we can rewrite this as

$$\Delta u_i(1) - \Delta u_i(2) = \int_{C^+}^{2^+} dx_j dx_j \left( u_{ij} - \omega_{ij} \right)$$
$$- \int_{C^-}^{1^-} dx_j \left( u_{ij} - \omega_{ij} \right). \quad (9.14)$$

The $\omega_{ij}$ pieces may be integrated by parts:

$$- (x_j - x_j(1^+)) \omega_{ij} \bigg|_{1^+}^{2^+} + \int_{1^+}^{2^+} dx_k \left( x_j - x_j(1^+) \right) \partial_k \omega_{ij}$$

**Figure 9.15** Geometry used in the derivation of Weingarten’s theorem [Eqs. (9.13)–(9.22)].
\[ + \left( x_j - x_j(1^-) \right) \omega_{ij} \left| \int_{1^-}^{2^-} dx_k \left( x_j - x_j(1^-) \right) \partial_k \omega_{ij} \right] \] (9.15)

\[ = \left[ - \left( x_j(2^+) - x_j(1^+) \right) \omega_{ij}(2^+) + \int_{1^+}^{2^+} dx_k \left( x_j - x_j(1^+) \right) \partial_k \omega_{ij} \right] - [+ \rightarrow -]. \]

Since

\[ x_j(1^+) = x_j(1^-), \quad x_j(2^+) = x_j(2^-), \]

we arrive at the relation

\[ \Delta u_i(1) - \Delta u_i(2) = - \left( x_j(1) - x_j(2) \right) \left( \omega_{ij}(2^-) - \omega_{ij}(2^+) \right) \]

\[ + \oint_{C^+} dx_k \{ u_{ik} + (x_j - x_j(1)) \partial_k \omega_{ij} \}, \quad (9.16) \]

where \( C^{+\neg} \) is the closed contour consisting of \( C^+ \) followed by \( -C^- \). Since \( C^+ \) and \( -C^- \) are running back and forth on top of each other, the closed contour integral can be rewritten as a single integral along \( -C^- \) with \( u_{ik} \) and \( \partial_k \omega_{ij} \) replaced by their discontinuities across the sheet \( S \). Moreover, the discontinuity of \( \partial_k \omega_{ij} \) can be decomposed in the following manner:

\[ \Delta \left( \partial_k \omega_{ij} \right) = \frac{1}{2} \partial_k \left( \partial_i u_j - \partial_j u_i \right) (x^-) - (x^- \rightarrow x^+) \]

\[ = \partial_i u_{kj}(x^-) - \partial_j u_{ki}(x^-) + \frac{1}{2} \left( \partial_k \partial_i - \partial_i \partial_k \right) u_j(x^-) \]

\[ - \frac{1}{2} \left( \partial_k \partial_j - \partial_j \partial_k \right) u_i(x^-) + \frac{1}{2} \left( \partial_j \partial_i - \partial_i \partial_j \right) u_k(x^-) - (x^- \rightarrow x^-). \]

Since above and below the sheet, the displacement field is smooth, the two derivatives in front of \( u(x^\pm) \) commute. Hence the integral in (9.16) becomes

\[ - \int_{C^-} dx_k \left\{ \Delta u_{ij} + (x_j - x_j)(1) \Delta \left( \partial_i u_{kj} - \partial_j u_{ki} \right) \right\}. \quad (9.18) \]

This expression vanishes due to the physical requirements (9.11) and (9.12). As a result we find that the discontinuities between two arbitrary points 1 and 2 on the sheet have the simple relation

\[ \Delta u_i(2) = \Delta u_i(1) - \Omega_{ij} \left( x_j(2) - x_j(1) \right), \quad (9.19) \]

where \( \Omega_{ij} \) is a fixed infinitesimal rotation matrix given by

\[ \Omega_{ij} = \Delta \omega_{ij}(2) = \omega_{ij}(2^-) - \omega_{ij}(2^+). \quad (9.20) \]

We now define the rotation vector \( \Omega \) with components

\[ \Omega_k = \frac{1}{2} \varepsilon_{ijk} \Omega_{ij} \quad (9.21) \]
in terms of which (9.19) takes the form
\[ \Delta \mathbf{u}(2) = \Delta \mathbf{u}(1) + \Omega \times (\mathbf{x}(2) - \mathbf{x}(1)). \] (9.22)

This forms the content of Weingarten’s theorem: The discontinuity of the displacement field across the cutting surface can only be a constant vector plus a fixed rotation.

Note that these are precisely the symmetry elements of a solid continuum. When looking back at the particular dislocation and disclination lines in Figs. 9.2–9.12 we see that all the discontinuities obey this theorem, as they should. The vector \( \Omega \) is the Frank vector of the disclination lines. For a pure disclination line, \( \Omega = 0 \) and \( \Delta \mathbf{u}(1) = \Delta \mathbf{u}(2) = \mathbf{b} \) is the Burgers vector.

### 9.8 Integrability Properties of Displacement Field

The rotation field \( \omega_{ij}(\mathbf{x}) \) has also nontrivial integrability properties. Taking Weingarten’s theorem (9.19) and forming derivatives, we see that the jump of the \( \omega_{ij}(\mathbf{x}) \) field is necessarily a constant, namely \( \Omega_{ij} \). Hence \( \omega_{kl} \) also satisfies the integrability condition
\[ \left( \partial_i \partial_j - \partial_j \partial_i \right) \omega_{kl} = 0, \] (9.23)
everywhere except on the defect line. The argument is the same as that for the vortex lines. We simply observe that the contour integral over a Burgers circuit
\[ \Delta \omega_{ij} = \oint_B d\omega_{ij} = \oint_B dx_k \partial_k \omega_{ij}, \] (9.24)
can be cast, by Stokes’ theorem, in the form
\[ \Delta \omega_{ij} = \int_{S^B} dS_m \varepsilon_{mkl} \partial_k \partial_l \omega_{ij}, \] (9.25)
where \( S^B \) is some surface enclosed by the Burgers circuit. Since the result is independent of the size, shape, and position of the Burgers circuit as long as it encloses the defect line \( L \), this implies
\[ \varepsilon_{mkl} \partial_k \partial_l \omega_{ij}(\mathbf{x}) = 0 \] (9.26)
everywhere away from \( L \), which is what we wanted to show.

In fact, the constancy of the jump in \( \omega_{ij} \) could have been derived somewhat more directly, without going through (9.23)–(9.26), by taking again the curves \( C^+, C^- \) on Fig. 9.15 and calculating
\[ \Delta \omega_{ij}(1) - \Delta \omega_{ij}(2) = \int_{C^+} dx_k \partial_k \omega_{ij} - \int_{C^-} dx_k \partial_k \omega_{ij} = -\int_{C^-} dx_k \Delta \left( \partial_k \omega_{ij} \right). \] (9.27)
From the assumptions (9.11) and (9.12), together with (9.18), we see that $\omega_{ij}(x)$ does not jump across the Volterra surface $S$. But then (9.27) shows us that $\Delta \omega_{ij}$ is a constant.

Let us now consider the displacement field itself. As a result of Weingarten’s theorem, the integral over the Burgers circuit $B_2$ in Fig. 9.15 gives

$$\Delta u_i(2) = \oint_{B_2} du_i = \Delta u(1) - \Omega_{ij} \left[ x_j(2) - x_j(1) \right], \quad \text{(9.28)}$$

$$\Delta u_i(1) - \Omega_{ij} \left[ x_j(2) - x_j(1) \right] = \oint_{B_2} dx_k \left\{ u_{ik} + \left[ x_j - x_j(2) \right] \partial_k \omega_{ij} \right\}. \quad \text{(9.29)}$$

Here we observe that the factors of $x_i(2)$ can be dropped on both sides by (9.24) and $\Delta \omega_{ij} = \Omega_{ij}$. By Stokes' theorem, the remaining equation then becomes an equation for the surface integral over $S_{B_2}$,

$$\Delta u_i(1) + \Omega_{ij} x_j(1) = \int_{S_{B_2}} dS_i \varepsilon_{lmk} \partial_m \left( u_{ik} + x_j \partial_k \omega_{ij} \right)$$

$$= \int_{S_{B_2}} dS_i \varepsilon_{lmk} \left[ \left( \partial_m u_{ik} + \partial_k \omega_{im} \right) + x_j \partial_m \partial_k \omega_{ij} \right]. \quad \text{(9.30)}$$

This must hold for any size, shape, and position of the circuit $B_2$ as long as it encircles the defect line $L$. For all these different configurations, the left-hand side of (9.30) is a constant. We can therefore conclude that

$$\int_S dS_i \varepsilon_{lmk} \left( \partial_m u_{ik} + \partial_k \omega_{im} \right) + x_j \varepsilon_{lmk} \partial_k \omega_{ij} = 0 \quad \text{(9.31)}$$

for any surfaces $S$ which does not enclose $L$. Moreover, from (9.23) we see that the last term cannot contribute. The first two terms, on the other hand, can be rewritten, using the same decomposition of $\partial_k \omega_{im}$ as in (9.18), in the form

$$- \int_S dS_i \varepsilon_{lmk} \left( S_{kmi} - S_{mik} + S_{ikm} \right) = \int_S dS_i \varepsilon_{lmk} S_{nki}, \quad \text{(9.32)}$$

where we have abbreviated

$$S_{kmi}(x) \equiv \frac{1}{2} \left( \partial_k \partial_m - \partial_m \partial_k \right) u_i(x). \quad \text{(9.33)}$$

Since this has to vanish for any $S$, we conclude that away from the defect line, the displacement field $u_i(x)$ also satisfies the integrability condition

$$\left( \partial_k \partial_m - \partial_m \partial_k \right) u_i(x) = 0. \quad \text{(9.34)}$$

On the line $L$, the integrability conditions for $u_i$ and $\omega_{ij}$ are, in general, both violated. Let us first consider $\omega_{ij}$. In order to give the constant result $\Delta \omega_{ij}(x) \equiv \Omega_{ij}$ in (9.25) the integrability condition must be violated by a singularity in the form of a $\delta$-function along the line $L$ (4.10), namely:

$$\varepsilon_{lmk} \partial_k \omega_{ij} = \delta_i(x; L). \quad \text{(9.35)}$$
Then (9.25) gives $\Delta \omega_{ij} = \Omega_{ij}$ via the formula

$$\int_{S^B} dS_l \delta_l(\mathbf{x}; L) = 1.$$  \hspace{1cm} (9.36)

In order to see how the integrability condition is violated for $u_i(\mathbf{x})$, consider now the integral (9.30) and insert the result (9.32). This gives

$$\Delta u_i(1) + \Omega_{ij} x_j(1) = \int_{S^{B_2}} dS_l \varepsilon_{lmk} \left(S_{mki} + x_j \partial_m \partial_k \omega_{ij}\right).$$  \hspace{1cm} (9.37)

The right hand side is a constant independent of the position of the surface $S^{B_2}$. This implies that the singularity along $L$ is of the form

$$\varepsilon_{lmk} \left(\partial_m \partial_k u_i + x_j \partial_m \partial_k \omega_{ij}\right) = b_i \delta_l(\mathbf{x}; L),$$  \hspace{1cm} (9.38)

where we have introduced the quantity

$$b_i \equiv \Delta u_i(1) + \Omega_{ij} x_j(1).$$  \hspace{1cm} (9.39)

Inserting (9.35) into (9.38) leads to the following violation of the integrability condition for $u_i(\mathbf{x})$ along $L$,

$$\varepsilon_{lmk} \partial_m \partial_k u_i = (b_i - \Omega_{ij} x_j) \delta_l(\mathbf{x}; L).$$  \hspace{1cm} (9.40)

In terms of the tensor (9.33), this reads

$$\varepsilon_{lmk} S_{mki} = (b_i - \Omega_{ij} x_j) \delta_l(\mathbf{x}; L).$$  \hspace{1cm} (9.41)

### 9.9 Dislocation and Disclination Densities

The violation of the integrability condition for displacement and rotation fields proportional to $\delta$-functions along lines $L$ is analogous to the situation in the multivalued description of the magnetic field in the last chapter. The analogy can be carried further. Consider, for example, the current density of magnetism in Eq. (4.7), which by Eqs. (4.30) and Eqs. (4.36) can be rewritten in the multivalued description as

$$j_i(\mathbf{x}) = \epsilon_{ijk} \partial_j B_k = \frac{I}{4\pi} \epsilon_{ijk} \partial_j \partial_k \Omega = I \delta_i(\mathbf{x}; L),$$  \hspace{1cm} (9.42)

where $\Omega$ is the solid angle under which the loop $L$ is seen from the point $\mathbf{x}$. By analogy, we introduce densities for dislocations and disclinations, respectively, as follows:

$$\alpha_{ij}(\mathbf{x}) \equiv \epsilon_{ikl} \partial_k \partial_l u_j(\mathbf{x}),$$  \hspace{1cm} (9.43)

$$\theta_{ij}(\mathbf{x}) \equiv \epsilon_{ikl} \partial_k \partial_l \omega_j(\mathbf{x}),$$  \hspace{1cm} (9.44)
where we have used the vector form of the rotation field $\omega_i = (1/2)\varepsilon_{ijk}\Omega_{jk}$, in order to save one index. For the general defect line along $L$, these densities have the form

$$\alpha_{ij}(x) = \delta_i(x; L) \left(b_i - \Omega_{jk}x_k\right), \quad (9.45)$$

$$\theta_{ij}(x) = \delta_i(x; L)\Omega_{ij}, \quad (9.46)$$

where $\Omega_i = (1/2)\varepsilon_{ijk}\Omega_{jk}$ is the Frank vector.

Note that in terms of the tensor $S_{ijk}$ of Eq. (9.33), the dislocation density (9.43) reads

$$\alpha_{ij}(x) \equiv \varepsilon_{ikl}S_{lkj}. \quad (9.47)$$

In (9.45) and (9.46) the rotation by $\Omega$ is performed around the origin. Obviously, the position of the rotation axis can be changed to any other point $x_0$ by a simple shift in the constant $b_j \to b'_j = (\Omega \times x_0)_j$. Then $\alpha_{ij}(x) = \delta_i(x; L) \left\{b'_i + (\Omega \times (x - x_0))_j\right\}$. Note that due to the identity

$$\partial_i\delta_i(x; L) = 0 \quad (9.48)$$

for closed lines $L$, the disclination density satisfies the conservation law

$$\partial_i\theta_{ij} = 0, \quad (9.49)$$

which implies that disclination lines are always closed. This is not true for media with a directional field, e.g., nematic liquid crystal. Such media are not considered here since they cannot be described by a displacement field alone. Differentiating (9.45) we find the conservation law for disclination lines $\partial_i\alpha_{ij} = -\Omega_{ij}\delta_i(x; L)$ which, in turn, can be expressed in the form

$$\partial_i\alpha_{ij} = -\varepsilon_{jkl}\theta_{kl}. \quad (9.50)$$

In terms of the tensor $S_{ijk}$, this becomes

$$\varepsilon_{jkl} \left(\partial_iS_{kli} + \partial_kS_{lmi} - \partial_lS_{kmi}\right) = -\varepsilon_{jkl}\theta_{kl}. \quad (9.51)$$

Indeed, inserting $S_{klj} = (1/2)\varepsilon_{kli}\alpha_{ij}$ from (9.47), this reduces to the conservation law (9.50) for the dislocation density.

From the linearity of the relations (9.43) and (9.44) in $u_j$ and $\omega_j$, respectively, it is obvious that these conservation laws remain true for any ensemble of infinitesimal defect lines. The conservation law (9.50) may, in fact, be derived by purely differential techniques from the first smoothness assumption (9.11). Using Stokes’ theorem, $\Delta u_{ij}$ can be expressed in the same way as $\Delta \omega_{ij}$ in (9.25), and by the same argument as the one used for $\omega_{ij}$ we conclude that the strain is an integrable function in all space and satisfies

$$(\partial_i\partial_k - \partial_k\partial_i) u_{ij}(x) = 0. \quad (9.52)$$
If we then look at $\alpha_{ij}$ in the general definition (9.43), rewrite it as

$$\alpha_{ij} = \varepsilon_{ikl} \partial_k \partial_j u_j = \varepsilon_{ikl} \partial_k \left( u_{ij} + \omega_{ij} \right)$$

and apply the derivative $\partial_i$, this gives directly (9.50).

In a similar way, the conservation law (9.49) can be derived by combining both smoothness assumptions (9.11) and (9.12). The first can be stated, via Stokes’ theorem, as an integrability condition for the derivative of strain, i.e.,

$$(\partial_l \partial_n - \partial_n \partial_l) \partial_k u_{ij}(x) = 0. \quad (9.54)$$

Let us recall that from the first assumption (9.11) we have concluded in (9.18) that $\partial_k \omega_{ij}(x)$ is also a completely smooth function across the surface $S$. Hence, $\partial_k \omega_{ij}$ must also satisfy the integrability condition

$$(\partial_l \partial_n - \partial_n \partial_l) \partial_k \omega_{ij}(x) = 0. \quad (9.55)$$

If we write down this relation three times, each time with $l,n,k$ exchanged cyclically, we find

$$\partial_l R_{nkij} + \partial_n R_{kl ij} + \partial_k R_{ln ij} = 0, \quad (9.56)$$

where $R_{nkij}$ is an abbreviation for the expression,

$$R_{nkij} = (\partial_n \partial_k - \partial_k \partial_n) \partial_j u_j(x). \quad (9.57)$$

Contracting $k$ with $i$ and $l$ with $j$ gives us

$$\partial_j R_{nijj} + \partial_n R_{ijij} + \partial_j R_{jinj} = 0. \quad (9.58)$$

Now we observe that because of (9.52), $R_{nkij}$ is anti-symmetric not only in $n$ and $k$ but also in $i$ and $j$ so that

$$2 \partial_j R_{nijj} - \partial_n R_{ijij} = 0.$$  

This, however, is the same as

$$2 \partial_j \left( \frac{1}{4} \varepsilon_{j pq} \varepsilon_{n kl} R_{pqkl} \right) = 0, \quad (9.59)$$

as can be verified using the identity

$$\varepsilon_{j pq} \varepsilon_{n kl} = \delta_{jn} \delta_{pk} \delta_{ql} + \delta_{jk} \delta_{pl} \delta_{qn} + \delta_{jl} \delta_{pn} \delta_{qk} - \delta_{jn} \delta_{pl} \delta_{qk} - \delta_{jk} \delta_{pn} \delta_{ql} - \delta_{jl} \delta_{pk} \delta_{qn}. \quad (9.60)$$

Recalling now the definition (9.57) and $\omega_n = (1/2) \varepsilon_{n kl} \partial_k u_l$, Eq. (9.59) becomes

$$2 \varepsilon_{j pq} \partial_p \partial_j \omega_m = 0, \quad (9.61)$$

and this is precisely the conservation law $\partial_j \theta_{jk} = 0$ for disclinations (9.49), which we wanted to prove. The educated reader will have noted the appearance of torsion and curvature in Eqs. (9.33) and (9.57), and Eqs. (9.51) and (9.56) as the linearized fundamental identity and the linearized Bianchi identity, to be discussed in detail in Sections 12.1 and 12.5 [see Eqs. (12.103) and (12.115), respectively].
There exists a simple mnemonic procedure for constructing the defect densities and their conservation laws. This we now explain.

Suppose we perform the Volterra cutting procedure on a closed surface $S$, dividing it mentally in two parts, joined along some line $L$ (see Fig. 9.16). On one part of $S$, say $S^+$, we remove material of thickness $b_i$ and on the other we add the same material. This corresponds to a simple translational movement of crystalline material by $b_i$, i.e., to a displacement field

$$ u_i(x) = -\delta(x; V)b_i, \quad (9.62) $$

where the $\delta$-function on a volume $V$ was defined in Eq. (4.29). By this transformation the elastic properties of the material are unchanged.

Consider now the distortion field $\partial_k u_i(x)$. Under (9.62), it changes by

$$ \partial_k u_i(x) \rightarrow \partial_k u_i(x) - \partial_i \delta(x; V)b_i. \quad (9.63) $$

The derivative of the $\delta$-function is singular only on the surface of the volume $V$. In fact, in Eq. (4.34) we already derived the formula $\partial_k \delta(x; V) = -\delta_k(x; S)$, so that (9.63) reads

$$ \partial_k u_i(x) \rightarrow \partial_k u_i(x) + \delta_k(x; S)b_i. \quad (9.64) $$

From this trivial transformation we can now construct a proper dislocation line by assuming $S$ to be no longer a closed surface but an open one, i.e. we may restrict $S$ to the shell $S^+$ with a boundary $L$. Then we can form the dislocation density

$$ \alpha_{il}(x) = \varepsilon_{ijk} \partial_j \partial_k u_l(x) = \varepsilon_{ijk} \partial_j \delta_k(x; S)b_l. \quad (9.65) $$
The superscript + was dropped. Using Stokes’ theorem for the $\delta_k(x; S)$-function in the form (4.23), this becomes simply

$$\alpha_{il}(x) = \delta_l(x; L)b_i. \tag{9.66}$$

For a closed surface, this vanishes.

For a general defect line, the starting point is the trivial Volterra operation of translating and rotating a piece of crystalline volume. This corresponds to a displacement field

$$u_l(x) = -\delta(x; V) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right). \tag{9.67}$$

If we now form the distortion, we find

$$\partial_k u_l(x) = \delta_k(x; S) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right) - \delta(x; V) \varepsilon_{lkq} \Omega_q. \tag{9.68}$$

In the expression it is still impossible to assume $S$ to be an open surface. If we, however, form the symmetric combination, the strain tensor

$$u_{kl} = \frac{1}{2} (\partial_k u_l + \partial_l u_k) = \frac{1}{2} \left[ \delta_k(x; S) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right) + (kl) \right] \tag{9.69}$$

is well defined for an open surface, in which case we shall refer to $u_{kl}$ as the plastic strain and denote it by $u^p_{kl}$. The field

$$\beta^p_{kl} = \delta_k(x; S) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right) \tag{9.70}$$

plays the role of a dipole density of the defect line across the surface $S$. It is usually called a plastic distortion. It is a single valued field (i.e., derivatives in front of it commute). In terms of $\beta^p_{kl}$, the plastic strain is simply

$$u^p_{kl} = \frac{1}{2} (\beta^p_{kl} + \beta^p_{lk}). \tag{9.71}$$

The full displacement field (9.67) is not defined for an open surface due to the $\delta(x; V)$ term. It is multi-valued. The dislocation density, however, is single valued. Indeed, we can easily calculate

$$\alpha_{il} = \varepsilon_{ijk} \partial_j \partial_k u_l(x) = \varepsilon_{ijk} \partial_j \left[ \delta_k(x; S) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right) - \delta(x; V) \varepsilon_{lkq} \Omega_q \right]$$

$$= \delta_l(x; L) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right), \tag{9.72}$$

and see that this is the same as (9.45).

Let us now turn to the disclination density $\theta_{pj} = \varepsilon_{pmn} \partial_m \partial_n \omega_j$. From (9.67) we find the gradient of the rotation field

$$\partial_n \omega_j = \frac{1}{2} \varepsilon_{jkl} \partial_n \partial_k u_l$$

$$= \frac{1}{2} \varepsilon_{jkl} \partial_n \left[ \delta_k(x; S) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right) - \delta(x; V) \varepsilon_{lkq} \Omega_q \right]$$

$$= \frac{1}{2} \varepsilon_{jkl} \partial_n \beta^p_{kl} + \delta_n(x; S) \Omega_j. \tag{9.73}$$
This gradient is defined for an open surface $S$ and called the field of plastic bend-twist and denoted by $\kappa_{nj}^p \equiv \partial_n \omega_j^p$. It is useful to define the plastic rotation

$$\phi_{nj}^p \equiv \delta_n (x; S) \Omega_j,$$

which plays the role of a dipole density for disclinations. With this, the plastic gradient of $\omega_j$ is given by

$$\kappa_{nj}^p = \partial_n \omega_j^p = \frac{1}{2} \varepsilon_{jkl} \partial_n \beta_{kl}^p + \phi_{nj}^p. \quad (9.75)$$

We can now easily calculate the disclination density:

$$\theta_{pj} = \varepsilon_{pmn} \partial_m \omega_j = \varepsilon_{pmn} \partial_m \kappa_{nj}^p = \frac{1}{2} \varepsilon_{jkl} \varepsilon_{pmn} \partial_m \partial_n \beta_{kl}^p + \varepsilon_{pmn} \partial_m \phi_{nj}^p. \quad (9.76)$$

In front of $\beta_{kl}^p$, the derivatives commute [see (9.70)] so that the first term vanishes. Use of Stokes’ theorem on the second term gives

$$\theta_{pj} = \varepsilon_{pmn} \partial_m \phi_{nj}^p = \delta_p (x; L) \Omega_j,$$

in agreement with (9.48).

Note that according to the second line of (9.72), the dislocation density can also be expressed in terms of $\beta_{kl}^p$ and $\phi_{li}^p$ as

$$\alpha_{il} = \varepsilon_{ijk} \partial_j \beta_{kl}^p + \delta_{il} \phi_{pp}^p - \phi_{li}^p. \quad (9.77)$$

In fact, this is a direct consequence of the decomposition (9.53), which can be written in terms of plastic strain and bend-twist fields as

$$\alpha_{ij} = \varepsilon_{ikl} \partial_k u_{lj}^p + \delta_{ij} \kappa_{qq}^p - \kappa_{ji}^p. \quad (9.78)$$

Expressing $u_{li}^p$ in terms of $\beta_{li}^p$ and $\kappa_{ij}^p$ in terms of $\phi_{ij}^p$ [see (9.71) and (9.75)] we find

$$\alpha_{ij} = \frac{1}{2} \varepsilon_{ikl} \partial_k \beta_{lj}^p + \delta_{ij} \phi_{qq}^p - \phi_{ji}^p + \frac{1}{2} \left( \varepsilon_{ijk} \partial_k \beta_{lj}^p + \delta_{ij} \varepsilon_{qkl} \partial_q \beta_{kl}^p - \varepsilon_{ikl} \partial_i \beta_{kl}^p \right).$$

But the quantity inside the parentheses is equal to $\frac{1}{2} \varepsilon_{ikl} \partial_k \beta_{lj}^p$, as can be seen from applying the identity

$$\delta_{ij} \varepsilon_{qkl} = \delta_{jq} \varepsilon_{ikl} + \delta_{jk} \varepsilon_{ilq} + \delta_{jl} + \delta_{ij} \varepsilon_{qkl}$$

to $\partial_q \beta_{kl}$. Thus $\alpha_{ij}$ takes again the form (9.77).
9.11 Defect Gauge Invariance

A given defect distribution can be derived from many plastic strains and rotations. This is an obvious consequence of the freedom in choosing the Volterra surfaces for the construction of the defect lines. They run along the boundary lines of these surfaces. The ambiguity can be formulated mathematically as a gauge freedom. This has already been seen in Subsection 4.6 in the context of a gradient representation of the magnetic field of a current loop. The current density was given in Eq. (4.91) by a curl of a \( \delta \)-function on a surface:

\[
j(x) = I \nabla \times \delta(x; S),
\]

and this representation was invariant under a gauge transformation, which arises when shifting the surface \( S \) to a new position \( S' \) with the same boundary line:

\[
\delta(x; S) \rightarrow \delta(x; S') = \delta(x; S) - \nabla \delta(x; V).
\]

This gauge invariance can be found in all expressions for defect densities which involve \( \delta \)-functions over Volterra surfaces \( S \). The relevant gauge transformations can be derived from the basic translational movement (9.63) which does not change the defect configurations but does change the plastic fields. Thus we perform what we may call the basic trivial Volterra transformation

\[
u_l(x) \rightarrow u_l(x) - \delta(x; V) b_l.
\]

According to Eq. (9.64), the plastic distortion field changes by

\[
\beta_{kl}^p \rightarrow \beta_{kl}^p - \partial_k \delta(x; V) b_l.
\]

This has the typical form of a gauge transformation, where the gradient of an arbitrary field is added to a gauge field. In the context of defects there is, however, an important difference: the function is not completely arbitrary but contains a \( \delta \)-function over an arbitrary volume. The Burgers vector accounts for the lattice properties.

Comparison with Eqs. (9.84) and (9.72) shows that the transformation (9.82) corresponds to a shift of the Volterra surface \( S \rightarrow S' \) if the defect is a pure dislocation line. If it contains also a Frank vector \( \Omega_k \), we must supplement the translation (9.62) by a rotation and transform

\[
u_l(x) \rightarrow u_l(x) - \delta(x; V) b_l - \epsilon_{lmn} \delta(x, V) \Omega_m x_n.
\]

Then we find from (9.84) that

\[
\partial_k u_l(x) \rightarrow \partial_k u_l(x) - \partial_k \delta(x; V) \left( b_l + \epsilon_{lqr} \Omega_q x_r \right) - \delta(x; V) \epsilon_{lqr} \Omega_q,
\]

so that the plastic strain transforms like

\[
u_{kl}^p \rightarrow u_{kl}^p - \frac{1}{2} \left[ \partial_k \delta(x; V) \left( b_l + \epsilon_{lqr} \Omega_q x_r \right) + (kl) \right].
\]
Comparison with (9.69) shows that this corresponds, via Eq. (9.80), precisely to a shift of the Volterra surface $S$ to $S'$. The change of the plastic distortion $\beta^p_{kl}$ in Eq. (9.70) is, however, not yet obtained, due to the last term in (9.84). This is removed if we accompany (9.83) by a trivial Volterra transformation of the rotation field

$$\omega_j = \frac{1}{2} \varepsilon_{jkl} \partial_k u_l \rightarrow \omega_j - \delta(x; V) \Omega_j.$$  

(9.86)

Adding this to (9.84), the last term is removed and we obtain the combined transformation

$$\partial_k u_l(x) \rightarrow \partial_k u_l(x) - \partial_k \delta(x; V) \left( b_l + \varepsilon_{lqr} \Omega_q x_r \right).$$  

(9.87)

This is precisely the gauge transformation of the plastic distortion (9.70) if the Volterra surface is shifted from $S$ to $S'$. Simultaneously, the gradient of the rotation field $\omega$ in Eq. (9.73) undergoes the defect gauge transformation

$$\partial_n \omega_j \rightarrow \partial_n \omega_j - \partial_n \delta(x; V) \Omega_j,$$  

(9.88)

corresponding to a gauge transformation of the plastic rotation (9.74) transforms like

$$\phi^p_{nj} \rightarrow \phi^p_{nj} - \partial_n \delta(x; V) \Omega_j.$$  

(9.89)

The vortex gauge transformations (9.82), (9.85), and (9.89) summarize the invariance of a defect line under the shift of the Volterra surface.

### 9.12 Branching Defect Lines

We recall that from the geometric point of view, the conservation laws state that disclination lines never end and dislocations end at most at a disclination line. Consider, for example, a branching configuration where a line $L$ splits into two lines $L$ and $L'$. Assign an orientation to each line and suppose that their disclination density is

$$\theta_{ij}(x) = \Omega_i \delta_j(x; L) + \Omega'_i \delta_j(x; L') + \Omega''_i \delta_j(x; L''),$$  

(9.90)

with their dislocation density being

$$\alpha_{ij}(x) = \delta_i(x; L) \left\{ b_j + [\Omega \times (x - x_0)]_j \right\} + \delta_i(x; L') \left\{ b_j' + [\Omega' \times (x - x'_0)]_j \right\} + \delta_i(x; L'') \left\{ b_j'' + [\Omega'' \times (x - x''_0)]_j \right\}. $$  

(9.91)

The conservation law $\partial_i \theta_{ij} = 0$ then implies that the Frank vectors satisfy the equivalent of Kirchhoff’s law for currents

$$\Omega_j + \Omega''_j = \Omega'_j.$$  

(9.92)
This follows directly from the identity for lines
\[ \partial_i \delta_i(x; L) = \int ds \frac{d\vec{x}}{ds} \delta^{(3)}(x - \vec{x}(3)) = \delta^{(3)}(x - x_i) - \delta^{(3)}(x - x_f), \]
where \(x_i\) and \(x_f\) are the initial and final points of the curve \(L\). The conservation law \(\partial_i \alpha_{ij} = \varepsilon_{ikl} \theta_{kl}\), on the other hand, gives
\[ b_i - [\Omega \times (x - x_0)]_i + b'_i - [\Omega' \times (x - x'_0)]_i = b'_i - [\Omega \times (x' - x'_0)]_i. \quad (9.93) \]
If the same position is chosen for all rotation axes, the Burgers vectors \(b_i\) satisfy again a Kirchhoff-like law:
\[ b_i + b''_i = b'_i. \quad (9.94) \]
But Burgers vectors can be compensated for by different rotation axes, for example, \(L'\) and \(L''\) could be pure disclination lines with different axes through \(x'_0, x''_0\) and \(L'\) a pure dislocation line with \(x' = -\Omega' \times (x'_0 - x''_0)\) which ends on \(L', L''\). Equation (9.92) renders different choices equivalent.

### 9.13 Defect Density and Incompatibility

As far as classical linear elasticity is concerned, the information contained in \(\alpha_{ij}\) and \(\theta_{ij}\) can be combined into a single symmetric tensor, called the defect density \(\eta_{ij}(x)\) [In higher gradient elasticity this is no longer true; see chapter 18.]. It is defined as the double curl of the strain tensor
\[ \eta_{ij}(x) \equiv \varepsilon_{ikl} \varepsilon_{jmn} \partial_k \partial_m u_{ln}(x). \quad (9.95) \]
In order to see its relation with \(\alpha_{ij}\) and \(\theta_{ij}\), we take (9.43) and contract the indices \(i\) and \(j\), obtaining
\[ \alpha_{ii} = 2 \partial_i \omega_i. \quad (9.96) \]
Using this, (9.53) can be written in the form
\[ \varepsilon_{ikl} \partial_k u_{ln} = \partial_n \omega_i - \left(-\alpha_{in} + \frac{1}{2} \delta_{in} \alpha_{kk} \right). \quad (9.97) \]
The expression in parentheses was first introduced by Nye and called contortion\(^2\)
\[ K_{ni} \equiv -\alpha_{in} + \frac{1}{2} \delta_{in} \alpha_{kk}. \quad (9.98) \]
\(^2\)In terms of the plastic quantities introduced in the last section the plastic part of \(K_{ij}\) reads
\[ K^p_{ij} = -\varepsilon_{ikl} \partial_k \beta^p_{li} + \frac{1}{2} \delta_{ij} \varepsilon_{nkli} \partial_k \beta^p_{ln} + \phi^p_{ij}. \]
The inverse relation is
\[ \alpha_{ij} = -K_{ji} + \delta_{ij}K_{kk}. \] (9.99)

Multiplying (9.95) by \( \varepsilon_{jmn} \partial_m \), we find with (9.44)
\[ \eta_{ij} = \varepsilon_{jmn} \varepsilon_{ikl} \partial_k u_{ln} = \varepsilon_{jmn} \partial_m \omega_i - \varepsilon_{jmn} \partial_m K_{ni} \]
\[ = \theta_{ij} - \varepsilon_{jmn} \partial_m K_{ni}. \] (9.100)

The final expression is not manifestly symmetric. Let us verify that it is, nevertheless. Contracting it with the antisymmetric tensor \( \varepsilon_{lij} \), we find
\[ \varepsilon_{lij} \theta_{ij} \partial_l K_{ii} - \partial_i K_{li} = \varepsilon_{lij} \theta_{ij} \theta_{ij} + \partial_i \alpha_{il}. \] But this vanishes due to the conservation law (9.50) for the dislocation density. Thus \( \eta_{ij} \) is symmetric.

There is yet another version of the decomposition (9.100) which is obtained after applying the identity
\[ \varepsilon_{ijn} \delta_{mq} + \varepsilon_{jmn} \delta_{iq} + \varepsilon_{min} \delta_{jq} = \varepsilon_{ijm} \delta_{nq} \] to \( \partial_m \alpha_{qn} \) giving
\[ \varepsilon_{njm} \partial_m \left( \alpha_{in} - \frac{1}{2} \delta_{in} \alpha_{kk} \right) = -\frac{1}{2} \partial_m \left( \varepsilon_{mnj} \alpha_{in} + (ij) + \varepsilon_{ijn} \alpha_{mn} \right). \] Hence
\[ \eta_{ij} = \theta_{ij} - \frac{1}{2} \partial_m \left( \varepsilon_{min} \alpha_{jn} + (ij) - \varepsilon_{ijn} \alpha_{mn} \right). \] (9.101)

This type of decomposition will be encountered in the context of general relativity later in Part IV.

The double curl operation is a useful generalization of the curl operation on vector fields to symmetric tensor fields. Recall that the vanishing of a curl of a vector field \( \mathbf{E} \) implies that \( \mathbf{E} \) can be written as the gradient of a scalar potential \( \phi(x) \) which satisfies the integrability condition \( (\partial_i \partial_j - \partial_j \partial_i) \phi(x) = 0 \):
\[ \nabla \times \mathbf{E} = 0 \Rightarrow \mathbf{E}_i = \partial_i \phi(x). \] (9.102)

The double curl operation implies a similar property for the symmetric tensor, as was shown a century ago by Riemann and by Christoffel. If the double curl of a symmetric tensor field vanishes everywhere in space, his field can be written as the strain of some displacement field \( u_i(x) \) which is integrable in all space [i.e., it satisfies (9.34)]. We may state this conclusion briefly as follows:
\[ \varepsilon_{ikl} \varepsilon_{jmn} \partial_k \partial_m u_{ln}(x) = 0 \Rightarrow u_{ij} = \frac{1}{2} \left( \partial_i u_j + \partial_j u_i \right). \] (9.103)

If the double curl of \( u_{ln}(x) \) is zero one says that \( u_{ln}(x) \) is compatible with a displacement field and calls the double curl the incompatibility, i.e.,
\[ (\text{inc } u)_{ij} \equiv \varepsilon_{ikl} \varepsilon_{jmn} \partial_k \partial_m u_{ln}. \] (9.104)

The proof of statement (9.103) follows from (9.102) for a vector field: we simply observe that every vector field \( V_k(x) \) vanishing at infinity and satisfying the integrability condition \( (\partial_i \partial_j - \partial_j \partial_i) V_k(x) = 0 \) can be decomposed into transverse and
longitudinal pieces, namely, a gradient whose curl vanishes and a curl whose gradient vanishes,

\[ V_i = \partial_j \varphi + \varepsilon_{ijk} \partial_j A_k, \quad (9.105) \]

both fields \( \varphi \) and \( A_k \) being integrable. Explicitly these are given by

\[ \varphi = \frac{1}{\partial^2} \partial_i V_i, \quad (9.106) \]

\[ A_k = -\frac{1}{\partial^2} \varepsilon_{klm} \partial_l V_m + \partial_k C, \quad (9.107) \]

where \( 1/\partial^2 \) is a short notation for the Coulomb Green function \((1/\partial^2)(x, x')\) which acts on an arbitrary function in the usual way:

\[ -\frac{1}{\partial^2} f(x) \equiv \int d^3 x \frac{1}{4\pi |x - x'|} f(x') \quad (9.108) \]

is the Coulomb Green function. Note that the field \( A_k \) is determined by (9.107) only up to an arbitrary pure gradient \( \partial_k C \).

By repeated application of this formula, we find the decompositions of an arbitrary, not necessarily symmetric, tensor \( u_{il} \):

\[ u_{il} = \partial_i \varphi_l' + \varepsilon_{ijk} \partial_j A_k' \]

\[ = \partial_i \varphi_l' + \varepsilon_{ijk} (\partial_j \varphi_k + \varepsilon_{lmn} \partial_m A_{kn}). \quad (9.109) \]

Setting

\[ \varphi''_i \equiv \varepsilon_{ijk} \partial_j \varphi_k, \quad (9.110) \]

this may be cast as

\[ u_{il} = \partial_i \varphi_l' + \partial_i \varphi''_l + \varepsilon_{ijk} \varepsilon_{lmn} \partial_j \partial_m A_{kn}. \quad (9.111) \]

For the special case that \( u_{il} \) is symmetric we can symmetrize this result and decompose it as

\[ u_{il} = \partial_i u_j + \partial_j u_i + \varepsilon_{ijk} \varepsilon_{lmn} \partial_j \partial_m A_{kn}^S, \quad (9.112) \]

where

\[ u_i = \frac{1}{2} \left( \varphi'_i + \varphi''_i \right), \quad (9.113) \]

and \( A_{kn}^S \) is the symmetric part of \( A_{kn} \), both being integrable fields. The first term in (9.112) has zero incompatibility, the second has zero divergence when applied to either index.
In the general case, i.e., when there is no symmetry, we can use the formulas (9.106), (9.107) twice and determine the fields $\phi'^l, \phi''^i, A_{kn}$ as follows:

$$\phi'^l = \frac{1}{\partial^2} \partial_k u_{kl},$$  \hspace{1cm} (9.114)  

$$A'_{kl} = -\frac{1}{\partial^2} \varepsilon_{kpq} \partial_k u_{ql} + \partial_k C'_l,$$  \hspace{1cm} (9.115)  

$$\phi^i_k = -\frac{1}{\partial^4} \varepsilon_{klm} \varepsilon_{nps} u_{mpq} \partial_k \left( \frac{1}{\partial^2} \varepsilon_{njl} \partial_j C'_l \right) + \partial_n D_k,$$  \hspace{1cm} (9.116)  

$$A_{kn} = -\frac{1}{\partial^4} \varepsilon_{klm} \varepsilon_{nps} u_{mpq} \partial_k \left( \frac{1}{\partial^2} \varepsilon_{njl} \partial_j C'_l \right) + \partial_n D_k,$$  \hspace{1cm} (9.117)  

so that from (9.109)

$$\phi''^i = -\frac{1}{\partial^4} \partial_i \partial_p \partial_q u_{pq} + \frac{1}{\partial^2} \partial_i u_{il}.$$  \hspace{1cm} (9.118)  

Reinserting this into decomposition (9.111) we find the identity

$$u_{il} = \frac{1}{\partial^2} (\partial_i \partial_k u_{kl} + \partial_k \partial_k u_{ik}) - \frac{1}{\partial^2} \partial_i \partial_i \left( \partial_p \partial_q u_{pq} \right)$$

$$+ \frac{1}{\partial^4} \varepsilon_{ijkl} \varepsilon_{lmn} \partial_j \partial_m \left( \varepsilon_{kpq} \varepsilon_{nps} \partial_p \partial_q u_{rs} \right),$$  \hspace{1cm} (9.119)  

which is valid for any tensor of rank two. This may be verified by working out the contractions of the $\varepsilon$ tensors.

While the statements (9.102) and (9.103) for vector and tensor fields are completely analogous to each other, it is important to realize that there exists an important difference between the two. For a vector field with no curl, the potential can be calculated uniquely (up to boundary conditions) from

$$\varphi = \frac{1}{\nabla^2} \partial_i E_i.$$  \hspace{1cm} (9.120)  

This is no longer true, however, for the compatible tensor field $u_{ij}$. The point of departure lies in the non-uniqueness of functions $\varphi'^i$ and $\varphi''^i$ in the decomposition (9.111). They are determined only modulo a common arbitrary local rotation field $\omega_i(x)$. In order to see this we perform the replacements

$$\partial_i \varphi'^i(x) \rightarrow \partial_i \varphi'^i(x) + \varepsilon_{ijq} \omega_q(x),$$  \hspace{1cm} (9.121)  

$$\partial_i \varphi''^i(x) \rightarrow \partial_i \varphi''^i(x) + \varepsilon_{ijq} \omega_q(x),$$  \hspace{1cm} (9.122)  

and see that (9.111) is still true. The field (9.113) is only a particular example of a displacement field which has the strain tensor equal to the given $u_{kl}$

$$u^0_{kl} = \frac{1}{2} \left( \partial_k u^0_{kl} + \partial_l u^0_{kl} \right) = u_{kl}.$$  \hspace{1cm} (9.123)
This displacement field may not, however, be the true displacement field $u_l(x)$ in the crystal, which also satisfies
\[ \frac{1}{2} (\partial_k u_l + \partial_l u_k) = u_{kl}. \] (9.124)
In order to find the latter, we need additional information on the rotation field
\[ \omega_{kl} = \frac{1}{2} (\partial_k u_l - \partial_l u_k). \] (9.125)
We must know both $u_{kl}(x)$ and $\omega_{kl}(x)$ to calculate
\[ \partial_k u_l(x) = u_{kl}(x) + \omega_{kl}(x) \] (9.126)
and solve this equation for $u_l(x)$.

In order to make use of this observation we have to be sure that $\omega_i = \frac{1}{2} \varepsilon_{ijk} \omega_{jk}$ can be written as the curl of a displacement field $u_i(x)$. This is possible if
\[ \partial_i \omega_i = \varepsilon_{ijk} \partial_j \partial_j u_k = 0, \] (9.127)
which implies that [see (9.96)]
\[ \alpha_{ii}(x) = 0. \] (9.128)
In later discussions we shall be confronted with the situation in which $u_{kl}$ and $\partial_j \omega_j$ are both given. In order to obtain $\omega_i$ from the latter we have to make sure that $\omega_i$ is an integrable field, which is assured by the constraint
\[ \theta_{ij} = \varepsilon_{ijk} \partial_k \partial_j \omega_j = 0. \] (9.129)
Thus we can state the following important result: Suppose a crystal is subject to a strain $u_{kl}(x)$ and a rotational distortion $\omega_{kl}(x)$. There exists an associated single-valued displacement field $u_l(x)$, if and only if the crystal possesses a vanishing defect density $\eta_{ij}(x)$, a vanishing disclination density $\theta_{ij}(x)$, and a vanishing $\alpha_{ii}(x)$, i.e.,
\[ \eta_{ij}(x) = 0, \quad \theta_{ij}(x) = 0, \quad \alpha_{ii}(x) = 0. \] (9.130)
Relation (9.101) implies that this is true if only two of these densities vanish:
\[ \eta_{ij}(x) = 0, \quad \alpha_{ij}(x) = 0, \] (9.131)
or
\[ \theta_{ij}(x) = 0, \quad \alpha_{ij}(x) = 0. \] (9.132)
Note that it is possible to introduce, into a given elastically distorted crystal, nonzero rotational and translational defects in such a way that $\theta_{ij}$ and $\alpha_{ij}$ in (9.101) cancel each other. Then the elastic distortions do, in fact, remain unchanged. It may, in fact, remain unchanged. The local rotation field, however, can be changed drastically. In particular it may no longer be integrable.
Notes and References


[3] For details on vortex lines see in
10

Defect Melting

In Chapter 5 we have seen that the phase transitions in superfluid helium and in superconductors can be explained by a the proliferation of vortex lines at the critical temperature. A similar proliferation mechanism of dislocation and disclination lines can be shown to lead to the melting of crystals.

10.1 Specific Heat

The specific heat of solids has several parallels with the specific heat of the $\lambda$-transition. For low temperature it starts out like $T^3$, as a signal for massless excitations [see Fig. 10.1]. These are the longitudinal and transverse phonons, which are the Goldstone modes caused by the fact that the crystalline ground state breaks spontaneously the translational symmetry of the Hamiltonian. For higher temperatures the specific heat saturates at a value $6 \times k_B N/2$ in accordance with the Dulong-Petit rule, which requires the value $1/2$ for each particle and harmonic degree of freedom (3 potential and three kinetic degrees).

The transition between the two regimes lies at the Debye temperature $\Theta_D$ which is determined by the longitudinal and transversal sound velocities $c_s^L$, $c_s^T$ and the particle density $n \equiv N/V$. For one atom per lattice cell end three equal sound velocities, it is given by

$$\Theta_D = 2\pi \frac{\hbar c_s}{k_B} \left( \frac{2n}{4\pi} \right)^{1/3}.$$  \hspace{1cm} (10.1)

The internal energy is given by the universal Debye function

$$D(z) = \frac{3}{z^3} \int_0^z \frac{x^3}{e^x - 1}$$  \hspace{1cm} (10.2)

as

$$U = 3Nk_BT D(\Theta_D/T).$$  \hspace{1cm} (10.3)

The specific heat follows from this:

$$C = \frac{\partial U}{\partial T} = 3Nk_BT \left[ D(\Theta_D/T) - (\Theta_D/T)D'(\Theta_D/T) \right].$$  \hspace{1cm} (10.4)
Using the limiting behavior

\[ D(z) = \begin{cases} \frac{\pi^2}{5z^3} - 3e^{-z} + \ldots & \text{for } z \gg 1, \\ 1 - \frac{3}{8}z^2 + \ldots & \text{for } z \ll 1, \end{cases} \] (10.5)

we find

\[ C = 3Nk_B \begin{cases} \frac{4\pi^4}{5} \left( \frac{T}{\Theta_D} \right)^3, & \text{for } T \ll \Theta_D, \\ 1, & \text{for } T \gg \Theta_D. \end{cases} \] (10.6)

The result agrees well with experiments as shown in Fig. 10.1.

\[ \frac{C_V}{3Nk_B} \]

**Figure 10.1** Specific heat of various solids. By plotting the data against the ratio \( T/\Theta_D \), where \( \Theta_D \) is the Debye temperature, the data fall on a universal curve. The insert lists \( \Theta_D \)-values and melting temperature \( T_m \).

### 10.2 Elastic and Plastic Energies

In crystals, the elastic energy is usually expressed in terms of a material *displacement field* \( u_i(x) \) as

\[ E = \int d^3x \left[ \mu u_{ij}^2(x) + \frac{\lambda}{2} u_i^2(x) \right], \] (10.7)

where \( \mu \) is the shear module, \( \lambda \) the Lamé constant, and

\[ u_{ij}(x) = \frac{1}{2} [\partial_i u_j(x) + \partial_j u_i(x)] \] (10.8)
is the strain tensor (9.69). The elastic energy goes to zero for infinite wave length since in this limit \( u_i(x) \) reduces to a pure translation and the energy of the system is translationally invariant. The crystallization process causes a spontaneous breakdown of the translational symmetry of the system. The elastic distortions describe the Nambu-Goldstone-modes resulting from this symmetry breakdown.

As discussed in the previous chapter, a crystalline material always contains defects. In their presence, the elastic energy is

\[
E = \int d^3x \left[ \mu (u_{ij} - u^p_{ij})^2 + \frac{\lambda}{2} (u_{ii} - u^p_{ii})^2 \right], \tag{10.9}
\]

where \( u^p_{ij} \) is so-called plastic strain tensor (9.71) describing the defects.

The above energy is the continuum limit of a lattice energy which we shall assume, for mathematical simplicity, to be simple cubic of spacing \( 2\pi \). Then the energy of \( u_i(x) \) and \( u_i(x) + 2\pi N_i(x) \) must be indistinguishable for any integer-valued field \( N_i(x) \), which correspond to permutations of the lattice sites.

The plastic strain tensor allows for this ambiguity. It guarantees the defect gauge invariance of the energy (10.9). By analogy with the superfluid in Eq. (5.175), we may define the expectation value

\[
O_i \equiv \langle O_i(x) \rangle = \langle e^{u_i(x)} \rangle \tag{10.10}
\]
as an order parameter of the system. It will be nonzero in the crystalline phase since \( u_i(x) \) fluctuates around zero, and zero in the molten state in which \( u_i(x) \) fluctuates through the entire crystal.

The plastic distortions contain three types of surfaces where the displacement field \( u_i(x) \) jumps by \( 2\pi \), one for each lattice direction. They are characterized by the three Burgers vectors \( b^{(1),(2),(3)} \) and described by the plastic distortion of Eq. (9.70)

\[
\beta^p_{il}(x) = \delta_i(x; S) b_l, \tag{10.11}
\]

where \( b_l \) are the components of any of the three Burgers vectors \( b^{(i)} \). The irrelevant surfaces \( S \) are the Volterra surfaces of the dislocation lines.

We can now calculate the partition function of lattice fluctuations governed by the energy (10.9) from the functional integral and the sum over all Volterra surfaces

\[
Z = \int \mathcal{D}u \sum_S e^{-H/k_BT} = e^{-\beta F}, \tag{10.12}
\]

where \( \beta \equiv 1/k_BT \). This is most easily done by Monte-Carlo simulations. The resulting specific heat near the melting transition is shown in Fig. 10.2.

It is possible to write down an elastic energy which disentangles dislocations and disclinations by including higher gradients of the displacements field. This energy reads [1, 2]

\[
E = \mu \int d^3x \left[ (u_{ij} - u^p_{ij})^2 + \ell^2 \left( \partial_i \omega_j - \kappa^p_{ij} \right)^2 \right]. \tag{10.13}
\]
The parameter $\ell$ is the length scale over which the crystal is rotationally stiff.

The partition function contains integrals over $u_i$ and sums over the jumping surfaces of dislocations and disclinations. By integrating out the $u_i$-fields, one obtains a Biot-Savart type of interaction energy between the defect lines in which dislocation line elements interact with each other via a Coulomb potential, whereas disclination line elements interact via a linearly rising potential.

It is again possible to eliminate the jumping surfaces from the partition function by introducing conjugate variables and associated stress gauge fields. For this we rewrite the elastic action of defect lines as \cite{1, 3}

$$E = \int d^3x \left[ \frac{1}{4\mu} \left( \sigma_{ij} + \sigma_{ji} \right)^2 + \frac{1}{8\ell^2} \tau_{ij}^2 \right. $$
$$+ i \sigma_{ij} \left( \partial_i u_j - \epsilon_{ijk} \omega_k - \beta_{ij}^p \right) + i \tau_{ij} \left( \partial_i \omega_j - \tilde{\omega}_{ij}^p \right) \right], \quad (10.14)$$

and form the partition function by integrating over $\sigma_{ij}, \tau_{ij}, u_i, \omega_j$ and summing over all jumping surfaces $S$ in the plastic fields. A functional integral over the antisymmetric part of $\sigma_{ij}$ has been introduced to obtain an independent integral over $\omega_i$. The integral over the antisymmetric part of $\sigma_{ij}$, enforce the relation $\omega_i = \frac{1}{2} \epsilon_{ijk} (\partial_j u_k + \beta_{ij}^p)$. By integrating out $\omega_j$ and $u_i$, we find the conservation laws

$$\partial_i \sigma_{ij} = 0, \quad \partial_i \tau_{ij} = -\epsilon_{jkl} \sigma_{kl}. \quad (10.15)$$

These are dual to the conservation laws for dislocation and disclination densities (9.50) and (9.49).
10.2 Elastic and Plastic Energies

Figure 10.3 Separation of first-order melting transition into two successive Kosterlitz-Thouless transitions in two dimension when increasing the length scale $\ell$ of rotational stiffness of the defect model (after Ref. [1]).

The conservation laws are guaranteed as Bianchi identities by introducing the stress gauge fields $A_{ij}$ and $h_{ij}$ and writing

$$\sigma_{ij} = \epsilon_{ikl} \partial_k A_{lj},$$
$$\tau_{ij} = \epsilon_{ikl} \partial_k h_{lj} + \delta_{ij} A_{ll} - A_{ji}. \tag{10.16}$$

This allows us to reexpress the energy as

$$E = \int d^3x \left[ \frac{1}{4} \left( \sigma_{ij} + \sigma_{ji} \right)^2 + \frac{1}{8\ell^2} \tau_{ij}^2 + A_{ij} \alpha_{ij} + h_{ij} \theta_{ij} \right]. \tag{10.17}$$

The stress gauge fields couple locally to the defect densities which are singular on the boundary lines of the Volterra surfaces. In the limit of a vanishing length scale $\ell$, $\tau_{ij}$ is forced to be identically zero and (10.16) allows us to express $A_{ij}$ in terms of $h_{ij}$. Then the energy becomes

$$E = \int d^3x \left[ \frac{1}{4} \left( \sigma_{ij} + \sigma_{ji} \right)^2 + h_{ij} \eta_{ij} \right], \tag{10.18}$$

where the defect density $\eta_{ij}$ contains dislocation and disclination lines.
Depending on the length parameter $\ell$ of rotational stiffness, the defect system was predicted to have either a single first-order transition (for small $\ell$), of two successive continuous melting transitions. In the first transitions, dislocation lines proliferate and destroy the translational order, in the second transition, disclination lines proliferate and destroy the rotational order [2] (see Figs. 10.3 and 10.4).

The existence of two successive continuous transitions was conjectured a long time ago [4, 5, 6, 7] for two dimensional melting, where these transitions would be of the Kosterlitz-Thouless type. However, the simplest lattice defect models constructed to illustrate this behavior displayed only a single first-order transition [8]. Only after introducing the angular stiffness $\ell$ was it possible to separate the first-order melting transition into two successive Kosterlitz-Thouless transitions [9], thus confirming theoretical predictions made by the author in 1983 [3]. The dependence on $\ell$ is shown in Figs. 10.3 and 10.4.

**Notes and References**


Relativistic Mechanics in Curvilinear Coordinates

The basic idea which led Einstein to his formulation of the theory of gravitation in terms of curved spacetime was the observation by Galileo Galilei (1564-1642) all bodies would fall with equal velocity if in the absence of air friction. This observation was confirmed with a higher accuracy by C. Huygens (1629-1695). In 1889 R. Eötvös found a simple trick to remove the air friction completely [1]. This enabled him to limit the relative difference between the falling speeds of wood and platinum to one part in $10^9$. This implies that the inertial mass $m$ which governs the acceleration of a body if subjected to a force $f(t)$ in Newton’s equation of motion

$$m \ddot{x}(t) = f(t),$$

which appears on the left-hand side of the equations of motion (1.2), and the gravitational mass on the right-hand side of Eq. (1.2) cannot differ by more than this extremely small amount.

11.1 Equivalence Principle

Einstein considered the result of the Eötvös experiment as evidence that inertial and gravitational masses are exactly equal. From this he concluded that the motion of all point particles under the influence of a gravitational field can be described completely in geometric terms. The basic thought experiment which led him to this conclusion consisted in imagining an elevator in a large sky scraper to fall freely. Since all bodies in it would fall with the same speed, they would appear weightless. Thus, for an observer inside the cabin, the gravitational attraction to the earth would have disappeared. Einstein concluded that gravitational forces can be removed by acceleration. This is the content of the equivalence principle.

Mathematically, the cabin is just an accelerating coordinate frame. If the original spacetime coordinates with gravity are denoted by $x^\mu$, the coordinates of the small cabin are given by a function $x^\mu(x^a)$. Hence the equivalence principle states that the behavior of particles under the influence of gravitational forces can be found
by going to a new coordinate frame $x^a(x^\mu)$ in which the motion within the cabin proceeds without gravity.

There is a converse way of stating this principle. Given an inertial frame $x^a$, we can simulate a gravitational field at a point by going to a small cabin with $x^\mu$, which is accelerated with respect to the inertial frame $x^a$. In the coordinates $x^\mu$, the motion of the particle looks the same as if a gravitational field were present.

This suggests a simple way of finding the equations of motion of a point particle in a gravitational field: one must simply transform the known equations of motion in an inertial frame to arbitrary curvilinear coordinates $x^\mu$. When written in general coordinates $x^\mu$, the flat-spacetime equations must be valid also in the presence of gravitational fields.

In formulating the equivalence principle it must be realized that by a coordinate transformation the gravitational field can only be removed at a single point. In a falling cabin, a point particle will remain at the same place only if it resides at the center of mass of the cabin. Particles in the neighborhood of this point will move slowly away from this point. The force causing this are called tidal forces. They are the same forces which give rise to the tidal waves of the oceans. Earth and moon circle around each other and their center of mass circles around the sun. The center of mass is in “free fall”, the gravitational attraction proportional to the gravitational mass being canceled by the centrifugal force proportional to the inertial mass. Any point on the earth which lies farther from the sun than the center of mass is pulled outwards by the centripetal force, those which lie closer are pulled inwards by the gravitational force.

It is important to realize that the existence of tidal forces makes it impossible to simulate gravitational forces by coordinate transformation in quantum mechanics. Due to Heisenberg’s uncertainty relation, quantum particles can never be localized to a point but always occur in the form of wave packets. These flow apart and are therefore increasingly sensitive to the tidal forces. If one wants to remove the gravitational forces for a wave packet, multivalued coordinate transformations will be necessary of the type used in the last chapter to create defects. These will supply us with a quantum equivalence principle to be derived in Chapter 12.

### 11.2 Free Particle in General Coordinates Frame

As a first application of Einstein’s equivalence principle, consider the action (2.19) of a free massive point particle in Minkowski spacetime:

$$\mathcal{A} = -mc \int_{\sigma_a}^{\sigma_b} ds$$

where

$$ds = d\sigma \sqrt{g_{ab} \dot{x}^a(\sigma) \dot{x}^b(\sigma)}$$

is the increment of the invariant length along the path $x(\sigma)$. The quantity

$$\tau \equiv s/c$$

(11.4)
is called the *proper time*.

A free particle moves along a straight line

\[ \ddot{x}^a(\tau) = 0, \quad (11.5) \]

which is the shortest spacetime path between initial and final points. This path extremizes the action:

\[ \delta m \mathcal{A} = 0. \quad (11.6) \]

When going to an arbitrary curvilinear description of the same Minkowski space in terms of coordinates \( x^\mu \) carrying Latin indices

\[ x^\mu = x^\mu(x^a), \quad (11.7) \]

the invariant length \( ds \) is given by

\[ ds = d\sigma \sqrt{g_{\mu\nu}(x(\sigma))\dot{x}^\mu(\sigma)\dot{x}^\nu(\sigma)}. \quad (11.8) \]

The \( 4 \times 4 \) spacetime-dependent matrix

\[ g_{\mu\nu}(q) = g_{ab} \frac{\partial x^a}{\partial x^\mu} \frac{\partial x^b}{\partial x^\nu} \quad (11.9) \]

plays the role of a *metric* in the spacetime. Note that the inverse metric is given by

\[ g^{\mu\nu}(x) = g^{ab} \frac{\partial x^\mu}{\partial x^a} \frac{\partial x^\nu}{\partial x^b}. \quad (11.10) \]

Since spacetime has not really changed, only its parametrization, the path is still straight. The equation of motion in the new curvilinear coordinates \( x^\mu \) can be found in two ways. One is to simply transform the free equation of motion in Minkowski space (11.5) to curvilinear coordinates. This will be done at the end of this section.

We begin with the more complicated but instructive way by extremizing the transformed action

\[ m \mathcal{A} = \int_{\sigma_a}^{\sigma_b} d\sigma \, m L (\dot{x}^\mu(\sigma)), \quad (11.11) \]

with the transformed Lagrangian [compare (2.19)]

\[ mL (\dot{x}^\mu(\sigma)) = -mc \frac{ds}{d\sigma} = -mc \left[ g_{\mu\nu}(x(\sigma))\dot{x}^\mu(\sigma)\dot{x}^\nu(\sigma) \right]^\frac{1}{2}. \quad (11.12) \]

As observed in Subsection (2.2), the action (11.11) is invariant under arbitrary reparametrizations

\[ \sigma \rightarrow \sigma' = f(\sigma). \quad (11.13) \]

Variation of the action yields

\[
\delta m \mathcal{A} = \int_{\sigma_a}^{\sigma_b} d\sigma \, \delta L (\dot{x}^\mu(\sigma))
\]

\[
= -m^2 c^2 \frac{1}{2} \int_{\sigma_a}^{\sigma_b} d\sigma \, \frac{ds}{L (\dot{x}^\mu(\sigma))} \left[ (\partial_\lambda g_{\mu\nu}) \delta x^\lambda (\dot{x}^\mu(\sigma)) \dot{x}^\nu(\sigma) + 2g_{\lambda\nu} \frac{d\delta x^\lambda}{d\sigma} \dot{x}^\nu(\sigma) \right]. (11.14)
\]
On the right-hand side we have used the property (2.7) that the variation of the
derivative is equal to the derivative of the variation.

The factor before the bracket is equal to
\[ \frac{d\sigma}{(-mc ds/d\sigma)} = -\frac{d\sigma}{ds} \frac{ds}{mc}. \]
Thus, if we choose \( \sigma \) to be the proper time \( \tau \) for which
\( \frac{d\sigma}{ds} = \frac{d\tau}{ds} = \frac{1}{c} \), we may rewrite the variations as

\[ \delta m^A = -m \frac{1}{2} \int_{\tau_a}^{\tau_b} d\tau \left[ (\partial_\lambda g_{\mu\nu}) \delta x^\lambda \dot{x}^\mu(\tau)\dot{x}^\nu(\tau) + 2g_{\lambda\nu} \frac{d\delta x^\lambda}{d\tau} \dot{x}^\nu(\tau) \right]. \] (11.15)

This shows that if we use the proper time \( \tau \) to parameterize the paths, the equations
of motion can alternatively be derived from the simpler acti
\[ m\bar{A} = \int_{\tau_a}^{\tau_b} d\tau \bar{L}(\dot{x}^\mu(\tau)), \] (11.16)
where
\[ \bar{L}(\dot{x}^\mu(\tau)) \equiv -\frac{m}{2} g_{\mu\nu}(x(\tau))\dot{x}^\mu(\tau)\dot{x}^\nu(\tau). \] (11.17)
This has the same form as the action of a nonrelativistic point particle in four-
dimensional spacetime parameterized by a pseudotime \( \tau \). Note that although the
action (11.16) has the same extrema as (11.11), it has only half the size.

The second integral in (11.15) can be performed by parts to yi
\[ 2g_{\lambda\nu}(x(\tau))\delta x^\lambda(\tau)\dot{x}^\nu(\tau) \bigg|_{\tau_a}^{\tau_b} - 2 \int_{\tau_a}^{\tau_b} d\tau \delta x^\lambda(\tau) \frac{d}{d\tau} [g_{\lambda\nu}(x(\tau))\dot{x}^\nu(\tau)]. \] (11.18)

According to the extremal principle of classical mechanics, we derive the equations
of motion by varying the action with vanishing variations of the paths \( \delta q^\mu \) at the
end points \( \tau \) [recall (2.3), which leads to the equat
\[ \frac{1}{2} \int_{\tau_a}^{\tau_b} d\tau \left[ (\partial_\lambda g_{\mu\nu} - 2\partial_\mu g_{\lambda\nu}) \dot{x}^\mu(\tau)\dot{x}^\nu(\tau) - 2g_{\lambda\nu}\ddot{x}^\nu(\tau) \right] \delta x^\lambda(\tau) = 0. \] (11.19)
This is valid for all \( \delta x^\mu(\tau) \) vanishing at the end points, in particular for the infinites-
imal local spikes:
\[ \delta x^\mu(\tau) = \epsilon \delta(\tau - \tau_0). \] (11.20)
Inserting these into (11.19) we obtain the equations of motion
\[ g_{\lambda\nu}\ddot{x}^\nu(\tau) + \left( \partial_\mu g_{\lambda\nu} - \frac{1}{2}\partial_\lambda g_{\mu\nu} \right) \dot{x}^\mu(\tau)\dot{x}^\nu(\tau) = 0. \] (11.21)
It is convenient to introduce a quantity called the Riemann connection of Christoffel symbol:
\[ \bar{\Gamma}_{\mu\nu\lambda} \equiv \{\mu\nu, \lambda\} = \frac{1}{2} \left( \partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu} \right). \] (11.22)
Then the equation of motion can be written as
\[ g_{\lambda\nu}\ddot{x}^\nu(\tau) + \bar{\Gamma}_{\mu\nu\lambda} \dot{x}^\mu(\tau)\dot{x}^\nu(\tau) = 0. \] (11.23)
By further introducing the modified Christoffel symbol
\[ \tilde{\Gamma}^{\kappa}_{\mu\nu} \equiv \begin{cases} \kappa \\ \mu
\nu \end{cases} = g^{\kappa\lambda}\tilde{\Gamma}^{\lambda}_{\mu\nu} = g^{\kappa\lambda}\{\mu\nu\lambda\}, \] (11.24)
we can bring Eq. (11.21) to the form
\[ \ddot{x}^\lambda(\tau) + \tilde{\Gamma}^{\kappa}_{\mu\nu}\dot{x}^\mu(\tau)\dot{x}^\nu(\tau) = 0. \] (11.25)

A path \( x^\lambda(\tau) \) satisfying this differential equation of shortest length is called a geodesic trajectory. It is Einstein’s postulate, that this equation describes correctly the motion of a point particle in the presence of a gravitational field.

Now we turn to the simpler direct derivation of the equation of motion applying the coordinate transformation \( x^a(x^\mu) \) to the straight-line equation of motion (11.5) in Minkowski space:
\[ \ddot{x}^a(\tau) = \frac{d}{dt}\left[ \frac{\partial x^a}{\partial x^\mu} \dot{x}^\mu(\tau) \right] = \frac{\partial x^a}{\partial x^\mu} \ddot{x}^\mu(\tau) + \left( \frac{d}{dt} \frac{\partial x^a}{\partial x^\mu} \right) \dot{x}^\mu(\tau) = 0, \] (11.26)
where we have written \( \frac{\partial x^a}{\partial x^\mu} \) for the coordinate transformation matrix evaluated on the trajectory \( x(\tau) \). Multiplying this by \( \partial x^\lambda/\partial x^a \) and summing over repeated indices \( a \) yields
\[ \ddot{x}^\lambda(\tau) + \frac{\partial x^\lambda}{\partial x^a} \left( \frac{d}{dt} \frac{\partial x^a}{\partial x^\mu} \right) \dot{x}^\mu(\tau) = \ddot{x}^\lambda(\tau) + \frac{\partial x^\lambda}{\partial x^a} (\partial_{\mu}\partial_{\nu}x^a) \dot{x}^\mu(\tau)\dot{x}^\nu(\tau) = 0. \] (11.27)

The second term can be processed using (11.10) as follows:
\[ \frac{\partial x^\lambda}{\partial x^a} (\partial_{\mu}\partial_{\nu}x^a) = g^{\lambda\sigma}(\partial_{\mu}x^\sigma)(\partial_{\nu}x^a). \] (11.28)
It takes a little algebra to verify that this is equal to \( \tilde{\Gamma}^{\lambda}_{\mu\nu} \), so that the transformed equation of motion (11.27) coincides, indeed, with the geodesic equation (11.25).

### 11.3 Minkowski Geometry formulated in General Coordinates

In Einstein’s theory, all gravitational effects can be completely described by a non-trivial geometry of spacetime. As a first step towards developing this theory it is important to learn to distinguish between inessential properties of the geometry which are merely due to the formulation in terms of general coordinates, as in the last section, and those which are caused by the presence of gravitational forces. For this purpose we study in more detail the mathematics of coordinate transformation in Minkowski spacetime.
11.3 Minkowski Geometry formulated in General Coordinates

11.3.1 Local Basis tetrads

As in Eq. (1.25) we use coordinates \( x^a \) \((a = 0, 1, 2, 3)\) to specify the points in Minkowski spacetime. From now on it will be convenient to use fat latin letters to denote four-vectors in spacetime. Thus we shall denote the four-dimensional basis vectors by \( e_a \), and an arbitrary four-dimensional vector with coordinates \( x^a \) by \( x = e_a x^a \). The basis vectors are orthonormal with respect to the Minkowski metric \( g_{ab} \) of Eq. (1.29):

\[
e_a e_b = g_{ab}.
\]

The basis vectors \( e_a \) define an inertial frame of reference.

Let us now reparametrize this Minkowski spacetime by a new set of coordinates \( x^\mu \) whose values are given by a mapping \( x^a \to x^\mu (x^a) \). Since \( x^\mu \) still labels the same spacetime we shall assume the function \( x^\mu (x^a) \) to possess an inverse \( x^a = x^a(x^\mu) \) and to be sufficiently smooth so that \( x^\mu (x^a) \) and \( x^a (x^\mu) \) have at least two smooth derivatives. These will always commute with each other. In other words, the general coordinate transformation (11.30) and their inverse \( x^a (x^\mu) \) will satisfy the integrability conditions of Schwartz:

\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) x^a (x^\kappa) = 0, \quad (11.31)
\]

\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\lambda x^a (x^\kappa) = 0. \quad (11.32)
\]

The conditions \( x^\mu (x^a) = \text{const.} \) define a network of new coordinate hypersurfaces whose normal vectors are given by (see Fig. 11.1)

\[
e_\mu (x) \equiv e_a e^a_\mu (x) = e_a \frac{\partial x^a}{\partial x^\mu}.
\]

These are called local basis vectors. Their components \( e^a_\mu (x) \) are called local basis tetrads. The difference vector between two points \( x' \) and \( x \) has, in the inertial frame

\[\]
of reference, the description \( \Delta \mathbf{x} = e_a \left( x'^a - x^a \right) \). When going to coordinates \( x'^\mu, x^\mu \), this becomes

\[
\Delta \mathbf{x} = e_a \int_x^{x'} e^a_\mu(x) dx^\mu. \tag{11.34}
\]

The length of an infinitesimal vector \( d\mathbf{x} \) is given by

\[
ds = \sqrt{d\mathbf{x}^2} = \sqrt{(e_\mu dx^\mu)^2} = \sqrt{e_\mu e_\nu dx^\mu dx^\nu}. \tag{11.35}
\]

The right-hand side shows that the metric in the curvilinear coordinates can be expressed as a scalar product of the local basis vectors:

\[
g_{\mu\nu}(x) = e_\mu(x)e_\nu(x) = g_{ab} e^a_\mu(x) e^b_\nu(x). \tag{11.36}
\]

Indeed, inserting here (11.33) we verify:

\[
g_{\mu\nu}(x) = e_\mu(x)e_\nu(x) = e^a_\mu(x) e^b_\nu(x) \delta^b_a = g_{ab} \frac{\partial x^a}{\partial x^\mu} \frac{\partial x^b}{\partial x^\nu}. \tag{11.37}
\]

as in (11.8).

In the sequel, we shall freely raise and lower the latin index using the Minkowski metric \( g_{ab} = g^{ab} \), and define

\[
e^a_\mu \equiv g^{ab} e^b_\mu, \quad e_{a\mu} \equiv g_{ab} e^b_\mu. \tag{11.38}
\]

Then we can rewrite (11.36) as

\[
g_{\mu\nu}(x) = e_\mu(x)e_\nu(x) = e^a_\mu(x) e^a_\nu(x). \tag{11.39}
\]

Since the general coordinate formulation (11.29) was assumed to have an inverse, we can also calculate the derivatives \( \partial x^\mu/\partial x^a \). These are orthonormal to the derivatives \( \partial x^a/\partial x^\mu \) in two ways:

\[
\frac{\partial x^a}{\partial x^\mu} \frac{\partial x^\mu}{\partial x^a} = \delta^a_a, \quad \frac{\partial x^a}{\partial x^\mu} \frac{\partial x^\mu}{\partial x^\nu} = \delta^a_\nu. \tag{11.40}
\]

It is useful to denote the inverse derivatives \( \partial x^\mu/\partial x^a \) by \( e_a^\mu \) and introduce the vectors \( \mathbf{e}^\mu = e_a g^{ab} e^b_\mu \), called the reciprocal multivalued basis tetrads (see also Fig. 11.2). With this notation, the equations in (11.40) become orthonormality and completeness relations of the tetrads:

\[
e^a_\mu(x) e^b_\mu(x) = \delta^a_b, \quad e^a_\mu(x) e^a_\nu(x) = \delta^\mu_\nu. \tag{11.41}
\]

The scalar product

\[
g^{\mu\nu}(x) = \mathbf{e}^\mu \mathbf{e}^\nu = e^a_\mu(x) e^a_\nu(x) \tag{11.42}
\]

is obviously the inverse metric, satisfying

\[
g^{\mu\nu}(x) g_{\nu\lambda}(x) = \delta^\nu_\lambda. \tag{11.43}
\]

The metric \( g_{\mu\nu}(x) \) and its inverse \( g^{\mu\nu}(x) \) can be used to freely lower and raise greek indices on any tensor, and to form invariants under coordinate transformations by contraction of all indices.
11.3 Minkowski Geometry formulated in General Coordinates

11.3.2 Vector- and Tensor Fields, and Lorentz Invariance

When formulating the laws of physics in Minkowski spacetime it is important to analyze physical quantities according to their transformation properties under Lorentz transformations, which change the coordinates of one inertial frame of reference to those of another

\[ x^a \rightarrow x'^a \equiv (\Lambda x)^a = \Lambda^a_b x^b. \]  \hspace{1cm} (11.44)

The 4 \times 4 matrices \( \Lambda^a_b \) preserve the metric \( g_{ab} \). The length elements \( ds = \sqrt{g_{ab}dx^adx^b} = \sqrt{g_{ab}dx'^adx'^b} \) are the same in both coordinate frames \( x^a \) and \( x'^a \). This implies that the transformation matrices \( \Lambda^a_b \) satisfy

\[ g_{ab} \Lambda^a_c \Lambda^c_b = (\Lambda^T g \Lambda)_{a'b'} = g_{a'b'}. \]  \hspace{1cm} (11.45)

Infinitesimally, we parametrize \( \Lambda^a_b \) and \( (\Lambda^{-1})^a_b \) as

\[ \Lambda^a_b = \delta^a_b + \omega^a_b, \]
\[ (\Lambda^{-1})^a_b = \delta^a_b - \omega^a_b, \]  \hspace{1cm} (11.46)

and the relation \( (g\Lambda)^T = g\Lambda^{-1} \) implies that

\[ \omega_{ab} \equiv g_{aad}\omega^d_{ab} \]  \hspace{1cm} (11.47)

is an antisymmetric matrix, i.e., \( \omega_{ab} = -\omega_{ba} \). It has six independent matrix elements. The three components \( \omega_k = \frac{1}{2} \epsilon_{kij}\omega^i_j \) parametrize infinitesimal rotations, where \( |\omega| \) is the angle, and \( \hat{\omega} \equiv \omega/|\omega| \) the axis of rotation. The three components \( \omega_{i0} = -\omega_{0i} = \zeta \) specify the small relative rapidity \( \zeta \) of the two coordinate frames.

Since the physical points are the same before and after a Lorentz transformation, the basis vectors \( e_a \) change according to the law

\[ e_a \rightarrow e'_a \equiv (e\Lambda^{-1})_a = e_b (\Lambda^{-1})^b_a = (eg\Lambda^T g)_a. \]  \hspace{1cm} (11.48)

This gives

\[ x \equiv e_a x^a = egx \rightarrow x' = e\Lambda^T g\Lambda x = egx = x \]  \hspace{1cm} (11.49)

showing that the vectors in the external frame are the same before and after the transformation.

Consider now a vector field \( v^a(x) \). It assigns to every point \( P \) a vector

\[ \mathbf{v}(P) = e_a v^a(x). \]  \hspace{1cm} (11.50)

Under a Lorentz transformation of the coordinates \( x^a \), the basis vectors \( e^a \) change. The observable vector \( \mathbf{v}(P) \), however, must remain the same at the same point in space, i.e.,

\[ \mathbf{v}'(P) = \mathbf{v}(P). \]  \hspace{1cm} (11.51)
Writing this as
\[ \mathbf{v}'(P) = e'_a v'^a(x') = \mathbf{v}(P) = e_a v^a(x) \] (11.52)
we see that the components of the vector in the two frames have to be related in the same way as the coordinate \(x'^a\) and \(x^a\), i.e.,
\[ v'^a(x') = \Lambda^a_{\ b} v'^b(x), \] (11.53)
or, written differently,
\[ v'^a(x) = \Lambda^a_{\ b} v'^b \left( \Lambda^{-1} x \right). \] (11.54)
For infinitesimal transformations,
\[ \Lambda^a_{\ b} x^b = (\delta^a_{\ b} + \omega^a_{\ b}) x^b \] (11.55)
with
\[ \left( \Lambda^{-1} x \right)^a = x^a - \omega^a_{\ b} x^b; \] (11.56)
the vector \(v^a(x)\) goes over into
\[ v'^a(x) = v^a(x) + \omega^a_{\ b} v^b(x) - \omega'^{\ b}_{\ b} x^b \partial_b v^a(x). \] (11.57)
The infinitesimal transformation law (11.57) is recognized to be a substantial variation defined in (3.6), so that we may write Eq. (11.57) also as
\[ \delta_s v(x) = v'(x) - v(x) = \omega^b_{\ b} v^b(x) - \omega'^{\ b}_{\ b} x^b \partial_b v^a(x). \] (11.58)
For reasons for constructing Lorentz-invariant equations of motion we introduce for every vector field \(v^a(x)\) a contravariant vector field \(\tilde{v}^a(x) = g_{ab} v^b(x)\). Its transformation properties correspond to those of the derivative with respect to the coordinates in Eq. (1.80). Infinitesimally, a substantial variation of \(\tilde{v}^a\) is equal to
\[ \delta_s \tilde{v}^a(x) = \omega^b_{\ b} \tilde{v}^b(x) - \omega'^{\ b}_{\ b} x^b \partial_b \tilde{v}^a(x). \] (11.59)
where we have introduced the matrix elements
\[ \omega^b_{\ b} = g_{ab} g^{a'b'} \omega^a_{\ b'} = g^{a'b'} \omega_{b'a}. \] (11.60)
The derivatives of a contravariant vector field \(v^a\) with respect to changes of \(x^a\) are higher tensor fields. Infinitesimally, derivatives transform via the sum of operations (11.59), one applied to each index. This follows directly from (11.59) and the commutation rule \([\partial_a, x_b] = g_{ab}\):
\[ \delta_s \partial_b v^a = \partial_b \delta_s v^a = \partial_b \left( \omega^a_{\ a'} v^{a'} + \omega^a_{\ c'} x^c \partial_{c'} v^a \right) \] (11.61)
\[ = \omega^a_{\ a'} \partial_b v^{a'} + \omega^b_{\ b'} \partial_{b'} v^a + \omega^c_{\ c'} x^c \partial_{c'} \partial_b v^a. \]
This simple rule can easily be extended to arbitrary higher derivatives thereby forming higher tensor fields. Note that since the arguments in \( f \) and \( f' \) in (3.6) are the same, the operation “substantial variation” commutes with the derivative.

Consider now the same physical objects but described in terms of curvilinear coordinates \( x^\mu(x^a) \). Then the components of \( \mathbf{v} \) are measured not with respect to the basis \( e_a \) but with respect to the local basis \( e_\mu(x) = e_a e^a_\mu(x) \). It is then natural to specify \( \mathbf{v}(x) \) in terms of its local components \( v^\mu(x) = v^a(x) e_a^\mu(x) \). On the fields \( v^\mu(x) \) one cannot only perform Lorentz transformations but any general coordinate transformation \( x^\mu \rightarrow x'^\mu(x^\nu) \) which, in the following, will shortly be referred to as \textit{Einstein transformations}.

Under Einstein transformations, the vectors \( e_a^\mu(x) \) being derivatives of the coordinate transformation functions \( x^\mu(x^a) \), undergo the following changes

\[
\begin{align*}
e_a^\mu \rightarrow e'^a_\mu(x') &\equiv \frac{\partial x'^\mu}{\partial x^a} = \frac{\partial x'^\nu}{\partial x^a} = \alpha^\mu_\nu(x) e_a^\nu(x) \\
e_\mu^a \rightarrow e'^\mu_\nu(x') &\equiv \frac{\partial x^\mu}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\mu} = \alpha_\mu^\nu(x) e^a_\nu(x).
\end{align*}
\]

The matrices

\[
\begin{align*}
\alpha_\mu^\nu(x) &\equiv \frac{\partial x'^\mu}{\partial x^\nu} \\
\alpha^\mu_\nu(x) &\equiv \frac{\partial x^\mu}{\partial x'^\nu}
\end{align*}
\]

(11.63)

are orthogonal to each other

\[
\begin{align*}
\alpha^\nu_\lambda \alpha_\nu^\mu &= \delta^\mu_\lambda, \\
\alpha_\mu^\nu \alpha^\lambda_\mu &= \delta^\lambda_\nu,
\end{align*}
\]

(11.64)
i.e.,

\[
(\alpha^{-1})^\nu_\lambda = \alpha_\lambda^\nu
\]

(11.65)
is a right- as well as a left-inverse of the matrix \( \alpha_\nu^\mu \). Infinitesimally, we may set

\[
x'^\mu \equiv x^\mu - \xi^\mu(x),
\]

(11.66)
and see that Einstein transformations can be interpreted as \textit{local} translations. The infinitesimal transformation matrices are

\[
\begin{align*}
\alpha^\lambda_\nu &\approx \delta^\lambda_\nu - \partial_\nu \xi^\lambda(x) \\
\alpha_\mu^\nu &\approx \delta_\mu^\nu + \partial_\mu \xi^\nu(x).
\end{align*}
\]

(11.67)
In the sequel, the substantial variations $\delta_4$ under Einstein transformations will be denoted by $\delta_E$. For the basis tetrads $e_a^\mu(x)$ and $e_a^\nu(x)$ the are

\[
\delta_E e_a^\mu = e_a^\mu(x) - e_a^\nu(x) = e_a^\mu(x') - e_a^\nu(x') \\
= e_a^\mu(x) - e_a^\mu(x') + e_a^\nu(x') - e_a^\nu(x) \\
= \xi^\lambda \partial_\lambda e_a^\mu(x) - \partial_\lambda \xi^\mu e_a^\mu(x),
\]

(11.68)

\[
\delta_E e_a^\nu = \xi^\lambda \partial_\lambda e_a^\nu(x) + \partial_\lambda \xi^\nu e_a^\nu(x).
\]

(11.69)

Analogous transformation laws can be derived for the components of the vector fields $v^\mu(x)$ and $v_\mu(x)$. They follow from the fact that the components $v^\mu(x), v_\mu(x)$ do not change under a change of the general coordinates from $x^\mu$ to $x'^\mu$. Thus we have the obvious relation

\[
v'^a(x^b) = v^a(x^b).
\]

(11.70)

When reparametrizing the point $x^b$ in the two different coordinates $x'^\mu$ and $x^\mu$, this relation takes the form

\[
v'^a(x^\prime) = v^a(x)
\]

(11.71)

where we have omitted the greek superscripts of $x'$ and $x$. Thus the substantial variations, i.e., the changes at the same values of the general coordinates $x^\mu$, are

\[
\delta_E v^a(x) = v'^a(x) - v^a(x) = \xi^\lambda \partial_\lambda v^a(x).
\]

(11.72)

Using this and (11.69), we derive from (11.71)

\[
v'^\mu(x^\prime) = \alpha^\mu v^\nu(x),
\]

\[
v_\mu(x^\prime) = \alpha^\nu v_\nu(x),
\]

(11.73)

with the substantial variations

\[
\delta_E v^\mu(x) = v'^\mu(x) - v^\mu(x) = \xi^\lambda \partial_\lambda v^\mu(x) - \partial_\lambda \xi^\nu v^\lambda
\]

(11.74)

\[
\delta_E v_\mu(x) = v'_\mu(x) - v_\mu(x) = \xi^\lambda \partial_\lambda v_\mu(x) - \partial_\lambda \xi^\nu v_\lambda.
\]

(11.75)

Any four-component field with these transformation properties is called **Einstein vector** or **world vector**.

This definition can trivially be extended to higher Einstein- or world tensors.

We merely apply separately the transformation matrices (11.67) to each index.

In particular, the metric $g_{\mu\nu}(x)$ transform as

\[
g'^{\lambda\kappa}(x^\prime) = \alpha^\lambda \alpha^\kappa g_{\mu\nu}(x),
\]

\[
g^{\lambda\kappa}(x^\prime) = \alpha^\lambda \alpha^\kappa g^{\mu\nu}(x),
\]

(11.76)

or, infinitesimally, as

\[
\delta_E g_{\mu\nu} = \xi^\lambda \partial_\lambda g_{\mu\nu} + \partial_\mu \xi^\lambda g_{\lambda\nu} + \partial_\nu \xi^\lambda g_{\mu\lambda},
\]

(11.77)

\[
\delta_E g^{\mu\nu} = \xi^\lambda \partial_\lambda g^{\mu\nu} - \partial_\mu \xi^\lambda g^{\lambda\nu} - \partial_\nu \xi^\lambda g^{\mu\lambda}.
\]

(11.78)
This can be rewritten in a manifestly covariant form as follows:

\[
\delta E g_{\mu\nu} = \bar{D}_\mu \xi_\nu + \bar{D}_\nu \xi_\mu, \quad (11.79)
\]
\[
\delta E g^{\mu\nu} = \bar{D}^\mu \xi^\nu + \bar{D}^\nu \xi^\mu. \quad (11.80)
\]

It is now obvious from (11.64) that one can multiply any set of world tensors with each other by a simple contraction of upper and lower indices and always obtain new world tensors. In particular, one obtains an *Einstein*- or *world invariant* if such a contraction is complete, i.e., if no index is left.

### 11.3.3 Affine Connections and Covariant Derivatives

The multiplication rules for world tensors are completely analogous to those for Lorentz tensors. There is, however, one important difference. Contrary to the Lorentz case, derivatives of world tensors are no longer tensors. In curvilinear coordinates, certain modifications of the derivatives are required in order to make them proper tensors. It is quite easy to find these modifications and construct objects analogous to the derivative tensors in the Lorentz frames. For this we rewrite the derivative tensors in terms of the general curvilinear components. Take, for example, the tensor \( \partial_b v_a(x) \). Going over to curvilinear components \( x^\mu \) we can write this as

\[
\partial_b v_a = \partial_b \left( e^a_\mu v_\mu \right). \quad (11.81)
\]

But if we take the derivative \( \partial_b \) past the basis tetrad \( e^a_\mu \) we find

\[
\partial_b v_a = e^a_\mu \partial_b v_\mu + \partial_b e^a_\mu v_\mu. \quad (11.82)
\]

Using the relation

\[
\partial_b = e^b_\lambda \partial_\lambda \quad (11.83)
\]

we see that

\[
\partial_b v_a = e^a_\mu e^b_\nu \partial_\nu v_\mu + \left( e^b_\nu \partial_\nu e^a_\lambda \right) v_\lambda. \quad (11.84)
\]

The right-hand side can be rewritten as

\[
\partial_b v_a = e^a_\mu e^b_\nu D_\nu v_\mu \quad (11.85)
\]

where the symbol \( D_\nu \) stands for the modified derivative

\[
D_\nu v_\mu = \partial_\nu v_\mu - e^c_\mu \partial_\nu e^c_\lambda v_\lambda \equiv \partial_\nu v_\mu - \Gamma_\nu^\lambda_\mu v_\lambda. \quad (11.86)
\]

The explicit form on the right-hand side follows from the simple relation

\[
\partial_\nu e^a_\lambda = -e^a_\mu \left( e^c_\lambda \partial_\nu e^c_\mu \right) \quad (11.87)
\]
\[
\partial_\nu e^a_\lambda = -e^a_\mu \left( e^c_\lambda \partial_\nu e^c_\mu \right) \quad (11.88)
\]
which, in turn, is a consequence of differentiating the orthogonality relation $e_a^\lambda e_b^\lambda = \delta_a^b$. Similarly, we can find the Einstein version of the derivative of a contravariant vector field $\partial^b v^a(x)$, which can be rewritten as

$$\partial^b v^a = \partial^b \left( e^a_\mu v^\mu \right) = e^a_\mu e^\nu_b \partial_\nu v^\mu + (e^\nu_b \partial_\nu e^a_\lambda) v^\lambda$$

(11.89)

and brought to the form

$$e^a_\mu e^\nu_b D_\nu v^\mu,$$

(11.90)

with a covariant derivative

$$D_\nu v^\mu = \partial_\nu v^\mu - e^c_\lambda \partial_\nu e^\mu_c v^\lambda = \partial_\nu v^\mu + e^c_\mu \partial_\nu e^\lambda_c v^\lambda \equiv \partial_\nu v^\mu + \Gamma^\lambda_{\nu\mu} v^\lambda.$$

(11.91)

The extra term appearing in (11.86) and (11.91):

$$\Gamma^\lambda_{\mu\nu} \equiv e^a_\lambda \partial_\mu e^\nu_a \equiv e^a_\nu \partial_\mu e^\lambda_a$$

(11.92)

is called the affine connection of the space. In general, a spacetime with a metric $g_{\mu\nu}$ and an affine connection to define covariant derivatives, is called an affine space, and the geometry carried by $g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}$, is referred to a metric-affine geometry. Note that by definition, the covariant derivatives of $e^a_\nu$ and $e^\nu_a$ vanish:

$$D_\mu e^a_\nu = \partial_\mu e^a_\nu - \Gamma^a_{\lambda\mu} e^\lambda_\nu = 0,$$

(11.93)

$$D_\mu e^\nu_a = \partial_\mu e^\nu_a + \Gamma^\nu_{\mu\lambda} e^a_\lambda = 0.$$ 

(11.94)

Since $g_{\mu\nu} = e^a_\mu e^\nu_a$, the same property holds for the metric tensor

$$D_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - \Gamma^\sigma_{\lambda\mu} g_{\sigma\nu} - \Gamma^\sigma_{\lambda\nu} g_{\mu\sigma} = 0,$$

(11.95)

$$D_\lambda g^{\mu\nu} = \partial_\lambda g^{\mu\nu} + \Gamma^\nu_{\lambda\sigma} g^{\mu\sigma} + \Gamma^\mu_{\lambda\sigma} g^{\nu\sigma} = 0.$$ 

(11.96)

It is worth noting that the metric satisfies once more relations like (11.96), in which the connections are replaced by Christoffel symbols. In fact, from the definition (11.22) we can verify directly that

$$\bar{D}_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - \Gamma^\sigma_{\lambda\mu} g_{\sigma\nu} - \Gamma^\sigma_{\lambda\nu} g_{\mu\sigma} = 0,$$

(11.97)

$$\bar{D}_\lambda g^{\mu\nu} = \partial_\lambda g^{\mu\nu} + \Gamma^\nu_{\lambda\sigma} g^{\mu\sigma} + \Gamma^\mu_{\lambda\sigma} g^{\nu\sigma} = 0.$$ 

(11.98)

Since the left-hand sides of (11.85) and (11.89) are tensors with respect to Lorentz transformations, the covariant derivatives $D_\nu v^\mu_\nu$ and $D_\nu v^\mu$ in Eqs. (11.86) and (11.91) must be tensors with respect to general coordinate transformation, i.e., world tensors. In fact, one can easily verify that they transform covariantly:

$$\partial^\prime_{\mu^\prime} v^\rho^\prime (x^\prime) = \alpha^\mu_{\mu^\prime} \alpha^\rho_{\rho^\prime} \partial_{\mu} v^\rho (x).$$

(11.99)

---

1In the present context where the spacetime is still flat and only reparametrized with curvilinear coordinates, this is a rather trivial statement. For the situation in general geometries see the remark [2] in Notes and References.
11.3 Minkowski Geometry formulated in General Coordinates

Working out the derivative on the left-hand side we obtain
\[
\partial'_\mu v'_\nu(x') = \alpha'^\mu_\mu \partial_\mu [a'_{\nu'\nu}(x)]
= \alpha'^\mu_\mu \alpha_{\nu'\nu} \partial_\mu v_\nu(x) + \alpha'^\mu_\mu \partial_\mu a_{\nu'\nu}.
\] (11.100)

The last term is an obstacle to covariance. It is compensated by a similar term in non-tensorial behavior of \(\Gamma^\lambda_{\mu\nu}^{}\):
\[
\Gamma'_{\mu'\nu'\nu}^{}(x') = e'_a^\nu \partial_{\mu'} e_\nu^a = \alpha^X_\lambda \alpha^\nu_{\mu'} \alpha^a_\lambda \partial_\mu (x_{\nu'\nu} e^a_\nu)
= \alpha^\nu_{\mu'} \Gamma^\lambda_{\mu\nu}^{}(x) + \alpha^\nu_{\nu'} \partial_\mu a_{\nu'\nu}.
\] (11.101)

Infinitesimally, the transformation matrices are \(\alpha_{\nu'}^\nu = \delta_{\nu'}^\nu + \partial_\nu \xi_{\nu'}^\nu\) and \(\alpha^\nu_{\nu'} = \delta^\nu_{\nu'} - \partial_\nu \xi_{\nu'}\), and we easily verify that the covariant derivatives \(\hat{D}_{\mu'} v_{\nu'}\), \(\hat{D}_{\mu'} v_{\nu'}\) have the correct substantial transformation properties of world tensors:
\[
\delta_E D_{\mu} v_{\nu} = \xi^\lambda \partial_\lambda D_{\mu} v_{\nu} + \partial_\mu \xi^\nu D_{\mu} v_{\lambda},
\delta_E D_{\mu'} v_{\nu'} = \xi^\lambda \partial_\lambda D_{\mu} v_{\nu'} + \partial_\mu \xi^\nu D_{\mu} v_{\lambda'},
\] (11.103)

The last non-covariant piece in
\[
\delta_E \partial_\mu v_{\nu} = \partial_\mu \delta_E v_{\nu} = \partial_\mu \left(\xi^\lambda \partial_\lambda v_{\nu} + \partial_\nu \xi^\nu v_{\lambda}\right)
= \xi^\lambda \partial_\lambda \partial_\mu v_{\nu} + \partial_\mu \xi^\nu \partial_\lambda v_{\nu} + \partial_\nu \xi^\nu \partial_\mu v_{\lambda} + \partial_\mu \partial_\nu \xi^\nu v_{\lambda}
\] (11.104)
is canceled by the last non-tensorial piece in \(\delta_E \Gamma^\kappa_{\mu\nu}^{}\):
\[
\delta_E \Gamma^\kappa_{\mu\nu}^{} = \xi^\lambda \partial_\lambda \Gamma_{\mu\nu}^\kappa + \partial_\mu \xi^\nu \Gamma_{\mu\nu}^\kappa + \partial_\nu \xi^\lambda \Gamma_{\mu\nu}^\kappa + \partial_\mu \partial_\nu \xi^\nu.
\] (11.105)

It is easily checked that the same cancellation occurs in the covariant derivative of an arbitrary tensor field, defined as
\[
D_{\mu} t_{\nu_1...\nu_n}^{} v'_{\nu'_1...\nu'_{n'}} \equiv \partial_\mu t_{\nu_1...\nu_n}^{} v'_{\nu'_1...\nu'_{n'}} - \sum_i \Gamma_{\mu\nu_i}^{} v_i^{} t_{\nu_1...\nu_i...\nu_n}^{} v'_{\nu'_1...\nu'_{n'}}
+ \sum_i \Gamma_{\mu\nu_i}^{} v'_{\nu_1...\nu_i...\nu_n}^{} t_{\nu_1...\nu_i...\nu_n}^{} v'_{\nu'_1...\nu'_{n'}}.
\] (11.106)

Note that in the antisymmetric part of the affine connection, the torsion tensor \(S_{\mu\nu}^\lambda\), the last term in (11.105) disappears:
\[
\delta_E S_{\mu\nu}^\kappa = \xi^\lambda \partial_\lambda S_{\mu\nu}^\kappa + \partial_\mu \xi^\nu S_{\mu\nu}^\kappa + \partial_\nu \xi^\lambda S_{\mu\nu}^\kappa,
\] (11.107)
which is the transformation law for any tensor of the same index configuration.
11.4 Covariant Time Derivative and Acceleration

In this context it is useful to introduce the concept of a covariant time derivative of a vector field in spacetime. The four-velocity $u^\mu(\tau) = q^\mu(\tau)$ transforms like a four-vector. By analogy with the covariant derivative of a vector field $v^\mu(x)$ along a trajectory $x = q(\tau)$ in Eqs. (11.86) and (11.91), we define the covariant time derivative of the vector field along a trajectory $x = q(\tau)$ as

$$\frac{D}{d\tau}v^\mu(\tau) \equiv \frac{d}{d\tau}v^\mu(\tau) + \Gamma^\lambda_\mu_\kappa v^\lambda(\tau)u^\kappa(\tau), \quad \frac{D}{d\tau}v_\mu(\tau) \equiv \frac{d}{d\tau}v_\mu(\tau) - \Gamma_\mu_\lambda_\kappa u^\lambda(\tau)v^\kappa(\tau), \quad (11.108)$$

where we have abbreviated $v^\mu(q(\tau))$ by $v^\mu(\tau)$.

If $v^a(x)$ were a parallel vector field in a Minkowski spacetime, the equation (11.108) would describe the change of the transformed components $v^\mu(x) = e^\mu_a(x)v^a(x)$. By analogy, we shall define (11.108) as the change under a parallel transport also for general metric-affine geometries. This will be discussed in detail in Subsection 12.3.2.

If the vector trajectory is the velocity trajectory of a point particle, the covariant derivative is the covariant acceleration.

We may also define a Riemann covariant derivative

$$\bar{D}\frac{d\tau}{d\tau}v^\mu(\tau) \equiv \frac{d}{d\tau}v^\mu(\tau) + \bar{\Gamma}^\lambda_\mu_\kappa v^\lambda(\tau)u^\kappa(\tau), \quad \bar{D}\frac{d\tau}{d\tau}v_\mu(\tau) \equiv \frac{d}{d\tau}v_\mu(\tau) - \bar{\Gamma}_\mu_\lambda_\kappa u^\lambda(\tau)v^\kappa(\tau), \quad (11.109)$$

which vanish along geodesic trajectories.

In Eq. (1.305) we have derived the time derivative of the spin four-vector of a spinning point particle in Minkowski space. The multivalued or nonholonomic mapping principle transforms this to a general affine geometry:

$$\frac{DS^\mu_\kappa}{d\tau} = S_\kappa\frac{Du^\kappa_{\mu}}{d\tau} \quad (11.110)$$

This equation shows that in the absence of external forces, the spin four-vector of a point particle remains always parallel to its initial orientation along the entire autoparallel trajectory:

$$\frac{DS^\mu_\kappa}{d\tau} = 0. \quad (11.111)$$

11.5 Torsion tensor

Since the coordinate transformations $x^\mu(x^a)$ and $x^a(x^\mu)$ were assumed to be integrable, the derivatives of the infinitesimal local translation field $\xi^a(x)$ commute with each other:

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) \xi^\lambda(x) = 0. \quad (11.112)$$

This has the consequence that the antisymmetric part of the connection

$$S^\lambda_\mu_\nu \equiv \frac{1}{2} \left(\Gamma^\lambda_\mu_\nu - \Gamma^\lambda_\nu_\mu\right) = e^\lambda_\sigma \partial_\mu e^\sigma_\nu - e^\lambda_\sigma \partial_\nu e^\sigma_\mu \quad (11.113)$$

H. Kleinert, GRAVITY WITH TORSION
transforms like a proper tensor. This follows directly from the transformation law (11.105). The additional derivative term $\partial_{\mu} \partial_{\nu} \xi^{\kappa}$ arising in the transformation of $\Gamma_{\mu\nu}^{\kappa}$ disappears in the antisymmetrized expression (11.113). For this reason, $S_{\mu\nu}$ is called the torsion tensor.

Minkowski spacetime has no torsion so that the tensor nature of $S_{\mu\nu}^{\lambda}$ implies that it vanishes in all curvilinear coordinates. In order to see that there is no torsion we describe Minkowski space in terms of coordinates $x^{a}$ which coincide with the inertial coordinates $x^{a'}$. Then the basis tetrads $e^{a}_{\mu} = \partial_{\mu}x^{a}$ are unit matrices, so that the connection vanishes, and so does its antisymmetric part, the torsion. If we now perform a general coordinate transformation to curvilinear coordinates $x^{\mu}(x^{a})$ the connection will in general become nonzero. The torsion, however, being a tensor, remains zero for all coordinate transformations of Minkowski space.

It is useful to realize that with the help of the torsion tensor, the connection can be decomposed into a Christoffel part, given by (11.24), which depends only on the metric $g_{\mu\nu}(x)$, and a second part, called the contortion tensor, which is a combination of torsion tensors.

To derive this decomposition, which is valid in spaces with torsion, let us define the modified connection

$$
\Gamma_{\mu\nu\lambda} \equiv \Gamma_{\mu\nu}^{\kappa}g_{\kappa\lambda} = e_{a\lambda}\partial_{\mu}e^{a}_{\nu},
$$

and decompose this trivially as follows:

$$
\Gamma_{\mu\nu\lambda} = \tilde{\bar{\Gamma}}_{\mu\nu\lambda} + \bar{K}_{\mu\nu\lambda},
$$

(11.114)

where

$$
\tilde{\bar{\Gamma}}_{\mu\nu\lambda} = \frac{1}{2}\left\{ e_{a\lambda}\partial_{\mu}e^{a}_{\nu} + \partial_{\mu}e_{a\lambda}e^{a}_{\nu} + e_{a\mu}\partial_{\nu}e^{a}_{\lambda} + e_{a\lambda}\partial_{\nu}e^{a}_{\mu} - e_{a\mu}\partial_{\lambda}e^{a}_{\nu} - \partial_{\lambda}e_{a\mu}e^{a}_{\nu} \right\},
$$

(11.115)

$$
\bar{K}_{\mu\nu\lambda} = \frac{1}{2}\left\{ e_{a\lambda}\partial_{\mu}e^{a}_{\nu} - e_{a\lambda}\partial_{\nu}e^{a}_{\mu} + e_{a\nu}\partial_{\lambda}e^{a}_{\mu} - e_{a\mu}\partial_{\lambda}e^{a}_{\nu} + e_{a\nu}\partial_{\lambda}e^{a}_{\mu} - e_{a\mu}\partial_{\lambda}e^{a}_{\nu} \right\}.
$$

(11.116)

The terms in the first expression can be combined to

$$
\tilde{\bar{\Gamma}}_{\mu\nu\lambda} = \frac{1}{2}\left\{ \partial_{\mu}(e_{a\lambda}e^{a}_{\nu}) + \partial_{\nu}(e_{a\mu}e^{a}_{\lambda}) - \partial_{\lambda}(e_{a\mu}e^{a}_{\nu}) \right\} = \frac{1}{2}\left( \partial_{\mu}g_{\lambda\nu} + \partial_{\nu}g_{\lambda\mu} - \partial_{\lambda}g_{\mu\nu} \right),
$$

(11.117)

which shows that $\tilde{\bar{\Gamma}}_{\mu\nu\lambda}$ is equal to the Riemann connection $\bar{\Gamma}_{\mu\nu\lambda}$ in Eq. (11.22). The second expression is a combination of three torsion tensors (11.113). Defining an associated torsion tensor $S_{\mu\nu\lambda} \equiv S_{\mu\nu}^{\kappa}g_{\kappa\lambda}$, we see that

$$
\bar{K}_{\mu\nu\lambda} \equiv K_{\mu\nu\lambda} \equiv S_{\nu\lambda\mu} - S_{\nu\lambda\mu} + S_{\lambda\mu\nu},
$$

(11.118)

The combination of the three torsion tensors $S_{\mu\nu\lambda}$ is the so-called contortion tensor $K_{\mu\nu\lambda}$. The order of the indices of the three torsion terms are easy to remember: The first starts out with the same indices as $K_{\mu\nu\lambda}$. The second
and third terms are shifted cyclically to the left with alternating signs. Note that
the antisymmetry of \( S_{\mu \nu \lambda} \) makes the contortion tensor \( K_{\mu \nu \lambda} \) antisymmetric in the
last two indices.

Summarizing, we have found that the full affine connection \( \Gamma_{\mu \nu \lambda} \) can be decom-
posed into a sum of a Riemann connection and a contortion tensor:

\[
\Gamma_{\mu \nu \lambda} = \tilde{\Gamma}_{\mu \nu \lambda} + K_{\mu \nu \lambda}.
\] (11.119)

### 11.6 Curvature Tensor as Covariant Curl of Affine Connection

In the last section we have seen that even though the connection \( \Gamma_{\mu \nu} \) is not a tensor,
its antisymmetric part, the torsion \( S_{\mu \nu} \), is a tensor. The question arises whether it
is possible to form a covariant object which contains information on the content
of gravitational forces in the symmetric Christoffel part of the connection. Such a
tensor does indeed exist.

When looking back at the transformation properties (11.113) we see that the
tensor character is destroyed by the last term which is additive in the derivative of
an arbitrary function \( \partial_{\mu} \partial_{\nu} \xi^\kappa(x) \). Such additive derivative terms were encountered
before in Subsection 2.4.4 in gauge transformations of electromagnetism. Recall that
the gauge field of magnetism transform with such an additive derivative term [recall
(2.103)]

\[
\delta A_a(x) = \partial_a \Lambda(x),
\] (11.120)

where \( \Lambda(x) \) are arbitrary gauge functions with commuting derivatives [recall (2.104)].
The experimentally measurable physical fields are given by the gauge invariant ant-
isymmetric combination of derivatives (2.80):

\[
F_{ab} = \partial_a A_b - \partial_b A_a.
\] (11.121)

The additional derivative terms (11.120) disappear in the antisymmetric combina-
tion (11.121). This suggests that a similar antisymmetric construction exists also
for the connection. The construction is slightly more complicated since the transform-
ination law (11.113) contains also contributions which are linear in the connection.

In a nonabelian gauge theory associated with an internal symmetry which is
independent of the spacetime coordinate \( x \), the covariant field strength \( F_{ab} \) is a
matrix. If \( g \) are the elements of the gauge group and \( D(g) \) a representation of \( g \) in
this matrix space, the field strength transforms like a tensor

\[
F_{ab} \rightarrow F'_{ab} = D(g) F_{ab} D^{-1}(g).
\] (11.122)

The gauge field \( A_a \) behaves under such transformations as

\[
A_a(x) \rightarrow A'_a(x) = D(g) A_a(x) D^{-1}(g) + [\partial_a D(g)] D^{-1}(g),
\] (11.123)
which is the generalization of the gauge transformations (2.103). The covariant field strength with the transformation property (11.122) is obtained from this by forming the nonabelian curl

$$F_{ab} = \partial_a A_b - \partial_b A_a - [A_a, A_b].$$

(11.124)

This kind of gauge transformations and covariant field strengths appear in the non-abelian gauge theories used to describe the vector bosons $W^0, \pm$ and $Z^0$ of weak interactions, where the gauge group is SU(2), and the octet of gluons $G^{1,\ldots,8}$ in the theory of strong interactions, where the gauge group is SU(3). In either case, the representation matrices $D(g)$ belong to the adjoint representation of the gauge group.

Now we observe that the transformation law (11.101) of the affine connection can be written in a way which is completely analogous to the transformation law (11.123) of a non-abelian gauge field. For this we consider $\Gamma^\lambda_{\mu\nu}$ as the matrix elements of a four $4 \times 4$ matrix $\Gamma_\mu$:

$$\Gamma_{\mu\nu}^\lambda = \left(\Gamma_\mu\right)^\lambda_\nu.\quad (11.125)$$

Then we can rewrite (11.102) as the matrix equation

$$\Gamma'_{\mu\nu}(x') = \alpha_\mu^\mu \left[ \alpha \Gamma_\mu(x) \alpha^{-1} + (\partial_\mu \alpha) \alpha^{-1} \right].$$

(11.126)

This equation is a direct generalization of Eq. (11.123) to the case that the symmetry group acts also on the spacetime coordinates. To achieve covariance, the vector index $\mu$ of the gauge field must be transformed accordingly.

Actually this is no surprise if we remember the original purpose of introducing the connection $\Gamma_{\mu\nu}^\lambda$. It served to form covariant derivatives (11.86) and (11.91). Equation (11.126) shows that the connection may be viewed as a non-abelian gauge field of the group of local general coordinate transformations $\alpha_\mu^\nu(x)$. Einstein vectors and tensors in curvilinear coordinates are the associated gauge covariant quantities.

By analogy with the field strength (11.124), we can immediately write down a covariant curl of the matrix field $\Gamma_\mu$:

$$R_{\mu\nu} \equiv \partial_\nu \Gamma_\mu - \partial_\mu \Gamma_\nu - \left[ \Gamma_\mu, \Gamma_\nu \right],$$

(11.127)

which should transform like a tensor under general coordinate transformations. In component form, this tensor reads

$$R_{\mu\nu\lambda}^\sigma = \partial_\mu \Gamma_{\nu\lambda}^\sigma - \partial_\nu \Gamma_{\mu\lambda}^\sigma - \Gamma_{\mu\lambda}^\delta \Gamma_{\nu\delta}^\sigma + \Gamma_{\nu\lambda}^\delta \Gamma_{\mu\delta}^\sigma.$$

(11.128)

The covariance properties of $R_{\mu\nu\lambda}^\kappa$ follow most easily by realizing that in terms of the basic tetrads $e^\mu_a$, the covariant curl has the simple representation

$$R_{\mu\nu\lambda}^\sigma = e^\sigma_a \left( \partial_\mu \partial_\nu e^a_\sigma - \partial_\nu \partial_\mu e^a_\sigma \right) e^\lambda_a = -e^\sigma_\lambda \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^\sigma_a.$$

(11.129)
The first line is obtained directly by inserting \( \Gamma_{\mu\nu}^\lambda = e_a^\lambda \partial_\mu e^a_{\nu} \) into (11.127), and executing the derivatives
\[
\left[ \partial_\mu \Gamma_{\nu\lambda}^\kappa - \left( \Gamma_{\mu\nu}^\lambda \right)_\lambda^\kappa \right] = [\mu \leftrightarrow \nu] \\
= \left( \partial_\mu e_a^\alpha \partial_\nu e^a_\lambda + e_a^\alpha \partial_\nu e_\lambda^a + e_b^\rho \partial_\mu e_\lambda^b \partial_\nu e_\kappa^a \right) - (\mu \leftrightarrow \nu) \\
= e_a^\kappa \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^a_\lambda.
\] (11.130)

The second line in (11.130) is obtained from the first by inserting \( \Gamma_{\mu\nu}^\lambda = -e_a^\alpha \partial_\mu e^a_{\lambda} \) or \( \Gamma_{\nu\mu}^\kappa = e_a^\kappa \partial_\nu e^a_{\mu} \).

We are now ready to realize another property of Minkowski space. Just as this spacetime had a vanishing torsion tensor for any curvilinear parametrization, it also has a vanishing curvature tensor. The representation (11.129) shows that a space has a vanishing torsion tensor for any curvilinear parametrization, it also follows from the obvious fact that
\[
R_{\mu\nu\lambda}^\kappa = e_a^\kappa \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^a_\lambda \equiv 0
\] (11.131)
for the trivial choice of the basis tetrad \( e_a^\kappa = \delta_a^\kappa \). Together with the tensor transformation law (11.132) we find that \( R_{\mu\nu\lambda}^\kappa \) remains identically zero in any curvilinear parametrization of Minkowski space.

From the tetrad expression for \( R_{\mu\nu\lambda}^\kappa \) the tensor transformation law is easily found [using (11.63)]
\[
R_{\mu\nu\lambda}^\kappa (x) \rightarrow R'_{\mu'\nu'\lambda'}^\kappa (x') \\
= e_a^\kappa (x') \left( \partial_{\mu'} \partial_{\nu'} - \partial_{\nu'} \partial_{\mu'} \right) e^a_\lambda (x) \\
= \alpha_{\kappa}^\kappa \alpha_{\mu}^\mu e_a^\kappa (x) \left( \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right) e^a_\lambda \\
= \alpha_{\mu}^\mu \alpha_{\nu}^\nu \alpha_{\lambda}^\lambda \alpha_{\kappa}^\kappa R_{\mu\nu\lambda}^\kappa (x) \\
+ \alpha_{\mu}^\mu \alpha_{\nu}^\nu \alpha_{\lambda}^\lambda \alpha_{\kappa}^\kappa \left[ \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right] \alpha_{\lambda}^\lambda.
\] (11.132)

Since general coordinate transformations are assumed to be smooth, the derivatives in front of \( \alpha_{\lambda}^\lambda \) commute and \( R_{\mu\nu\lambda}^\kappa \) is a proper tensor. It is called the curvature tensor.

By constructing, this curvature tensor is antisymmetric in the first index pair. A property that is not so easy to see is the antisymmetry with respect to the second index pair, if the last index is lowered to \( R_{\mu\nu\kappa} \equiv R_{\mu\nu\lambda}^\kappa g_{\kappa\sigma} \):
\[
R_{\mu\nu\kappa} = -R_{\mu\kappa\nu}.
\] (11.133)

Indeed, if we calculate the difference between the two sides using the definition (11.129) we find
\[
R_{\mu\nu\lambda} + R_{\mu\kappa\lambda} = e_a^\kappa \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^a_\lambda + e_a^\lambda \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^a_\kappa \\
= \partial_\mu \partial_\nu \left( e_a^\kappa e^a_\lambda \right) - \partial_\nu \partial_\mu \left( e_a^\kappa e^a_\lambda \right) \\
= \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) g_{\kappa\lambda}.
\] (11.134)
11.6 Curvature Tensor as Covariant Curl of Affine Connection

The physical observability requires the metric

\[ g_{\lambda \kappa}(x) = \frac{\partial x^a}{\partial x^\lambda} \frac{\partial x^a}{\partial x^\kappa} \quad (11.135) \]

to be a smooth single-valued function, so that it satisfies the integrability condition

\[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) g_{\lambda \kappa} = 0. \quad (11.136) \]

Inserting this into Eq. (11.134) proves that the Riemannian-Cartan curvature tensor is indeed antisymmetric in the last two indices.\(^2\) According to the definition given after Eq. (2.88), this antisymmetry is therefore a Bianchi identity.

An integrability assumption of the type (11.136) must also be imposed upon the affine connection to make it a physically observable field:

\[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \Gamma_{\lambda \kappa \delta} = 0. \quad (11.137) \]

This integrability condition gives rise to the famous Bianchi identity of Riemann-Cartan spacetimes to be derived in Section 12.5.

It should be pointed out that a nonvanishing curvature tensor has the consequence that covariant derivatives no longer commute. If we form

\[ D_\mu D_\nu v_\lambda - D_\nu D_\mu v_\lambda \quad (11.138) \]

we find using (11.86), (11.91), and (11.128):

\[ D_\nu D_\mu v_\kappa - D_\mu D_\nu v_\kappa = -R_{\nu \mu \lambda}^{\ k} v_\kappa - 2S_{\nu \mu}^\rho R_\rho v_\lambda, \]

\[ D_\nu D_\mu v_\kappa - D_\mu D_\nu v_\kappa = R_{\nu \mu \lambda}^{\ k} v_\lambda - 2S_{\nu \mu}^\rho D_\rho v_\kappa. \quad (11.139) \]

Since \( R_{\mu \nu \lambda}^{\ k} \) is a tensor it can be contracted with the metric tensor to form covariant quantities of lower rank. There are two possibilities

\[ R_{\nu \mu} \equiv R_{\kappa \nu \mu} \quad (11.140) \]

called the \textit{Ricci tensor} and

\[ R = R_{\nu \mu} g^{\mu \nu} \quad (11.141) \]

called the \textit{scalar curvature}. A combination of both

\[ G_{\nu \mu} \equiv R_{\nu \mu} - \frac{1}{2} g_{\nu \mu} R \quad (11.142) \]

was introduced by Einstein and is therefore called the \textit{Einstein curvature tensor}. It can also be written as

\[ G^{\nu \mu} = \frac{1}{4} \epsilon^{\mu \alpha \beta \gamma} \epsilon_\alpha^\nu \delta_\gamma^\delta \delta_\tau^\tau R_{\beta \gamma \delta \tau}, \quad (11.143) \]

where \( \epsilon^{\mu \nu \lambda \kappa} \) is the contravariant version of the Levi-Civita tensor defined in Appendix 11A. The equality between (11.143) and (11.142) follows directly from the curved-spacetime version of the identity (1A.24).

\(^2\)In more general geometries, where \( D_\mu g_{\lambda \kappa} = -Q_{\mu \lambda \kappa} \neq 0 \) (see the remark [2] in Notes and References), there may also exist a nonzero symmetric part

\[ R_{\mu \nu \lambda \kappa} + R_{\mu \nu \kappa \lambda} = [D_\mu Q_{\nu \lambda \kappa} - (\nu \leftrightarrow \mu)] + 2S_{\mu \nu}^\rho Q_{\rho \lambda \kappa}. \]
11.7 Riemann Curvature Tensor

Actually, Einstein worked with a related tensor which deals exclusively with the Riemann part of the connection and the curvature tensor. Since the contortion $K_{\mu\nu}^\lambda$ is a tensor, the Riemann part $\tilde{\Gamma}_{\mu\nu}^\lambda$ of $\Gamma_{\mu\nu}^\lambda$ has the same transformation properties (11.105) as $\Gamma_{\mu\nu}^\lambda$, and we can form the Riemann curvature tensor

$$\bar{R}_{\mu\nu\lambda\kappa} = \partial_\mu \bar{\Gamma}_{\nu\lambda\kappa} - \partial_\nu \bar{\Gamma}_{\mu\lambda\kappa} - \bar{\Gamma}_{\mu\lambda\rho} \bar{\Gamma}_{\nu\rho\kappa} + \bar{\Gamma}_{\nu\lambda\rho} \bar{\Gamma}_{\mu\rho\kappa}.$$ (11.144)

Contrary to $R_{\mu\nu\lambda\kappa}$ in Eq. (11.128), this curvature tensor can be expressed completely in terms of derivatives of the metric [recall (11.22), (11.24)]. The difference between the two tensors is the following function of the contortion tensor

$$R_{\mu\nu\lambda\kappa} - \bar{R}_{\mu\nu\lambda\kappa} = \bar{D}_\mu K_{\nu\lambda\kappa} - \bar{D}_\nu K_{\mu\lambda\kappa} - \left( K_{\mu\lambda\rho} K_{\nu\rho\kappa} - K_{\nu\lambda\rho} K_{\mu\rho\kappa} \right),$$ (11.145)

where $\bar{D}_\mu$ denotes a covariant derivative which is formed with only the Christoffel part of the connection. Note that the Riemann part of the curvature tensor $R_{\mu\nu\lambda\kappa}$ has the same antisymmetry in the first and second index pairs as $R_{\mu\nu\kappa\lambda}$. For the first pair $\mu\nu$ this follows directly from the definition (11.144). For the second index pair $\lambda\kappa$ it follows from (11.133) and the antisymmetry of the contortion tensor $K_{\nu\lambda\kappa}$ in the second index pair.

In addition the curvature tensor $\bar{R}_{\mu\nu\lambda\kappa}$ is symmetric under the exchange of the first and the second index pair

$$\bar{R}_{\mu\nu\lambda\kappa} = \bar{R}_{\lambda\kappa\mu\nu}.$$ (11.146)

This can be shown by expressing the first two terms in (11.144) as derivatives of the metric tensor

$$\bar{R}_{\mu\nu\lambda\kappa} = \frac{1}{2} \left( g_{\kappa\delta} \partial_\mu \frac{g^{\delta\sigma}}{2} \left( \partial_\nu g_{\lambda\sigma} + \partial_\lambda g_{\nu\sigma} - \partial_\sigma g_{\nu\lambda} \right) - [\mu \leftrightarrow \nu] \right) - g_{\kappa\delta} \left( \bar{\Gamma}_{\mu\lambda\rho} \bar{\Gamma}_{\nu\rho\delta} - \bar{\Gamma}_{\nu\lambda\rho} \bar{\Gamma}_{\mu\rho\delta} \right),$$ (11.147)

and using (11.98) to express $\partial_\mu g^{\delta\sigma}$ in terms of Christoffel symbols,

$$g_{\kappa\delta} \partial_\mu g^{\delta\sigma} = - \left( \partial_\mu g_{\kappa\delta} \right) g^{\delta\sigma} = - \left( \bar{\Gamma}_{\mu\kappa} \sigma g_{\tau\delta} + \bar{\Gamma}_{\mu\delta} \sigma g_{\kappa\tau} \right) g^{\delta\sigma} = - \bar{\Gamma}_{\mu\kappa} \sigma - \bar{\Gamma}_{\mu\delta} \sigma g^{\delta\sigma}$$ (11.148)

we derive

$$\bar{R}_{\mu\nu\lambda\kappa} = \frac{1}{2} \left( \left( \partial_\mu \partial_\nu g_{\lambda\sigma} - \partial_\nu \partial_\mu g_{\lambda\sigma} \right) - (\mu \leftrightarrow \nu) \right) - \left( \bar{\Gamma}_{\mu\kappa} \sigma g_{\tau\delta} + \bar{\Gamma}_{\mu\delta} \sigma g_{\kappa\tau} \right) g^{\delta\sigma} - \bar{\Gamma}_{\mu\kappa} \sigma - \bar{\Gamma}_{\mu\delta} \sigma g^{\delta\sigma}$$ (11.149)

A further use of relation (11.98) brings the second line to

$$- \frac{1}{2} \left( \left( \bar{\Gamma}_{\mu\kappa} \sigma + g^{\delta\sigma} \bar{\Gamma}_{\mu\delta\kappa} \right) \left( \bar{\Gamma}_{\nu\lambda\sigma} + \bar{\Gamma}_{\nu\sigma\lambda} \right) + (\lambda \leftrightarrow \nu) - \bar{\Gamma}_{\sigma\nu\lambda} - \bar{\Gamma}_{\sigma\lambda\nu} \right) - \{\mu \leftrightarrow \nu\},$$
and we find that almost all terms cancel, due to the symmetry of $\bar{\Gamma}_{\mu \nu \lambda}$ in $\mu \nu$. Only

$$- \left( \bar{\Gamma}_{\mu \kappa}^{\sigma} \bar{\Gamma}_{\nu \lambda \sigma} + \bar{\Gamma}_{\mu \delta \kappa}^{\delta} \bar{\Gamma}_{\nu \nu \lambda} \right) + (\mu \leftrightarrow \nu)$$

survives, whose second term cancels the third line in (11.149), bringing $\bar{R}_{\mu \nu \lambda \kappa}$ to the form

$$\bar{R}_{\mu \nu \lambda \kappa} = \frac{1}{2} \left[ \left( \partial_\mu \partial_\lambda g_{\nu \kappa} - \partial_\mu \partial_\kappa g_{\nu \lambda} \right) - (\mu \leftrightarrow \nu) \right] - \left( \bar{\Gamma}_{\mu \kappa}^{\sigma} \bar{\Gamma}_{\nu \lambda \sigma} - \bar{\Gamma}_{\nu \kappa}^{\sigma} \bar{\Gamma}_{\mu \lambda \sigma} \right). \quad (11.150)$$

This expression shows manifestly the symmetry $\mu \nu \leftrightarrow \lambda \kappa$ as a consequence of the integrability property

$$\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) g_{\lambda \kappa} = 0.$$

The same property makes $\bar{R}_{\mu \nu \lambda \kappa}$ antisymmetric under $\mu \leftrightarrow \nu$ by Eq. (11.134), as deduced from Eqs. (11.134) and (11.145).

By contracting (11.150) with $g^{\nu \lambda} g^{\mu \kappa}$, we can derive the following compact expression for the curvature scalar

$$\sqrt{-g} \bar{R} = \bar{\lambda} \left[ \left( g^{\nu \lambda} \sqrt{-g} \left( \bar{\Gamma}_{\mu \nu}^{\lambda} - \delta_\mu^{\lambda} \bar{\Gamma}_{\nu \kappa} \right) \right) + \sqrt{-g} g^{\mu \nu} \left( \bar{\Gamma}_{\mu \nu}^{\lambda} \bar{\Gamma}_{\nu \kappa} - \bar{\Gamma}_{\mu \nu}^{\lambda} \bar{\Gamma}_{\kappa \nu} \right) \right]. \quad (11.151)$$

It is instructive to check this equation as follows: We use the identity (11.148) in the form

$$\partial_\kappa g_{\mu \nu} = g_{\mu \sigma} g_{\nu \tau} \partial_\kappa g^{\sigma \tau} = \bar{\Gamma}_{\kappa \mu \nu} + \bar{\Gamma}_{\kappa \nu \mu}, \quad (11.152)$$

and another identity

$$\partial_\lambda \sqrt{-g} = \frac{1}{2} \sqrt{-g} g^{\sigma \tau} \partial_\lambda g_{\sigma \tau} = \bar{\Gamma}_{\lambda \mu}^{\mu}, \quad (11.153)$$

which follows directly from Eq. (11A.24), to derive

$$\partial_\lambda \left( g^{\mu \nu} \sqrt{-g} \right) = \sqrt{-g} \left[ -g^{\mu \sigma} \bar{\Gamma}_{\lambda \sigma}^{\nu} - g^{\nu \sigma} \bar{\Gamma}_{\lambda \sigma}^{\mu} + g^{\mu \nu} \bar{\Gamma}_{\lambda \sigma}^{\sigma} \right]. \quad (11.154)$$

This allows us to rewrite the first term in (11.151) as

$$\partial_\lambda \left( g^{\mu \nu} \sqrt{-g} \right) \left( \bar{\Gamma}_{\mu \nu}^{\lambda} - \delta_\mu^{\lambda} \bar{\Gamma}_{\nu \kappa} \right) \quad (11.155)$$

$$= \sqrt{-g} \left[ -g^{\mu \sigma} \bar{\Gamma}_{\sigma \lambda}^{\nu} - g^{\nu \sigma} \bar{\Gamma}_{\sigma \lambda}^{\mu} + g^{\mu \nu} \bar{\Gamma}_{\sigma \lambda}^{\sigma} \right] \left( \bar{\Gamma}_{\mu \nu} - \delta_\mu^{\lambda} \bar{\Gamma}_{\nu \kappa} \right)$$

$$= -2 \sqrt{-g} \left[ \Gamma_{\mu \sigma} g^{\mu \lambda} g^{\nu \sigma} \bar{\Gamma}_{\lambda \kappa}^{\sigma} - \bar{\Gamma}_{\lambda \sigma}^{\sigma} g^{\lambda \kappa} \Gamma_{\mu \nu} \right].$$

Using this, Eq. (11.151) becomes

$$\sqrt{-g} \bar{R} = \sqrt{-g} \left[ g^{\mu \nu} g^{\lambda \kappa} \left( \partial_\lambda \Gamma_{\kappa \mu \nu} - \partial_\mu \Gamma_{\nu \lambda \kappa} - g^{\sigma \tau} \bar{\Gamma}_{\lambda \sigma \mu} \bar{\Gamma}_{\kappa \tau \nu} + \bar{\Gamma}_{\sigma \lambda}^{\sigma} \Gamma_{\kappa \lambda \nu} \right) \right], \quad (11.156)$$

which is the contraction of the defining equation (11.144) for the Riemann curvature tensor with $\delta_\mu^{\sigma} g^{\nu \lambda}$. 

Appendix 11A  Curvilinear Versions of Levi-Civita Tensor

In Appendix 1A we have listed to the properties of the Levi-Civita tensor $\epsilon^{a_1\ldots a_D}$ in Euclidean as well as Minkowski space. These properties acquire little change if the spaces are reparametrized with curvilinear coordinates. To be specific, we consider only a four-dimensional Minkowski spacetime whose metric arises from a coordinate transformation of (1A.21). The same formulas hold also if the spacetime is curved. The curvilinear Levi-Civita tensor is

$$\epsilon^{\mu_1\ldots\mu_D} = \frac{1}{\sqrt{-g}} \epsilon^{\mu_1\ldots\mu_D},$$  \hspace{1cm} (11A.1)$$

where

$$-g \equiv \det (-g_{\mu\nu})$$  \hspace{1cm} (11A.2)$$

is the positive determinant of $-g_{\mu\nu}$, and $\sqrt{-g}$ is the positive square root of it. Just as $\epsilon^{a_1\ldots a_D}$ was a pseudotensor under Lorentz transformations [recall (1A.11)], $\epsilon^{\mu_1\ldots\mu_D}$ is a pseudotensor under general coordinate transformations, which transform

$$\epsilon^{\mu_1\ldots\mu_D} \rightarrow \alpha^{\mu_1}_{\nu_1} \cdots \alpha^{\mu_D}_{\nu_D} \epsilon^{\nu_1\ldots\nu_D} = \det (\alpha) \epsilon^{\mu_1\ldots\mu_D}. $$ \hspace{1cm} (11A.3)$$

Since $g_{\mu\nu}$ is transformed as

$$g_{\mu\nu} \rightarrow \alpha_{\mu}^{\lambda} \alpha_{\nu}^{\kappa} g_{\lambda\kappa},$$ \hspace{1cm} (11A.4)$$

its determinant behaves like

$$g \rightarrow \det \left( \alpha_{\mu}^{\lambda} \right)^2 g = \det (\alpha_{\nu}^{\mu})^{-2} g.$$ \hspace{1cm} (11A.5)$$

Hence

$$\epsilon^{\mu_1\ldots\mu_D} \rightarrow \frac{\det (\alpha_{\mu}^{\nu})}{\det (\alpha_{\nu}^{\mu})} \epsilon^{\mu_1\ldots\mu_D},$$ \hspace{1cm} (11A.6)$$

showing the pseudotensor property.

The same thing holds for the tensor

$$e_{\mu_1\ldots\mu_D} = \sqrt{-g} \epsilon_{\mu_1\ldots\mu_D},$$ \hspace{1cm} (11A.7)$$

It arises from $\epsilon^{\nu_1\ldots\nu_D}$ by multiplication with $g_{\mu_1\nu_1} \cdots g_{\mu_D\nu_D}$ as it should.

The co- and contravariant antisymmetric tensors $e_{\mu_1\ldots\mu_D}$, $\epsilon^{\mu_1\ldots\mu_D}$ share an important property with the symmetric tensors $g_{\mu_1\nu_2}$, $g^{\mu_1\nu_2}$. Just as those, they are invariant under covariant differentiation:

$$D_\lambda e_{\mu_1\ldots\mu_D} = 0, \quad D_\lambda \epsilon^{\mu_1\ldots\mu_D} = 0.$$ \hspace{1cm} (11A.8)$$

Indeed, since $e_{\mu_1\ldots\mu_D}$ is a tensor, we can write this equation explicitly as

$$\partial_\lambda e_{\mu_1\ldots\mu_D} = \Gamma^{\nu_1}_{\lambda\mu_1} e_{\nu_2\ldots\mu_D} + \Gamma^{\nu_2}_{\lambda\mu_2} e_{\nu_1\nu_2\ldots\mu_D} + \cdots + \Gamma^{\nu_D}_{\lambda\mu_D} e_{\mu_1\nu_2\ldots\nu_D}. $$ \hspace{1cm} (11A.9)$$

H. Kleinert, GRAVITY WITH TORSION
Using $e_{\mu_1...\mu_D} = \sqrt{-g} e_{\mu_1...\mu_D}$, the left-hand side can be rewritten as
\[
\frac{1}{\sqrt{-g}} \left( \partial_\lambda \sqrt{-g} \right) e_{\mu_1...\mu_D},
\]
from which the equality follows by using the trivial identity
\[
\delta_{\sigma\tau} e_{\mu_1...\mu_D} = \delta_{\sigma\mu_1} e_{\tau\mu_2...\mu_D} + \delta_{\sigma\mu_2} e_{\mu_1\tau...\mu_D} + \ldots + \delta_{\sigma\mu_D} e_{\mu_1\mu_2...\tau},
\]
after turning it into the covariant form
\[
g_{\sigma\tau} e_{\mu_1...\mu_D} = g_{\sigma\mu_1} e_{\tau\mu_2...\mu_D} + g_{\sigma\mu_2} e_{\mu_1\tau...\mu_D} + \ldots + g_{\sigma\mu_D} e_{\mu_1\mu_2...\tau},
\]
and multiplying it by $g_{\sigma\delta} \Gamma^\lambda_{\delta\tau}$.

An important consequence of the vanishing covariant derivative of the antisymmetric tensors in Eq. (11A.8) is that antisymmetric products satisfy the covariant version chain rule of differentiation without an extra term. For instance, the vector product in three curved dimensions
\[
(x \times w)_{\mu} = e_{\mu\lambda\kappa} x^{\lambda} w^{\kappa}
\]
has the covariant derivative
\[
D_\sigma (x \times w) = D_\sigma x \times w + x \times D_\sigma w,
\]
just as in flat space. The same rule applies, of course, to the scalar product
\[
x \cdot w = g_{\mu\nu} v^{\mu} w^{\nu},
\]
as a consequence of the vanishing covariant derivative of the metric in Eq. (11.95):
\[
D_\sigma (x \cdot w) = D_\sigma x \cdot w + x \cdot D_\sigma w.
\]
The determinant of arbitrary tensor $t_{\mu\nu}$ is given by a formula similar to (1A.9)
\[
\det (t_{\mu\nu}) = \frac{1}{D!} e_{\mu_1...\mu_D}^{\nu_1...\nu_D} t_{\mu_1\nu_1} \ldots t_{\mu_D\nu_D} = \frac{g}{D!} e_{\mu_1...\mu_D}^{\nu_1...\nu_D} t_{\mu_1\nu_1} \ldots t_{\mu_D\nu_D}.
\]
The determinant of $t_{\mu}^{\nu}$, on the other hand, is equal to
\[
\det (t_{\mu}^{\nu}) = \frac{1}{D!} e_{\mu_1...\mu_D}^{\nu_1...\nu_D} t_{\mu_1}^{\nu_1} \ldots t_{\mu_D}^{\nu_D} = \frac{1}{D!} e_{\mu_1...\mu_D}^{\nu_1...\nu_D} t_{\mu_1}^{\nu_1} \ldots t_{\mu_D}^{\nu_D},
\]
in agreement with the relation $\det (t_{\mu}^{\nu}) = \det (t_{\mu\lambda} g^{\lambda\nu}) = \det (t_{\mu\nu}) g^{-1}$.

The covariant tensors $e_{\nu_1...\nu_D}^{\mu_1...\mu_D}$ are useful for writing down explicitly the cofactors $M_{\nu}^{\mu}$ in the expansion of a determinant.
\[
\det (t_{\mu}^{\nu}) = \frac{1}{D!} t_{\mu}^{\nu} M_{\nu}^{\mu}.
\]
By comparison with Eq. (11A.18) we identify:

\[ M_{\nu_1}^{\mu_1} = -\frac{1}{(D-1)!}\epsilon^{\mu_1\ldots\mu_D}\epsilon_{\nu_1\ldots\nu_D} t_{\mu_2\nu_2} \ldots t_{\mu_D\nu_D}. \]  

(11A.20)

The inverse of the matrix \( t_{\mu}^{\nu} \) has then the explicit form

\[ \left( t^{-1} \right)^\mu_\nu = \frac{1}{\det t_{\mu}^{\nu}} M^{\mu_\nu}. \]  

(11A.21)

For a determinant \( \det (t_{\mu\nu}) \) we find, similarly,

\[ \det (t_{\mu\nu}) = \frac{1}{D} t_{\mu\nu} M^{\mu\nu}, \]  

(11A.22)

with

\[ M^{\mu_1\nu_1} = \frac{1}{(D-1)!}\epsilon^{\mu_1\ldots\mu_D}\epsilon^{\nu_1\ldots\nu_D} t_{\mu_2\nu_2} \ldots t_{\mu_D\nu_D} = \det (t_{\mu\nu}) \left( t^{-1} \right)^{\mu_1\nu_1}. \]  

(11A.23)

This equation is useful for calculating variations of the determinant \( g \) upon variations of the metric \( g_{\mu\nu} \), which will be needed later in Eq. (15.23). Inserting \( g_{\mu\nu} \) into (11A.17) and using the first line of (11A.23), we find immediately

\[ \delta g = \frac{1}{D!}\epsilon^{\mu_1\ldots\mu_D}\epsilon^{\nu_1\ldots\nu_D}\delta \left( g_{\mu_1\nu_1} g_{\mu_2\nu_2} \ldots g_{\mu_D\nu_D} \right) \]
\[ = \frac{1}{(D-1)!}\epsilon^{\mu_1\ldots\mu_D}\epsilon^{\nu_1\ldots\nu_D}\delta g_{\mu_1\nu_1} g_{\mu_2\nu_2} \ldots g_{\nu_D\nu_D} = \delta g_{\mu\nu} M^{\mu\nu} = \det (g_{\mu\nu}) g^{\mu\nu} \delta g_{\mu\nu} = \delta g_{\mu\nu} \]  

(11A.24)

The identity \( g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda \) implies opposite signs of co- and contravariant variations:

\[ g^{\lambda\mu}\delta g_{\nu\mu} = -g_{\nu\mu} \delta g^{\lambda\mu}, \]  

(11A.25)

so that \( \delta g^{\mu\nu} = -g^{\lambda\mu}g^{\nu\lambda}\delta g_{\mu\nu} \) and

\[ \delta g = gg^{\mu\nu} \delta g_{\mu\nu} = -gg_{\mu\nu} \delta g^{\mu\nu}. \]  

(11A.26)

Another way of deriving this result employs the identity valid for any nonsingular matrix \( A \):

\[ \det A = e^{\text{tr} \log A}, \]  

(11A.27)

from which we find

\[ \delta \det A = \det A \delta (\text{tr} \log A) = \det A \text{tr}(A^{-1} \delta A). \]  

(11A.28)

Replacing \( A \) by the metric gives directly (11A.24).
Notes and References

Torsion and Curvature from Defects

In the last chapter we have seen that a Minkowski space has neither torsion nor curvature. The absence of torsion follows from its tensor property, which was a consequence of the commutativity of derivatives in front of the infinitesimal translation field

\[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \xi^\kappa (x) = 0. \]  

(12.1)

The absence of curvature, on the other hand, was a consequence of the integrability condition (11.32) of the transformation matrices

\[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \alpha^\kappa \lambda (x) = 0. \]  

(12.2)

Infinitesimally, this implies that

\[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\lambda \xi^\kappa (x) = 0, \]  

(12.3)

i.e., that derivatives commute in front of derivatives of the infinitesimal translation field.

The situation is similar to those in electromagnetism in Chapter 4. Arbitrary gauge transformations (2.103) whose gauge function \( \Lambda(x) \) has commuting derivatives [see (2.104)] do not change the electromagnetic fields in spacetime. In particular, a field-free spacetime remains field-free. In Subsection 4.3 we have seen however, that it is possible to generate thin nonzero magnetic field tubes in a field-free space by performing multivalued gauge transformations which violate Schwarz' integrability conditions. It is useful to imagine these coordinate transformations as being plastic distortions of a world crystal. The ordinary single-valued coordinate transformations correspond to elastic distortions of the world crystal which do not change the geometry represented by the defects.

In Chapter 9 we have shown that the theoretical description of crystals with defects is very similar to that of electromagnetism in terms of a multivalued scalar field. This suggests a simple way of constructing general affine spaces with torsion or curvature or both from a Minkowski spacetime by performing multivalued coordinate transformations which do not satisfy (12.1), (12.3).
12.1 Multivalued Infinitesimal Coordinate Transformations

Let us study the properties of a spacetime at which we can arrive from basis tetrads $e_a^\mu = \delta_a^\mu$ via such *infinitesimal multivalued* coordinate transformations $\xi^a(x)$. According to (11.69), the new basis tetrads are

$$
e_a^\mu = \delta_a^\mu - \partial_a \xi^\mu$$

$$e^a_\mu = \delta^a_\mu + \partial_a \xi^\mu$$

and the metric is

$$g_{\mu \nu} = e_a^\mu e_a^\nu = \eta_{\mu \nu} + \left( \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \right),$$

where $\eta_{\mu \nu}$ denotes the Minkowski metric (1.29) which needs here a different notation due to our convention that Greek subscripts refer to curvilinear coordinates.

Inserting the basis tetrads (12.4) into Eq. (11.92) we find the affine connection

$$\Gamma_{\mu \nu}^\lambda = \partial_\mu \partial_\nu \xi^\lambda,$$

and from this the torsion and curvature tensors

$$S_{\mu \nu}^\lambda = \frac{1}{2} \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \xi^\lambda,$$

$$R_{\mu \nu \lambda}^\kappa = \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\lambda \xi^\kappa.$$  

(12.7)

Since $\xi^\lambda$ are infinitesimal displacements, we can lower the last index in both equations with a mistake quadratic in $\xi^\kappa$, and thus negligible for small $\xi^\kappa$, so that

$$\Gamma_{\mu \nu \lambda} = \partial_\mu \partial_\nu \xi_\lambda, \quad S_{\mu \nu \lambda} = \frac{1}{2} \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \xi_\lambda, \quad R_{\mu \nu \lambda \kappa} = \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\lambda \xi_\kappa.$$

(12.8)

The curvature tensor is trivially antisymmetric in the first two indices [as in (11.134)].

For singular $\xi(x)$, the metric and the connection are, in general, also singular. This would cause difficulties in performing consistent length measurements and parallel displacements. To avoid such difficulties, Einstein postulated that the metric $g_{\mu \nu}$ and the connection $\Gamma_{\mu \nu}^\lambda$ should be smooth enough to permit two differentiations which commute with each other as stated earlier in (11.136) and (11.137). For the infinitesimal expressions (12.4) and (12.5), these properties imply that we must consider only such singular coordinate transformations which satisfy the condition

$$\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \left( \partial_\lambda \xi_\kappa + \partial_\kappa \xi_\lambda \right) = 0,$$

$$\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\sigma \partial_\lambda \xi_\kappa = 0.$$  

(12.9)

The integrability conditions (12.9) show again, now for the linearized metric, that the curvature tensor (12.13) is antisymmetric in the last two indices [recall (11.134)].
For completeness, let us also write down the pure Christoffel part of the connection as it follows from inserting (12.5) into (11.24):

\[\bar{\Gamma}_{\mu\nu\kappa} = \frac{1}{2} \{\mu\nu,\kappa\} = \frac{1}{2} \left[ \partial_{\mu} \left( \xi_{\kappa} + \partial_{\kappa} \xi_{\nu} \right) + \partial_{\nu} \left( \xi_{\mu} + \partial_{\mu} \xi_{\kappa} \right) - \partial_{\kappa} \left( \partial_{\mu} \xi_{\nu} + \partial_{\nu} \xi_{\mu} \right) \right]\]

(12.11)

thus negligible such that

\[\Gamma_{\mu\nu\lambda} = \partial_{\mu} \partial_{\nu} \xi_{\lambda}, \quad S_{\mu\nu\lambda} = \frac{1}{2} \left( \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right) \partial_{\lambda} \xi_{\kappa}, \quad R_{\mu\nu\lambda\kappa} = \left( \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right) \partial_{\lambda} \xi_{\kappa}.\]

(12.12)

For completeness, let us also write down the decomposition (11.119) of the connection into the Christoffel part and the contortion tensor obtained by inserting (12.5) into (11.24):

\[\Gamma_{\mu\nu\kappa} = \{\mu\nu,\kappa\} + K_{\mu\nu\kappa}\]

(12.13)

with

\[\{\mu\nu,\kappa\} = \frac{1}{2} \partial_{\mu} \left( \partial_{\nu} \xi_{\kappa} + \partial_{\kappa} \xi_{\nu} \right) + \frac{1}{2} \partial_{\nu} \left( \xi_{\mu} + \partial_{\mu} \xi_{\kappa} \right) - \frac{1}{2} \partial_{\kappa} \left( \partial_{\mu} \xi_{\nu} + \partial_{\nu} \xi_{\mu} \right)\]

\[K_{\mu\nu\lambda} = \frac{1}{2} \left( \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right) \xi_{\lambda} - \frac{1}{2} \left( \partial_{\nu} \partial_{\lambda} - \partial_{\lambda} \partial_{\nu} \right) \xi_{\mu} + \frac{1}{2} \left( \partial_{\lambda} \partial_{\mu} - \partial_{\mu} \partial_{\lambda} \right) \xi_{\nu} = \frac{1}{2} \partial_{\mu} \left( \partial_{\nu} \xi_{\lambda} - \partial_{\lambda} \xi_{\nu} \right) + \frac{1}{2} \partial_{\nu} \left( \partial_{\mu} \xi_{\lambda} + \partial_{\lambda} \xi_{\mu} \right) - \frac{1}{2} \partial_{\lambda} \left( \partial_{\mu} \xi_{\nu} + \partial_{\nu} \xi_{\mu} \right).\]

(12.14)

From the Christoffel symbol we find the Riemann curvature tensor

\[\bar{R}_{\mu\nu\lambda\kappa} = \frac{1}{2} \partial_{\mu} \left[ \partial_{\nu} \left( \partial_{\lambda} \xi_{\kappa} + \partial_{\kappa} \xi_{\lambda} \right) + \partial_{\lambda} \left( \partial_{\nu} \xi_{\kappa} + \partial_{\kappa} \xi_{\nu} \right) - \partial_{\kappa} \left( \partial_{\nu} \xi_{\lambda} + \partial_{\lambda} \xi_{\nu} \right) \right] = -\frac{1}{2} \partial_{\nu} \left[ \partial_{\mu} \left( \partial_{\lambda} \xi_{\kappa} + \partial_{\kappa} \xi_{\lambda} \right) + \partial_{\lambda} \left( \partial_{\nu} \xi_{\kappa} + \partial_{\kappa} \xi_{\nu} \right) - \partial_{\kappa} \left( \partial_{\mu} \xi_{\lambda} + \partial_{\lambda} \xi_{\mu} \right) \right].\]

(12.15)

Due to the integrability condition (12.10) the first terms in each line cancel and this becomes

\[\bar{R}_{\mu\nu\lambda\kappa} = \frac{1}{2} \left\{ \left[ \partial_{\mu} \partial_{\lambda} \left( \partial_{\nu} \xi_{\kappa} + \partial_{\kappa} \xi_{\nu} \right) - \left( \mu \leftrightarrow \nu \right) \right] \right\}.\]

(12.16)

In order to understand the geometric properties of such a space generated by the infinitesimal singular transformations

\[x^{a} \rightarrow x^{\mu} = \left[ x^{a} - \xi^{a} \left( x^{b} \right) \right] \delta_{a}^{\mu}\]

(12.17)

we like to point out that such transformations are encountered in the context of crystalline defects. There, one considers infinitesimal displacements of atoms given by

\[x_{i} \rightarrow x'_{i} = x_{i} + u_{i}(x)\]

(12.18)
where $x'_i$ are the shifted positions, as seen from an ideal reference crystal. If we change the point of view to an intrinsic description, i.e., if we measure coordinates by counting the number of atomic steps *within* the distorted crystal, then the atoms of the ideal reference crystal are displaced by

$$x_i \to x'_i = x_i - u_i(x).$$  \hfill (12.19)

This is the same as (12.17). Hence the non-commutativity of derivatives in front of singular coordinate changes $\xi^a(x^\lambda)$ is completely analogous to that in front of crystal displacements $u_i(x)$. In the crystals this was a signal for the presence of defects. For the purpose of a better visualization, let us restrict our consideration to the three-dimensional Euclidian subspace of the Minkowski space. Then we have to identify the physical coordinates of material points $x^a$ for $a = 1, 2, 3$ with the previous spatial coordinates $x_i$ for $i = 1, 2, 3$ and $\partial^a = \partial/\partial x^a (a = i)$ with the previous derivatives $\partial_i$. The infinitesimal translations in (11.142), $\xi^a(x^\lambda)$ are equal to the displacements $u_i(x)$ such that the basis tetrads are

$$e'_a = \delta^a_i - \partial^a u_i, \quad e^a_i = \delta^a_i + \partial^a u_i$$  \hfill (12.20)

and the metric becomes, to linear approximation,

$$g_{ij} = e_at^a_j = \delta_{ij} + \partial_i u_j + \partial_j u_i.$$  \hfill (12.21)

Apart from the trivial unit matrix it coincides with twice the strain tensor $u_{ij} = \frac{1}{2} \left( \partial_i u_j + \partial_j u_i \right)$. The connection is simply

$$\Gamma_{ijk} = \partial_j \partial_k u_i$$  \hfill (12.22)

with torsion and curvature tensors

$$S_{ijk} = \frac{1}{2} \left( \partial_i \partial_j - \partial_j \partial_i \right) u_k, \quad R_{ijkl} = \left( \partial_i \partial_j - \partial_j \partial_i \right) \partial_k u_l.$$  \hfill (12.23)

The integrability conditions read

$$\left( \partial_i \partial_j - \partial_j \partial_i \right) \partial_k (u_l + \partial_l u_k) = 0,$$  \hfill (12.24)

$$\left( \partial_i \partial_j - \partial_j \partial_i \right) \partial_k (u_i + \partial_i u_k) = 0,$$  \hfill (12.25)

$$\left( \partial_i \partial_j - \partial_j \partial_i \right) \partial_k (u_i - \partial_i u_k) = 0.$$  \hfill (12.26)

They state that the strain tensor, its derivative, and the derivative of the local rotation field are all twice-differentiable single-valued functions everywhere. It was

\footnote{When working with four-vectors it is conventional to consider the upper indices as physical components. In purely three dimensional calculations one usually employs the metric $g_{ab} = \delta_{ab}$ such that $x^{a=1}$ and $x_i$ are the same.}
argued that this should be true in a crystal. We can take advantage of the first
condition and write the curvature tensor alternatively as

\[ R_{ijkl} = \left( \partial_i \partial_j - \partial_j \partial_i \right) \frac{1}{2} \left( \partial_k u_l - \partial_l u_k \right). \]  

(12.27)

The antisymmetry in \( ij \) and \( kl \) suggests, in three dimensions, the introduction of a
tensor of second rank analogous to (11.143)

\[ G_{ji} \equiv \frac{1}{4} \epsilon_{ikl} \epsilon_{jmn} R^{klmn}, \]  

(12.28)

where \( \epsilon_{ijk} \) and was defined in Eq. (11A.7). The tensor \( G_{ji} \) coincides with the Einstein
tensor (11.142) due to the identity (1A.17). To linear approximation, \( G_{ij} \) becomes
due to (12.23):

\[ G_{ij} = \epsilon_{ikl} \partial_k \partial_l \left( \frac{1}{2} \epsilon_{jmn} \partial_m u_n \right). \]  

(12.29)

The second factor is the local rotation \( \omega_j = \frac{1}{2} \epsilon_{jmn} \partial_m u_n \), and we see that the Einstein
curvature tensor can be written as

\[ G_{ji} = \epsilon_{ikl} \partial_k \partial_l \omega_j. \]  

(12.30)

Let us also form the Einstein tensor \( \tilde{G}_{ij} \) associated with the Riemann curvature
tensor \( \tilde{R}_{ijkl} \). Using (12.16) we find

\[ \tilde{G}_{ji} = \epsilon_{ikl} \epsilon_{jmn} \partial_k \partial_m \frac{1}{2} \left( \partial_l u_n + \partial_n u_l \right). \]  

(12.31)

In the discussion of crystal defects we have introduced the following measures for
the non-commutativity of derivatives. The dislocation density

\[ \alpha_{ij} = \epsilon_{ikl} \partial_k \partial_l u_j \]  

(12.32)

the disclination density

\[ \theta_{ij} = \epsilon_{ikl} \partial_k \partial_l \omega_j \]  

(12.33)

and the defect density

\[ g_{ij} = \epsilon_{ikl} \epsilon_{jmn} \partial_k \partial_m u_{lm}. \]  

(12.34)

Comparison with (12.16) shows that \( \alpha_{ij} \) is directly related to the torsion tensor
\( S_{kl}^i = \frac{1}{2} \left( \Gamma_{ki}^i - \Gamma_{ik}^i \right) \):

\[ \alpha_{ij} \equiv \epsilon_{ikl} \Gamma_{klj} \equiv \epsilon_{ikl} S_{klj}. \]  

(12.35)
Hence torsion is a measure of the translational defects contained in the multivalued coordinate transformations, which may be pictured as combinations of elastic plus plastic distortions of a world crystal.

We can also use the decomposition (11.119) and write, due to the symmetry of the Christoffel symbol \( \{kl,j\} \) in \( kl \):

\[
\alpha_{ij} = \epsilon_{ikl} K_{klj}, \quad (12.36)
\]

where \( K_{klj} \) is the contortion tensor. In terms of the displacement field \( u(\mathbf{x}) \),

\[
K_{ijk} = \frac{1}{2} \partial_j \left( \partial_j u_k - \partial_k u_j \right) - \frac{1}{2} \left[ \partial_j \left( \partial_k u_j + \partial_j u_k \right) - (j \leftrightarrow k) \right]
= \partial_i \omega_{jk} - \left( \partial_j u_{ki} - \partial_k u_{ji} \right). \quad (12.37)
\]

Since \( K_{ijk} \) is antisymmetric in \( lj \), it is useful to introduce the tensor of second rank called Nye’s contortion tensor

\[
K_{ln} = \frac{1}{2} K_{klj} \epsilon_{ljn}. \quad (12.38)
\]

Inserting this into (12.36) we see that

\[
\alpha_{ij} = -K_{ji} + \delta_{ij} K_{ll}. \quad (12.39)
\]

In terms of the displacement and rotation fields, one has

\[
K_{il} = \partial_i \omega_l - \epsilon_{ljk} \partial_j u_{kj}. \quad (12.40)
\]

Consider now the disclination density \( \theta_{ij} \). Comparing (12.34) with (12.30) we see that it coincides exactly with the Einstein tensor \( G_{ji} \) formed from the full curvature tensor

\[
\theta_{ij} \equiv G_{ji}. \quad (12.41)
\]

The defect density (12.34), finally, coincides with the Einstein tensor formed from the Riemann curvature tensor.

\[
g_{ij} = \bar{G}_{ij}. \quad (12.42)
\]

Hence we can conclude: A spacetime with torsion and curvature can be generated from a Minkowski spacetime via singular coordinate transformations and is completely equivalent to a crystal which has undergone plastic deformation and is filled with dislocations and disclinations.

In Minkowski space, the trajectories of free particles are straight lines. In the defected space, they choose the shortest path which is no longer straight since defects may lie in its way. According to Einstein’s theory, the motion of mass points in a gravitational field is governed by the principle of shortest path as defined by the
defected metric $g_{\mu\nu}$. This defected metric contains all gravitational effects which are a consequence of the defects presented in the world crystal. The natural length scale of gravitation is the Planck length which is the following combination of Newton’s gravitational constant $G_N \approx 6.673 \times 10^{-8}$ cm$^3$/g s$^2$, recall (1.3)] with the light velocity $c (\approx 3 \times 10^{10}$ cm/s) and Planck’s constant $\hbar (\approx 1.05459 \times 10^{-27}$ erg/s):

$$l_P = \left( \frac{c^3}{G_N \hbar} \right)^{-1/2} \approx 1.616 \times 10^{-33} \text{cm.} \quad (12.43)$$

The Planck length is an extremely small quantity. It is by a factor $10^{-25}$ smaller than an atom, which is roughly the ratio between the radius of an atom ($\approx 10^{-8}$ cm) and the radius of the solar system ($\approx 10^{10}$ km). Such small distances are at present beyond any experimental resolution. The Planck length $l_P$ may easily be imagined as the lattice constant of a world crystal with defects, without running into experimental contradictions.

The mass whose Compton wavelength is $l_p$,

$$m_p = \frac{\hbar}{cl_p} = \sqrt{\frac{\hbar c}{G_N}} = 1.221047(79) \times 10^{19} \text{GeV} \approx 0.0217671(14) \text{mg} = 1.30138(6) \times 10^{19} m_{\text{proton}}, \quad (12.44)$$

is extremely large.

### 12.2 Examples for Nonholonomic Coordinate Transformations

It may be useful to give a few explicit examples of multivalued mappings $x^\mu(x^a)$ leading from a flat spacetime to a spacetime with curvature and torsion. We shall do so by appealing to actual physical situations. For simplicity, we consider two dimensions. Imagine an ideal crystal with atoms placed at $x^a = (n_1, n_2, n_3) \cdot b$ with infinitesimal lattice constant $b$.

#### 12.2.1 Dislocation

The simplest example for a crystalline defect is the edge dislocation and the edge disclination shown in Fig. 12.2.1. The mapping transforms the lattice points to new distorted positions of which $x^\mu(x^a)$ are the Cartesian coordinates. There exists no one-to-one mapping between the two figures since the excessive atoms in the middle horizontal layer $x^a < 0, x^2 = 0$ have no correspondence in $x^a$ space. In the continuum limit of an infinitesimally small Burgers vector, the mapping can be described by the multivalued function

$$\bar{x}^1 = x^1, \quad \bar{x}^2 = x^2 - \frac{b}{2\pi} \phi, \quad (12.45)$$
12.2 Examples for Nonholonomic Coordinate Transformations

Figure 12.1 Edge dislocation in a crystal associated with a missing semi-infinite plane of atoms. The multivalued mapping from the ideal crystal to the crystal with the dislocation introduces a $\delta$-function type torsion in the image space.

where the

$$\phi(x) = \arctan \frac{x^2}{x^1}$$

with the multivalued definition of the $\arctg$, which on the physical Riemann sheet is equal to $\pm \pi$ for $x^1 = 0, x^2 = \pm \epsilon$. Its differential version is

$$d\bar{x}^1 = dx^1$$

$$d\bar{x}^2 = dx^2 + \frac{b}{2\pi (x^1)^2 + (x^2)^2} \left( x^2 dx^1 - x^1 dx^2 \right)$$

with the basis diads

$$e^a_\mu = \frac{\partial \bar{x}^a}{\partial x^\mu}$$

and

$$e^a_\mu = \begin{pmatrix} 1 & \frac{x^2}{2\pi (x^1)^2 + (x^2)^2} - \frac{b}{2\pi (x^1)^2 + (x^2)^2} \\ 0 & x^1 \end{pmatrix}.$$  

We have used the notation $\bar{x}^a \equiv x^a$ in order to distinguish $x^{a=1,2}$ from $x^{\mu=1,2}$.

Let us now integrate $dx^\mu$ over a Burgers circuit which consists of a closed circuit $C(x^\mu)$ in $x^\mu$-space around the origin,

$$b^a = \oint_{C(x^\mu)} dx^\mu e^a_\mu = \oint_{C(x^\mu)} dx^\mu \frac{\partial \bar{x}^a}{\partial x^\mu} = \oint_{C(x^\mu)} dx^\mu e^a_\mu.$$  

Inserting (12.47) and (12.48) we see that

$$b^1 = \oint_{C(x^\mu)} dx^1 = \oint_{C(x^\mu)} dx^\mu \frac{\partial x^1}{\partial x^\mu} = \oint_{C(x^\mu)} dx^\mu e_1^\mu = 0,$$

$$b^2 = \oint_{C(x^\mu)} dx^2 = \oint_{C(x^\mu)} dx^\mu \frac{\partial x^2}{\partial x^\mu} = \oint_{C(x^\mu)} dx^\mu e^2_\mu = -b.$$  

It is easy to calculate the torsion tensor $S^a_{\mu\nu}$ associated with the multivalued mapping (12.47) and (12.48). Because of its antisymmetry, only $S^1_{12}$ and $S^2_{12}$ are independent. These become

$$S^2_{12} = \partial_1 e^2_2 - \partial_2 e^2_1 = \partial_1 \frac{\partial x^2}{\partial x^1} - \partial_2 \frac{\partial x^2}{\partial x^1} = -b \delta^{(2)}(x),$$
\[ S_{12}^{1} = \partial_{1}e_{2}^{1} - \partial_{2}e_{1}^{1} = \partial_{1}\frac{\partial x^{1}}{\partial x^{2}} - \partial_{2}\frac{\partial x^{2}}{\partial x^{1}} = 0. \quad (12.53) \]

We may write this result with the Burgers vector \( b^{a} = (0, b) \) in the form
\[ S_{\mu\nu}^{a} = b^{a}\delta^{(2)}(x). \quad (12.54) \]

Let us now calculate the curvature tensor for this defect which is
\[ R_{\mu\nu\lambda\kappa} = e_{a\kappa}(\partial_{\mu}\partial_{\nu}e^{a}_{\lambda} - \partial_{\nu}\partial_{\mu}e^{a}_{\lambda}). \quad (12.55) \]

Since \( e^{a}_{\mu} \) in (12.49) is single-valued, derivatives in front of it commute. Hence \( R_{\mu\nu\lambda\kappa} \) vanishes identically,
\[ R_{\mu\nu\lambda\kappa} \equiv 0. \quad (12.56) \]

A pure dislocation gives rise to torsion but not to curvature.

### 12.2.2 Disclination

As a second example for a multivalued mapping, we generate curvature by the transformation

\[ x^{i} = \delta^{i}_{\mu}[x^{\mu} + \Omega\epsilon^{\mu}_{\nu}x^{\nu}\phi(x)], \quad (12.57) \]

with the multi-valued function (12.46). The symbol \( \epsilon_{\mu\nu} \) denotes the antisymmetric Levi-Civita tensor. The transformed metric
\[ g_{\mu\nu} = \delta_{\mu\nu} - \frac{2\Omega}{x^{2}x^{\nu}}\epsilon_{\mu\nu}\epsilon^{\lambda\kappa}x^{\lambda}x^{\kappa}. \quad (12.58) \]

is single-valued and has commuting derivatives. The torsion tensor vanishes since \( (\partial_{1}\partial_{2} - \partial_{2}\partial_{1})x^{1,2} \) is proportional to \( x^{2,1}\delta^{(2)}(x) = 0 \). The local rotation field \( \omega(x) \equiv \frac{1}{2}(\partial_{1}x^{2} - \partial_{2}x^{1}) \), on the other hand, is equal to the multi-valued function \( -\Omega\phi(x) \), thus having the noncommuting derivatives:
\[ (\partial_{1}\partial_{2} - \partial_{2}\partial_{1})\omega(x) = -2\pi\Omega\delta^{(2)}(x). \quad (12.59) \]
To lowest order in $\Omega$, this determines the curvature tensor, which in two dimensions possesses only one independent component, for instance $R_{1212}$. Using the fact that $g_{\mu\nu}$ has commuting derivatives, $R_{1212}$ can be written as:

$$R_{1212} = (\partial_1 \partial_2 - \partial_2 \partial_1)\omega(x). \quad (12.60)$$

In defect physics, the mapping (12.57) is associated with a disclination which corresponds to an entire section of angle $\alpha$ missing in an ideal atomic array (see Fig. 10.2).  

### 12.3 Differential Geometric Properties of Affine Spaces

Up to now we have studied only such affine spaces which were obtained from a Minkowski spacetime by introducing an infinitesimal amount of defects. In reality, defects can pile up and the spacetime must be described by the full non-linear formulation of affine spaces. At the linear level we have learned how dislocations and disclinations manifest themselves by certain non-vanishing contour integrals around Burgers circuits. In this section we would like to discuss these geometric aspects at the non-linear level.

#### 12.3.1 Integrability of Metric and Affine Connection

The general affine spacetime will be characterized by the same type of integrability conditions as the spacetime with infinitesimal defects stated in Eqs. (12.9) and (12.10). In the nonlinear formulation, these conditions are imposed upon metric and affine connection:

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) g_{\lambda\kappa} = 0, \quad (12.61)$$

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) \Gamma^\kappa_{\sigma\lambda} = 0. \quad (12.62)$$

Remember that the first condition ensures the antisymmetry of the curvature tensor in the last two indices [see (11.134)]. By antisymmetrizing the second condition in $\sigma\lambda$ it can also be replaced by an integrability condition for the torsion

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) S^{\kappa}_{\sigma\lambda} = 0. \quad (12.63)$$

Moreover, using the decomposition (11.119), the Christoffel symbol is seen to be integrable as well:

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) \left\{ \kappa \atop \sigma\lambda \right\} = 0. \quad (12.64)$$

---

2Ibid., Eq. 2.86 on p. 1359.

3Ibid., Fig. 2.2 on p. 1366.
Since $\partial_\mu g_{\lambda \kappa}$ can be expressed in terms of products of Christoffel symbols and metric tensors, and since products of integrable functions are integrable,\(^4\) the derivatives of $g_{\lambda \kappa}$ satisfy the nonlinear version of (12.25):

\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \partial_\sigma g_{\lambda \kappa} = 0. \tag{12.65}
\]

Conversely, with the Christoffel symbol consisting of products of $g_{\lambda \kappa}$ and $\partial_\mu g_{\lambda \kappa}$, this condition implies (12.64) and thus is completely equivalent to it due to the necessary validity of (12.61).

### 12.3.2 Local Parallelism

In order to understand the geometric properties of such a general affine spacetime let us first introduce the concept of local parallelism.

Consider a vector field $v(x) = e_a v^a(x)$ which is parallel in the inertial frame in the naive sense that all vectors point in the same direction. This simply means $\partial_b v(x) = e_a \partial_b v^a = 0$. But when changing to the coordinates $x^\mu$ we find

\[
\partial_b v^a = \partial_b e^a_\mu v^\mu = e_b^\nu \partial_\nu \left( e^a_\mu v^\mu \right) = e_b^\nu e^a_\mu D^\nu v^\mu = 0. \tag{12.66}
\]

Thus parallel vector fields have their local components $v^\mu$ change in such a way that their covariant derivatives vanish:

\[
D^\nu v^\mu = \partial^\nu v^\mu + \Gamma^\nu_\lambda^\mu v^\lambda = 0. \tag{12.67}
\]

Similarly we find:

\[
D^\nu v_\mu = \partial^\nu v_\mu - \Gamma^\nu_\mu^\lambda v_\lambda = 0. \tag{12.68}
\]

Note that the basis tetrads $e^\nu_\alpha$, $e_\nu^a$ are parallel vector fields, by construction [see (11.94)].

Let us study this type of situation in general: Given an arbitrary connection $\Gamma^\lambda_\mu^\nu$ we first ask the question under what condition it is possible to find a parallel vector field in the whole space. For this we consider the vector field $v^\mu(x)$ at a point $x_0$ where it has the value $v^\mu(x_0)$. Let us now move to the neighboring position $x_0 + dx$. There the field has components

\[
v^\mu \left( x_0 + dx \right) = v^\mu(x_0) + \partial_\nu v^\mu(x_0) dx^\nu. \tag{12.69}
\]

If $v^\mu(x)$ is a parallel vector field with $D^\nu v^\mu = 0$ the derivative satisfies

\[
\partial_\nu v^\mu = -\Gamma^\nu_\mu^\kappa v^\kappa. \tag{12.70}
\]

\(^4\)This follows from the chain rule of differentiation

\[
(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu) \left( fg \right) = \left[ \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) f \right] g + f \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) g.
\]
This differential equation is integrable over a finite region of spacetime if and only if Schwarz’s criterion is fulfilled which says

\[ (\partial_\lambda \partial_\nu - \partial_\nu \partial_\lambda) v^\mu = 0 \]  

(12.71)

terms of second order in \( dx \).

If we calculate

\[ (\partial_\lambda \partial_\nu - \partial_\nu \partial_\lambda) v^\mu = -\partial_\lambda (\Gamma^\mu_{\nu\kappa} v^\kappa) + \partial_\nu (\Gamma^\mu_{\lambda\kappa} v^\kappa) \]  

(12.72)

we find

\[ - (\partial_\lambda \Gamma^\mu_{\nu\kappa} - \partial_\nu \Gamma^\mu_{\lambda\kappa}) v^\kappa - \Gamma^\mu_{\nu\kappa} \partial_\lambda v^\kappa + \Gamma^\mu_{\lambda\kappa} \partial_\nu v^\kappa \]  

(12.73)

and thus, using once more (12.70),

\[ (\partial_\lambda \partial_\nu - \partial_\nu \partial_\lambda) v^\mu = -R^\mu_{\lambda\nu\kappa} v^\kappa. \]  

(12.74)

Thus the parallel field \( v^\mu(x) \) exists in the whole spacetime if and only if the curvature tensor vanishes everywhere.

If \( R^\mu_{\lambda\nu\kappa} \) is non-zero, the concept of parallel vectors cannot be carried over from Minkowski space to the general affine spacetime over any finite distance. Such spaces are called curved. One says that in curved spaces there exists no teleparallelism.

We have illustrated before, that this is the case in the presence of disclinations. Disclinations generate curvature, i.e., a crystal containing disclinations is curved in the differential geometric sense.

This is in accordance with the previous observation that the disclination density \( \theta_{ij} \) coincides with the Einstein curvature tensor \( G_{ij} \).

In the illustration we have also seen that even in the presence of a disclination it still is meaningful to define a vector field as locally parallel. The condition for this is that the covariant derivatives vanish at that point \( x_0 : D_\nu v^\mu(x_0) = 0 \). If this condition is satisfied, the neighboring vector \( v^\mu(x) \), close to \( x_0 \) differs from \( v^\mu(x_0) \) at most by terms of the order \( (x - x_0)^2 \) rather than \( (x - x_0) \) for non-parallel vectors. In order to see this in more detail let us draw an infinitesimal quadrangle \( ABCD \) in the coordinate frame \( x^\mu \) spanned by \( AB = dx^\mu = DC \) and \( BC = dx^\nu = AD \) (see Fig. 12.3). Now we compare the directions of \( v^\mu(x) \) before and after going around the circumference. When passing from \( A \) at \( x^\mu \) to \( B \) at \( x^\mu + dx^\nu \) the vector components change from \( v^\mu_{B} = v^\mu(x) \) to

\[ v^\mu_{B} = v^\mu \left( x^\mu + x^\nu \right) = v^\mu_1 + \partial_\nu v^\mu dx^\nu v_{A}^\mu - \frac{\Gamma^A_{\nu\lambda}}{} v^\lambda v_{1}^\mu d x^\nu. \]  

(12.75)

When continuing to \( C \) at \( x^\mu + dx^\rho + dx^\tau \) we have

\[ v^\mu_{C} = v^\mu_{B} - \frac{\Gamma^B_{\tau\kappa}}{} v^\kappa v_{B}^{\rho} dx^\tau \]
Figure 12.3 Illustration of parallel transport of a vector around a closed circuit ABCD.

\[
\begin{align*}
&= v_A^\mu - A_{\nu\lambda}^\mu v^\nu_1 d x_1^\nu - B_{\tau\kappa}^\mu v^\kappa_2 d x_2^\tau + B_{\tau\kappa}^A v^A_1 d x_1^\tau d x_2^\tau \\
&= v_A^\mu - A_{\nu\lambda}^\mu v^\nu_1 (d x_1^\nu + d x_2^\nu) - \partial_\nu A_{\tau\kappa}^\mu v^\kappa_1 d x_1^\nu d x_2^\tau + \\
&\quad + A_{\tau\kappa}^A v^A_1 d x_1^\nu d x_2^\tau + O \left( d x^3 \right). \\
&= v_A^\mu - A_{\nu\lambda}^\mu v^\nu_1 (d x_1^\nu + d x_2^\nu) - \partial_\nu A_{\tau\kappa}^\mu v^\kappa_1 d x_1^\nu d x_2^\tau + O \left( d x^3 \right). \\
&= v_A^\mu - A_{\nu\lambda}^\mu v^\nu_1 (d x_1^\nu + d x_2^\nu) - \partial_\kappa A_{\tau\kappa}^\mu v^\lambda_1 d x_1^\nu d x_2^\tau + O \left( d x^3 \right). \\
&= v_A^\mu - A_{\nu\lambda}^\mu v^\nu_1 (d x_1^\nu + d x_2^\nu) - \partial_\kappa A_{\tau\kappa}^\mu v^\lambda_1 d x_1^\nu d x_2^\tau + O \left( d x^3 \right).
\end{align*}
\]

(12.76)

We can now repeat the same procedure along the line ADC, and we find the same result with interchanged \( d x_1 \) and \( d x_2 \). The difference between the two results is

\[
\begin{align*}
&= v_A^\mu - v_A^\mu = - \frac{1}{2} R_{\nu\tau\kappa}^\mu v^\kappa_1 d s^\nu_1 \tau + O \left( d x^3 \right)
\end{align*}
\]

(12.77)

where \( d s^\nu_1 \tau = \left( d x_1^\nu d x_1^\tau - d x_2^\nu d x_2^\tau \right) \) is the infinitesimal surface element of the quadrangle.

There exists a similar geometric illustration of the torsion property \( S_{\nu\mu}^\lambda \neq 0 \). Consider a crystal with an edge dislocation (see Fig. 12.2). Let us focus attention upon a closed circuit with the form of a parallelogram in the ideal reference crystal (i.e., in the coordinate frame \( e^a \)) and suppose its image in the \( e^\mu \)-frame encloses the dislocation line (see Fig. 12.4).

Volterra process of constructing the dislocation, the reference crystal was cut open, and a layer of atoms was inserted. In this process, the original parallelogram is opened such that the dislocation crystal has a gap between the open ends. The gap vector is precisely the Burgers vector. To be specific, let the parallelogram in the ideal reference crystal be spanned by the vectors \( AB = d^1_a, AD = d^2_a, D^a_C = d^a_1, BC = d^a_2 \). Since \( d^a_1, d^a_2 \) are parallel in the ideal reference crystal, they are parallel.
vectors, i.e., the vectors \( v^\mu(x) = d^\mu \frac{x^\mu}{2}, v^\mu \left(x^\mu + d^\mu \right) \) satisfy (12.70) when going from A to B, i.e., \( \partial_\nu d^\mu \frac{x^\mu}{2} = -\Gamma^\mu_{\nu\lambda} d^\lambda \frac{x^\lambda}{2} \) and hence

\[
dx^\mu_2 = d \frac{x^\mu}{2} - \Gamma^\mu_{\nu\lambda} d \frac{x^\nu}{2} d \frac{x^\lambda}{2}.
\]

(12.78)

Similarly the vectors \( d^\mu \frac{x^\mu}{1} \) and \( d^\mu \frac{x^\mu}{1} \) are parallel and therefore related by

\[
d^\mu \frac{x^\mu}{1} = d \frac{x^\mu}{1} - \Gamma^\mu_{\nu\lambda} d \frac{x^\nu}{2} d \frac{x^\lambda}{1}.
\]

(12.79)

From this it follows that

\[
b^\mu = \left( \frac{d^\mu x^\mu}{2} + \frac{d^\mu x^\mu}{1} \right) = \left( \frac{d^\mu x^\mu}{2} + \frac{d^\mu x^\mu}{2} \right) = -S^\mu_{\nu\lambda} d s^\nu \lambda.
\]

(12.80)

In a Minkowski space, the torsion vanishes and the image is again a closed parallelogram. Einstein assumed the vanishing of torsion in gravitational spacetime.

### 12.4 Circuit Integrals in Affine Spaces with Curvature and Torsion

In order to establish contact with the circuit definitions of disclinations and dislocations in crystals, let us phrase the differential results (12.77) and (12.80) in terms of contour integrals.

#### 12.4.1 Closed Contour Integral over Parallel Vector Field

Given a vector field \( v^\mu(x) \) which is locally parallel, i.e., which has \( D^\nu v^\mu(x) = 0 \), consider the change of \( v^\mu(x) \) while going around a closed contour which is

\[
\Delta v^\mu = \oint dx^\mu(x) = \oint dx^\nu \partial_\nu v^\mu(x).
\]

(12.81)
By decomposing $C$ into a large set of infinitesimal surface elements we can apply (12.77) and find

$$
\Delta \nu^\mu = \oint_{C(x^\nu)} dx^\nu \partial_\nu \nu^\mu = -\frac{1}{2} \int_{S(x^\nu)} ds^\tau \tau^\nu \tau^\mu \nu^\kappa(x^\kappa). \quad (12.82)
$$

Note that the tetrad fields $e_a^\mu$ are locally parallel by definition such that they satisfy

$$
\Delta e_a^\mu = -\oint_{C(x^\nu)} dx^\nu \partial_\nu e_a^\mu = -\frac{1}{2} \int_{S(x^\nu)} ds^\tau \tau^\nu \tau^\mu e_a^\kappa(x). \quad (12.83)
$$

Actually, this relation follows directly from Stokes' theorem:

$$
\Delta e_a^\mu = \oint_{C(x^\nu)} dx^\nu \partial_\nu e_a^\mu = \oint_{C(x^\nu)} dx^\nu \partial_\nu e_a^\mu = -\frac{1}{2} \int_{S(x^\nu)} ds^\tau \tau^\nu \tau^\mu e_a^\kappa. \quad (12.84)
$$

For an infinitesimal circuit, we can remove the tetrad from the integral and have

$$
\Delta e_a^\mu \approx \left\{ -\frac{1}{2} \int_{S(x^\nu)} ds^\tau \tau^\nu \tau^\mu e_a^\kappa \right\} e_a^\kappa \equiv \omega^\mu_\kappa e_a^\kappa. \quad (12.85)
$$

The matrix $\omega^\mu_\kappa$ has the property that $\omega^\mu_\kappa = g^\mu_\lambda \omega^\lambda_\kappa$ is antisymmetric, due to the antisymmetry of $R^\epsilon_\nu_\kappa_\mu$ in $\kappa_\mu$. Hence $\omega^\mu_\kappa$ can be interpreted as the parameters of an infinitesimal local Lorentz transformation. In three dimensions, this is a local rotation in agreement with what we observed previously:

Curvature is a signal for disclinations and these are rotational defects.

### 12.4.2 Closed Contour Integral over Coordinates

Let us now give an integral characterization of torsion. For this we consider an arbitrary closed contour $C(x^a)$ in the inertial frame (which generalizes the parallelogram used in the previous discussion). In the defected spacetime this contour has an image $C'(x^a)$ which does not necessarily close. In order to find how much is missing we form the integral

$$
\Delta x^\mu = \oint_{C(x^a)} dx^\mu = \oint_{C(x^a)} dx^a \partial_a x^\mu = \oint_{C(x^a)} dx^a e_a^\mu(x^a). \quad (12.86)
$$

By Stokes’ theorem, this becomes

$$
\frac{1}{2} \int_{C(x^a)} ds^{ab} (\partial_a e_b^\mu - \partial_b e_a^\mu) = \frac{1}{2} \oint_{C(x^a)} ds^{ab} (e_a^\nu \partial_\nu e_b^\mu - (a \leftrightarrow b)) = -\oint ds^{ab} S_{ab}^\mu. \quad (12.87)
$$

The quantity

$$
S_{ab}^\mu = -\frac{1}{2} e_a^\nu \left[ \partial_a e_b^\mu - (a \leftrightarrow b) \right] \quad (12.88)
$$

H. Kleinert, GRAVITY WITH TORSION
is called *anholonomity* of the mapping $x^a \rightarrow x^\mu$. It is related to the torsion $S^\mu_{\lambda\kappa}$ conversion of the lower indices from the local to the inertial form

\[
S^\mu_{ab} = e^\lambda_a e^\kappa_b S^\mu_{\lambda\kappa}
\]

\[
= -\frac{1}{2} \left\{ e^\lambda_a e^\kappa_b \left[ e^\epsilon_{\kappa} \partial_{\lambda} e^\epsilon_{\mu} - (a \leftrightarrow b) \right] \right\}
\]

\[
\equiv -\frac{1}{2} \left[ e^\lambda_a \partial_{\lambda} e^\epsilon_{\mu} - (a \leftrightarrow b) \right].
\] (12.89)

If the tetrad vectors are known as functions of the external coordinates $x^a$, we can also use $e^\lambda_a \partial_{\lambda} = \partial_a$ and write the anholonomity in the form

\[
S^\mu_{ab} \equiv -\frac{1}{2} \left[ \partial_a e^\epsilon_{\mu} - (a \leftrightarrow b) \right].
\] (12.90)

Sometimes one also converts the upper Einstein index $\mu$ into a local Lorentz index $c$ and works with

\[
S^c_{ab} \equiv e^c_{\mu} S^\mu_{ab} = -\frac{1}{2} \left[ e^c_{\mu} \partial_a e^\epsilon_{\mu} - (a \leftrightarrow b) \right].
\] (12.91)

If there is no torsion, the integral (12.87) vanishes. Otherwise the image of the closed contour $C(x^a)$ has a gap and thus is defined as the Burgers vector

\[
b^\mu = \int_{C'(x^\nu)} dx^\mu = -\oint_{C(x^a)} ds^\nu S^\nu_{ab} S^\mu_{ab}.
\] (12.92)

### 12.4.3 Closure Failure and Burgers Vector

It should be mentioned that the circuit integrals measuring curvature and torsion may be executed in the opposite way by forming closed circuits $C'(x^\nu)$ around the defect in the space $x^\mu$ and studying the properties of the image circuit $C'(x^a)$ in the ideal reference crystal. The torsion measures how much the image $C'(x^a)$ fails to close. The *closure failure* is given by the Burgers vector

\[
b^a = \int_{C'(x^\nu)} dx^a = \int_{c'(x^\mu)} dx^\mu \frac{\partial x^a}{\partial x^\mu} = \int_{c(x^a)} dx^\mu e^a_{\mu}
\] (12.93)

which, by Stokes’ theorem can be rewritten as

\[
b^a \int_{S(x^\nu)} ds^\nu \partial_a e^\epsilon_{\mu} = \int_{S(x^\nu)} ds^\nu S^\lambda_{\nu\mu} e^\lambda_{\mu}(x).
\] (12.94)

The tensor $S^a_{\mu} = S^\lambda_{\nu\mu} e^a_{\lambda} = 1/2 \left( \partial_\mu e^a_{\nu} - \partial_\nu e^a_{\mu} \right)$ is obviously a converse form of the anholonomity (12.92), with Einstein indices exchanged by local Lorentz indices.
12.4.4 Alternative Circuit Integral for Curvature

There is an analogous circuit integral characterizing the curvature from the standpoint of the coordinates $x^a$. For this we introduce the local Lorentz tensor related to $R_{\mu\nu\rho\kappa}$:

$$R_{abc}^d \equiv e_a^\mu e_b^\nu e_c^\lambda e_\kappa R_{\mu\nu\lambda\kappa}$$  \hspace{1cm} (12.95)

Then the circuit integral reads

$$\Delta e_a^\mu = -\frac{1}{2} \int_{S(x^\nu)} ds^d R_{\kappa\rho} e_b e_{\kappa} e_a^\mu.$$  \hspace{1cm} (12.96)

If one wants to calculate $R_{abc}^d$ directly in $x^a$ spacetime using differentiations one has to keep in mind that under the anholonomic mapping $x^a \rightarrow x^\mu R$, is not a tensor.

In fact, a simple manipulation shows

$$R_{\mu\nu\lambda\kappa} = e_a^\kappa \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) e^d_\lambda$$  \hspace{1cm} (12.97)

where

$$\tilde{R}_{ab\lambda} = R_{\mu\nu\lambda\kappa} e_a^\mu \partial_\nu - e_b^\mu \partial_\mu \partial_\lambda$$  \hspace{1cm} (12.98)

is evaluated in the same way as $R_{\mu\nu\lambda\kappa}$ in Eq. (11.129), but by forming $\partial_\mu$ derivatives rather than $\partial_\mu$. Expressing also the torsion $S_{\mu\nu}^\sigma$ in terms of derivatives $\partial/\partial x^a = \partial_a$ as in (12.90) we can write

$$S_{\mu\nu}^\sigma = e_a^\mu e_b^\nu e_c^\sigma S_{abc}.$$  \hspace{1cm} (12.99)

For the affine connection we may define, similarly,

$$\Gamma_{\mu\nu}^\sigma \equiv e_a^\mu e_b^\nu e_c^\sigma \Gamma_{abc}.$$  \hspace{1cm} (12.100)

with

$$\Gamma_{ab}^c = e_a^\mu e_b^\nu e_c^\lambda \Gamma_{\mu\nu\lambda} = -e_a^\mu e_b^\nu e_c^\lambda \partial_\mu e_\lambda = -e_a^\mu e_c^\lambda \partial_\mu e_\lambda$$  \hspace{1cm} (12.101)

Then $R_{abc}^d$ of (12.95) can be written as

$$R_{abc}^d = \tilde{R}_{abd}^c + 2 S_{ab}^e \Gamma_{ec}^d.$$  \hspace{1cm} (12.102)
12.4.5 Parallelism in World Crystal

From the standpoint of a world crystal with defects, parallelism has a simple meaning. Consider Fig. 11.1b. We identify the dashed curves $x^a(x^\mu) = \text{const.}$ with the crystal planes of an elastically distorted crystal as seen from the local frame with coordinates $x^\mu$. An observer living on the distorted crystal orients himself by the planes $x^a(x^\mu) = \text{const.}$ He measures distances and directions by counting atoms along the crystal directions. The above definition of parallelism amounts to vectors being defined as parallel if they are so from his point of view, i.e., if they correspond to parallel vectors in the undistorted crystal. Thus the normal vectors to the dashed coordinate planes $x^a(x^\mu) = \text{const.}$ are parallel to each other. Indeed, they form the vector fields $e^\mu_a(x)$, which always satisfy $D_\nu e^\mu_a = 0$ [see (11.94)].

If the mapping $x^a(x^\mu)$ contains defects it is, in general, impossible to find a global definition of parallelism. Consider, for example, a wedge disclination which is shown in Fig. 12.2, say the $-90^\circ$ one. The crystal has been cut from the left, and new crystalline material has been inserted in the Volterra construction process. The crystalline coordinate planes define parallel lines. With the right-hand piece stemming from the original crystal, there exists a completely consistent definition of parallelism. For example, the almost horizontal planes are all parallel. The lines cutting these vertically are also parallel by definition. On the left-hand side, the vertical lines continue smoothly into the inserted new crystalline material. In the middle, however, they meet and turn suddenly out to be orthogonal. Still, the coordinate planes define parallelism in any small region inside the original as well as the inserted material except on the disclination line.

12.5 Bianchi Identities for Curvature and Torsion Tensors

Because of their physical importance, let us derive a few important properties of curvature and torsion tensors. As noted before, the curvature tensor is antisymmetric in $\mu \nu$, by construction, and in $\lambda \kappa$, due to the integrability condition (11.19) for the metric tensor. In addition, it satisfies the so-called fundamental identity. This follows directly from the representation (11.128) by adding terms in which $\mu \nu \lambda$ are interchanged cyclically:

$$R_{\mu\nu\lambda}^\kappa = 2D_\nu S_{\mu\lambda}^\kappa - 4S_{\mu\rho}^\nu S_{\rho\lambda}^\kappa$$

where the symbol $\longrightarrow$ denotes a sum of cyclic permutations of the indicated subscripts. The derivation of the fundamental identity requires commuting derivatives in front of the metric tensor, i.e., it requires that the metric $g_{\mu\nu}$ satisfies the integrability condition

$$(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu)g_{\lambda\kappa} = 0.$$  

The fundamental identity is therefore a Bianchi identity [recall the definition given after Eq. (2.88)].
In symmetric spaces where \( S_{\mu\nu\lambda} = 0 \) and \( R_{\mu\nu\lambda\kappa} = \bar{R}_{\mu\nu\lambda\kappa} \), the fundamental identity implies the additional symmetry property of the Riemann tensor
\[
\bar{R}_{\mu\nu\lambda\kappa} + \bar{R}_{\nu\lambda\mu\kappa} + \bar{R}_{\lambda\mu\nu\kappa} = 0.
\]
(12.105)

Using the antisymmetry in \( \mu\nu \) and \( \lambda\kappa \) leads once more to the property (11.146):
\[
\bar{R}_{\mu\nu\lambda\kappa} = \bar{R}_{\lambda\kappa\mu\nu}.
\]
(12.106)

Another important identity is the original Bianchi identity which has given the name to all similar identities in this book which are based on the intergrability condition of observable fields. The original Bianchi identity follows from the assumption of the single-valuedness of the affine connection which implies that it satisfies the integrability condition
\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \Gamma_{\lambda\kappa}^\rho = 0.
\]
(12.107)

Consider the vector
\[
R_{\sigma\nu\mu} \equiv (\partial_\sigma \partial_\nu - \partial_\nu \partial_\sigma) e_\mu,
\]
(12.108)
which determines the curvature tensor \( R_{\sigma\nu\mu}^\lambda \) via the scalar product with \( e^\lambda \) [recall (11.129)]. Applying the covariant derivative gives
\[
D_\tau R_{\sigma\nu\mu} = \partial_\tau R_{\sigma\nu\mu} - \Gamma_{\tau\sigma}^\kappa R_{\kappa\nu\mu} - \Gamma_{\tau\nu}^\kappa R_{\sigma\kappa\mu}.
\]
(12.109)
Performing cyclic sums over \( \tau\sigma\nu \) and using the antisymmetry of \( R_{\sigma\nu\mu} \) in \( \sigma\nu \) leads to
\[
D_\tau R_{\sigma\nu\mu} = \partial_\tau R_{\sigma\nu\mu} - \Gamma_{\tau\mu}^\kappa R_{\sigma\nu\kappa} + 2 S_{\tau\sigma} R_{\nu\kappa\mu}.
\]
(12.110)
Now we use
\[
\partial_\sigma \partial_\nu e_\mu = \partial_\sigma \left( \Gamma_{\nu\mu}^\alpha e_\alpha \right) = \Gamma_{\nu\mu}^\kappa e_\kappa,
\]
(12.111)
to derive
\[
\partial_\tau \partial_\sigma \partial_\mu e_\nu = \partial_\tau \Gamma_{\nu\mu}^\kappa \partial_\sigma e_\kappa + (\tau\sigma) + \partial_\tau \partial_\sigma \Gamma_{\nu\mu}^\kappa e_\alpha + \Gamma_{\nu\mu}^\kappa \partial_\tau \partial_\sigma e_\kappa.
\]
(12.112)
Antisymmetrizing in \( \sigma\tau \) gives
\[
\partial_\tau \partial_\sigma \partial_\mu e_\nu - \partial_\sigma \partial_\tau \partial_\mu e_\nu = \Gamma_{\nu\mu}^\alpha R_{\tau\sigma\alpha} + \left[ (\partial_\tau \partial_\sigma - \partial_\sigma \partial_\tau) \Gamma_{\nu\mu}^\alpha \right] e_\alpha.
\]
(12.113)
This is the place where we make use of the integrability condition for the connection (12.107) to drop the last term, resulting in
\[
\partial_\tau R_{\sigma\nu\mu} - \Gamma_{\nu\mu}^\alpha R_{\tau\sigma\alpha} = 0
\]
(12.114)
Inserting this into (12.110) and multiplying by $e^n$ we obtain an expression involving the covariant derivative of the curvature tensor

$$D_\tau R_{\sigma\mu}^\kappa - 2S_\tau^\lambda R_{\nu\lambda\mu}^\kappa = 0.$$  \hspace{1cm} (12.115)

This is the Bianchi identity, which guarantees the integrability of the connection.

Within the defect interpretation of torsion and curvature, we are now prepared to demonstrate that these two identities have a simple physical interpretation. They are the non-linear versions of the conservation laws for dislocation and disclination densities. These read\(^5\)

$$\partial_i \alpha_{ij} = -\epsilon_{jkl} \theta_{kl},$$  \hspace{1cm} (12.116)

$$\partial_i \theta_{ij} = 0.$$  \hspace{1cm} (12.117)

They state that disclination lines never end while dislocation lines can end at most at a disclination line.

Consider now Eq. (12.115). Linearizing this gives

$$\partial_\tau R_{\sigma\mu}^\lambda + \partial_\sigma R_{\nu\tau\mu}^\lambda + \partial_\nu R_{\tau\mu\sigma}^\lambda = 0.$$  \hspace{1cm} (12.118)

Contracting $\nu$ and $\kappa$ and $\tau$ with $\lambda$ we obtain

$$2 \left( \partial_\nu S_{\mu\lambda}^\kappa + \partial_\mu S_{\lambda\nu}^\kappa + \partial_\lambda S_{\nu\mu}^\kappa \right) = R_{\nu\mu\lambda}^\kappa + R_{\mu\lambda\nu}^\kappa + R_{\lambda\nu\mu}^\kappa.$$  \hspace{1cm} (12.120)

where we have used the antisymmetry of $R_{\nu\mu\lambda\kappa}$ in the last two indices which is a consequence of the integrability condition for the metric tensor. The right-hand side is the same as $G_{\mu\lambda} - G_{\lambda\mu}$.

In three dimensions we can contract this equation with the $\epsilon$-tensor and find

$$\epsilon_{jkl} \left( \partial_i S_{kli} + \partial_k S_{lin} - \partial_l S_{kin} \right) = \epsilon_{ijkl} G_{kl}.$$  \hspace{1cm} (12.122)

Inserting here $S_{kij} = (1/2)\epsilon_{kli} \alpha_{ij}$ from (12.35), and Eq. (12.41) for $G_{ik}$, this becomes the conservation law (12.116) for the dislocation density.

\(^5\)See Eqs. (11.91) and (11.92) in Part III of the textbook [4].
12.6 Special Coordinates in Riemann Spacetime

12.6.1 Geodesic Coordinates

To a local observer, curved spacetime looks flat in his immediate neighborhood. After all, this is why men believed for a long time that the earth has the form of a flat disc. In four-dimensional spacetime the equivalent statement is that, in a freely falling elevator cabin, people would not experience any gravitational force as long as the cabin is small enough to make higher non-linear effects negligible. The cabin constitutes an inertial frame of reference for the motion of a mass point. From Eq. (11.22) we see that its coordinates in an arbitrary geometry can be determined from the requirement of a vanishing Christoffel symbol \( \{ \mu' \lambda', \kappa' \} = 0 \), which amounts to

\[
\begin{align*}
\partial_{\lambda'} g_{\mu' \lambda'}(x') &= 0, \quad \text{(12.123)} \\
\partial_{\lambda'} g^{\mu' \lambda'}(x') &= -g^{\mu' \sigma'} g^{\lambda' \nu'} \partial_{\lambda'} g_{\sigma' \nu'}(x') = 0. \quad \text{(12.124)}
\end{align*}
\]

Given an arbitrary set of coordinates \( x \), the derivatives are connected by

\[
\begin{align*}
\partial_{\lambda'} g^{\mu' \lambda'}(x') &= \partial_{\lambda} \left[ g^{\mu \nu}(x) \alpha_{\mu}^{\prime} \alpha_{\nu}^{\prime} \right] \alpha_{\lambda'}^{\lambda}, \\
&= \partial_{\lambda} g^{\mu \nu}(x) \alpha_{\mu}^{\prime} \alpha_{\nu}^{\prime} \alpha_{\lambda'}^{\lambda} + g^{\mu \nu} \partial_{\lambda} \alpha_{\mu}^{\prime} \alpha_{\nu}^{\prime} \alpha_{\lambda'}^{\lambda}.
\end{align*}
\]

Recall that derivative symbols \( \partial_{\mu} \) are meant to act only on the first function behind it. Equations (12.123) or (12.124) provide us with \( D^2(D + 1)/2 \) partial differential equations for the \( D \) coordinates \( x^{\mu'}(x) \) which do not, in general, have a solution over a finite region. If \( \partial_{\lambda'} g^{\mu' \lambda'} \) were to vanish over a finite region, the spacetime would necessarily be Euclidean. So we can, at best, achieve

\[
\partial_{\lambda'} g^{\mu' \nu'}(x'_0) = 0 \quad \text{(12.126)}
\]

at some point \( x_0 \). This implies, via (12.123), that also \( \partial_{\lambda} g_{\sigma' \nu'}(x'_0) = 0 \) and thus the vanishing of the Christoffel symbols at that point. Then a mass point will move force-free at \( x_0 \). In any neighborhood of \( x_0 \) there are gravitational forces of order \( O(x - x_0) \).

Let us try and solve (12.126) by an expansion

\[
\begin{align*}
x^{\mu'} &= x_0^{\mu} + a^{\mu}_{\lambda}(x - x_0)^{\lambda} + \frac{1}{2!} a^{\mu}_{\lambda \kappa}(x - x_0)^{\lambda}(x - x_0)^{\kappa} \\
&\quad + \frac{1}{3!} a^{\mu}_{\lambda \kappa \delta}(x - x_0)^{\lambda}(x - x_0)^{\kappa}(x - x_0)^{\delta} + \ldots. \quad \text{(12.127)}
\end{align*}
\]

The associated transformation matrix \( \alpha_{\mu}^{\prime \nu} \equiv \partial x^{\mu'}/\partial x^{\mu} \) satisfies

\[
\begin{align*}
\alpha_{\mu}^{\prime \nu} &= a_{\mu}^{\prime \nu} + a_{\mu \lambda}^{\prime \nu}(x - x_0)^{\lambda} + \frac{1}{2!} a_{\mu \lambda \kappa}^{\prime \nu}(x - x_0)^{\lambda}(x - x_0)^{\kappa} + \ldots, \\
\partial_{\lambda} \alpha_{\mu}^{\prime \nu} &= a_{\mu \lambda}^{\prime \nu} + a_{\mu \lambda \kappa}^{\prime \nu}(x - x_0)^{\kappa} + \ldots, \\
\partial_{\kappa} \partial_{\lambda} \alpha_{\mu}^{\prime \nu} &= a_{\mu \lambda \kappa}^{\prime \nu} + \ldots. \quad \text{(12.128)}
\end{align*}
\]

H. Kleinert, GRAVITY WITH TORSION
Inserting this into (12.125) we find
\[
\partial_\lambda g^{\mu\nu} = \partial_\lambda g^{\mu\nu}(x_0) a_\mu \nu' + g^{\sigma\tau}(x_0) \left[ a_\sigma \lambda \nu' a_\tau \mu' + (\mu' \leftrightarrow \nu') \right] = 0 + \mathcal{O}(x - x_0).
\] (12.129)

This is solved by
\[
a_\mu \mu' = g_\mu \mu'(x_0), \quad a_\lambda \kappa \mu = \frac{1}{2} \bar{\Gamma}_{\lambda \kappa} \mu(x_0),
\] (12.130)
in accordance with (11.98). Hence the coordinates which are locally geodesic at \(x_0\) are given by
\[
x'^\mu = x_0^\mu + (x - x_0)^\mu + \frac{1}{2} \bar{\Gamma}_{\lambda \kappa} \mu(x - x_0)^\lambda(x - x_0)^\kappa + \mathcal{O}\left((x - x_0)^3\right).
\] (12.131)

Note that while the Christoffel symbols vanish in the geodesic frame at \(x_0\), their derivatives are nonzero if the curvature is nonzero at \(x_0\).

In order to complete the construction of a freely falling coordinate system we just note that in the neighborhood of the point, the geodesic coordinates can always be brought to a Minkowski-form by a further linear transformation
\[
(x' - x_0)^\mu \rightarrow L_\alpha^\mu (x^\mu - x_0)^\alpha
\] (12.132)
which transforms \(g_{\mu\nu}\) into \(g_{\alpha\beta}\), i.e.,
\[
g_{\alpha\beta} = L_\alpha^\mu L_\beta^\nu g_{\mu\nu} = g_{\alpha\beta}.
\] (12.133)

Such a linear transformation does not change the geodesic property of the coordinates such that the coordinates \(x'' - x_0\) are a local inertial frame, which is what we wanted to find.

As far as the crystalline defects are concerned, the possibility of constructing geodesic coordinate is related to the fact that, in the regions between defects, the crystal can always be distorted elastically to form a regular array of atoms. In the continuum limit, these regions shrink to zero but so do the Burgers’ vectors of the defects. Therefore even though any small neighborhood does contain some defects, these themselves are infinitesimal such that the perfection of the crystal is disturbed only infinitesimally.

**12.6.2 Canonical Geodesic Coordinates**

The condition of being geodesic determined the coordinates transformation (12.127) up to the coefficients of the quadratic terms.
\[
x'^\mu = x_0^\mu + (x - x_0)^\mu + \frac{1}{2} \bar{\Gamma}_{\lambda \kappa} \mu(x - x_0)^\lambda(x - x_0)^\kappa
\]
\[
+ \frac{1}{3!} a_{\lambda \kappa \delta}^\mu (x - x_0)^\lambda(x - x_0)^\kappa(x - x_0)^\delta + \ldots
\] (12.134)
By construction, the transformation matrix
\[
\alpha^\mu_\nu = \frac{\partial \bar{x}^\mu}{\partial x^\nu} = \delta^\mu_\nu + \bar{\Gamma}^\mu_{\nu\lambda} (x - x_0)^\lambda + \frac{1}{2} a^\mu_{\lambda\nu} (x - x_0)^\lambda (x - x_0)^\kappa + \ldots \tag{12.135}
\]
has the property of making the Christoffel symbol of the point \(x_0\) vanish. It is obvious that the higher coefficients \(a^\mu_{\lambda\nu}\) must have an influence upon the derivatives of the Christoffel symbols. In general, these cannot be made zero since the curvature tensor at the point \(x_0\) where Gamma\(\mu'\nu'\lambda\) vanishes is
\[
R_{\mu'\nu'\lambda'}^{\kappa'} = \partial^\mu_\mu' \bar{\Gamma}_{\nu'\lambda'}^{\kappa'} - \left( \mu' \leftrightarrow \nu' \right). \tag{12.136}
\]
This implies that only in a flat spacetime can one find \(a^\mu_{\lambda\nu}\) to have also \(\partial^\mu_\mu' \bar{\Gamma}_{\nu'\lambda'}^{\kappa'} = 0\).

Even though the derivatives cannot be brought to zero, there is a most convenient coordinate system referred to as canonical, in which the derivatives satisfy the following relation
\[
\partial^\mu_\mu' \bar{\Gamma}_{\nu'\lambda'}^{\kappa'} + \partial^\lambda_\lambda' \bar{\Gamma}_{\nu'\mu'}^{\kappa'} = 0. \tag{12.137}
\]
Before we show how to find such a system, let us first see what its advantages are. The canonical condition allows us to invert the relation (12.136) for \(R_{\mu'\nu'\lambda'}^{\kappa'}\) and express (always at a geodesic coordinate point) the derivatives of the Christoffel symbols uniquely in terms of the curvature tensor
\[
\partial^\nu_\nu' \bar{\Gamma}_{\mu'\lambda'}^{\kappa'} = -\frac{1}{3} \left( R_{\mu'\nu'\lambda'}^{\kappa'} + R_{\lambda'\nu'\mu'}^{\kappa'} \right). \tag{12.138}
\]
This has the consequence, the metric \(g_{\mu'\nu'}(x')\) can in the neighborhood of the point \(x_0^\mu\), be expanded uniquely up to second order in terms of the curvature tensor. In order to see this we recall Eqs. (11.98). Differentiating these once more we find the second-order derivatives
\[
\partial^\kappa_\kappa' \partial^\lambda_\lambda' g_{\mu'\nu'} = \partial^\kappa_\kappa' \bar{\Gamma}_{\lambda'\mu'}^{\sigma} g_{\sigma\nu'} + \partial^\lambda_\lambda' \bar{\Gamma}_{\kappa'\nu'}^{\sigma} + \bar{\Gamma}_{\kappa'\lambda'}^{\sigma} \partial^\lambda_\kappa g_{\sigma\nu'} + \bar{\Gamma}_{\lambda'\nu'}^{\sigma} \partial^\kappa_\kappa' g_{\mu'\sigma}. \tag{12.139}
\]
At a point where the coordinates are geodesic, this becomes simply
\[
\partial^\kappa_\kappa' \partial^\lambda_\lambda' g_{\mu'\nu'} = -\frac{1}{3} \left( R_{\lambda'\kappa'\nu'}^{\mu'} + R_{\kappa'\nu'\lambda'}^{\mu'} \right) g_{\sigma'\nu'} - \frac{1}{3} \left( R_{\lambda'\kappa'\nu'}^{\mu'} + R_{\kappa'\nu'\lambda'}^{\mu'} \right) g_{\mu'\sigma'} = \frac{1}{3} \left( R_{\kappa'\nu'\lambda'}^{\mu'} + R_{\kappa'\nu'\lambda'}^{\mu'} \right). \tag{12.140}
\]
Hence the metric has the expansion
\[
g_{\mu'\nu'}(x') = g_{\mu'\nu'}(x_0) + \frac{1}{2} \partial^\kappa_\kappa' \partial^\lambda_\lambda' g_{\mu'\nu'}(x_0) (x' - x_0)^\kappa (x' - x_0)^\lambda + \ldots
\]
\[
= g_{\mu'\nu'}(x_0) + \frac{1}{3} R_{\kappa'\nu'\lambda'}^{\mu'} (x' - x_0)^\kappa (x' - x_0)^\lambda + \ldots
\]
\[
= g_{\mu'\nu'}(x_0) + \frac{1}{3} R_{\kappa'\nu'\lambda'}^{\mu'} (x' - x_0)^\kappa (x' - x_0)^\lambda + \ldots. \tag{12.141}
\]
Let us now turn to the construction of these canonical coordinates. For this we take the transformation law for the Christoffel symbols (11.102)
\[
\Gamma_{\mu'\nu'}^{\lambda'} = a_{\mu}^{\alpha} a_{\nu}^{\beta} \alpha^{\lambda'} \Gamma_{\mu
u}^{\lambda} - a_{\mu}^{\alpha} a_{\nu}^{\beta} \partial_{\lambda} \alpha_{\lambda}^{\nu'},
\]
and differentiate once more by
\[
\frac{\partial}{\partial x^\kappa} = \frac{\partial x^\kappa}{\partial x^\kappa} = a_{\kappa'}^{\kappa} \frac{\partial}{\partial x^\kappa}.
\]
This gives
\[
\frac{\partial_{\kappa'} \Gamma_{\mu'\nu'}^{\lambda'}}{\partial x^\kappa} = \kappa' \mu \bar{\Gamma}_{\lambda\kappa}^{\mu} + \kappa' \nu \bar{\Gamma}_{\lambda\kappa}^{\nu} \alpha_{\kappa}^{\lambda'} + \kappa' \nu \Gamma_{\mu\nu}^{\lambda} + \kappa' \mu \Gamma_{\mu\lambda}^{\kappa} - \kappa' \nu \partial_{\lambda} \alpha_{\lambda}^{\nu'} + \kappa' \mu \partial_{\mu} \alpha_{\kappa}^{\lambda'} + \kappa' \nu \alpha_{\kappa}^{\lambda'} (12.144)
\]

Besides the known transformation coefficient \( a_{\mu}^{\alpha} \nu \), this formula also involves the inverse coefficients \( a_{\kappa}^{\lambda} \). Since for \( x \sim x_0 \), \( a_{\kappa}^{\lambda} \) are close to unity, the inverse is simply (recall (11.67))
\[
a_{\mu}^{\nu} = \frac{\partial x^\nu}{\partial x^\mu} = \delta_{\mu}^{\nu} - \Gamma_{\lambda\nu}^{\mu} (x - x_0)^{\lambda} + \ldots (12.145)
\]

Indeed,
\[
a_{\mu}^{\nu} a_{\mu'}^{\nu'} = \delta_{\mu}^{\mu'} + \Gamma_{\lambda\nu}^{\mu} (x - x_0)^{\lambda} - \Gamma_{\lambda\nu}^{\mu} (x - x_0)^{\lambda} + \ldots = \delta_{\mu}^{\mu'} (12.146)
\]

Inserting \( a_{\mu}^{\nu} \) and \( a_{\nu}^{\nu} \) into the above transformation law gives
\[
\partial_{\kappa'} \Gamma_{\mu'\nu'}^{\lambda'} = [\partial_{\kappa} \Gamma_{\mu}\nu\lambda + \Gamma_{\mu\nu}^{\lambda} - a_{\kappa\mu\nu\lambda}]. \quad (12.147)
\]

Note the appearance of the coefficients \( a_{\kappa\mu\nu\lambda} \) of the cubic expansion terms. Interchanging on the left-hand side \( \kappa' \mu' \nu' \) cyclically and adding the three expansions gives
\[
\partial_{\kappa'} \Gamma_{\mu'\nu'}^{\lambda'} + 2 \text{ cyclic perms of } (\kappa' \mu' \nu') \quad (12.148)
\]

By setting the left-hand side equal to zero we obtain the desired equation for \( a_{\kappa\mu\nu\lambda} \).

Thus given an arbitrary coordinate frame \( x^\mu \), the coefficients \( a_{\kappa\mu\nu\lambda} \) can indeed be chosen such as to make the geodesic coordinate frame \( x'^\mu \) in the neighborhood of the point \( x_0 \) canonical, and thereby determining \( g_{\mu\nu}(x') \) in this neighborhood up to quadratic order uniquely in terms of the curvature tensor as stated in Eq. (12.141).

12.6.3 Harmonic Coordinates

While geodesic properties of coordinates can be enforced at most at one point there exists a way of fixing the choice of coordinates in the entire spacetime by choosing what are called harmonic coordinates. These were introduced first by T. DeDonder and C. Lanczos and extensively used by V. Fock in his gravitational work [2]. Given an arbitrary set of coordinates $x^\mu$, one asks for $d$ independent scalar functions $f^a(x)$ ($a = 0, 1, 2, d$) which satisfy the Laplace equation in curved space

$$D^2 f^a(x) = g^{\mu\nu} D_\mu D_\nu f^a(x) = 0. \quad (12.149)$$

Since $f^a(x)$ are supposed to be scalar functions, we calculate

$$D_\mu D_\nu f = D_\mu \partial_\nu f = \left( \partial_\mu \partial_\nu - \Gamma_{\mu\nu}^\lambda \partial_\lambda \right) f \quad (12.150)$$

and hence

$$D^2 f = g^{\mu\nu} D_\mu D_\nu f = \left( g^{\mu\nu} \partial_\mu \partial_\nu - \Gamma^\lambda_{\mu\nu} \partial_\lambda \right) f \quad (12.151)$$

where we have introduced the contracted affine connection

$$\Gamma^\lambda \equiv \Gamma_{\mu\lambda}^\mu, \quad (12.152)$$

for brevity. In a symmetric space

$$\Gamma^\lambda = \frac{1}{2} g^{\mu\nu} g^{\lambda\kappa} \left( \partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - \partial_\kappa g_{\mu\nu} \right)$$

$$= g^{\mu\nu} g^{\lambda\kappa} \partial_\nu g_{\mu\kappa} - \frac{1}{2} g^{\lambda\kappa} g^{\mu\nu} \partial_\kappa g_{\mu\nu}$$

$$= -\frac{1}{\sqrt{-g}} \partial_\kappa \left( \sqrt{-g} g^{\lambda\kappa} \right) \quad (12.153)$$

and

$$D^2 f = \left[ g^{\mu\nu} \partial_\mu \partial_\nu + \frac{1}{\sqrt{-g}} \partial_\kappa \left( \sqrt{-g} g^{\lambda\kappa} \right) \right] f = \Delta f, \quad (12.154)$$

where

$$\Delta \equiv \frac{1}{\sqrt{-g}} \left( \partial_\mu g^{\mu\nu} \sqrt{-g} \partial_\nu \right) \quad (12.155)$$

is the Laplace-Beltrami operator $\Delta$ in curved space. The Laplace operator in a spacetime with torsion is related to the Laplace-Beltrami operator by

$$D^2 f = \Delta f - S_{\mu\nu} \partial_\lambda f \quad (12.156)$$

Suppose we have found $D$ functions $f^a(x)$ which satisfy (12.149), then we introduce the harmonic coordinates $X^a$ as

$$X^a = f^a(x). \quad (12.157)$$

When transforming the Laplace equation (12.149) from coordinates $x^\mu$ to the harmonic coordinates $X^a$, we obtain

$$\left( g^{bc} \partial_b \partial_c - \Gamma^c_{\ bc} \partial_c \right) X^a = -\Gamma^c_{\ c} \delta^a_c = 0. \quad (12.158)$$

Thus, harmonic coordinates are characterized by vanishing $\Gamma^a$ ($a = 1, \ldots, d$).
12.6 Special Coordinates in Riemann Spacetime

12.6.4 Coordinates with $\det(g_{\mu\nu}) = 1$

A further choice of coordinates which was favored by Einstein, since it simplifies some formulas, is one in which the determinant of the metric is constant and has the Minkowski value $-1$ in all space. Since

$$\bar{\Gamma}_{\mu\nu}^{\lambda} = \frac{1}{2} g^{\mu\rho} \left( \partial_{\mu} g_{\nu\lambda} + \partial_{\nu} g_{\mu\lambda} - \partial_{\lambda} g_{\mu\nu} \right)$$

$$= \frac{1}{2} g^{\mu\rho} \partial_{\nu} g_{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_{\nu} \sqrt{-g} = \partial_{\nu} \log \sqrt{-g}$$

we can state this condition also in the form

$$\bar{\Gamma}_{\mu\nu}^{\lambda} = 0. \tag{12.160}$$

Given an arbitrary coordinate system $x^\nu$, the special ones $\bar{x}^\mu$ are found by a transformation

$$\bar{x}^\mu = \alpha^\mu_{\nu} x^\nu \tag{12.161}$$

which fulfills the condition

$$\sqrt{-g} = |\det (\alpha^\mu_{\nu})| \sqrt{-g} = |\det (\alpha^\mu_{\nu})|. \tag{12.162}$$

Taking the logarithm and differentiating gives

$$\bar{\Gamma}_{\mu\nu}^{\lambda} = \partial_{\nu} \log \det (\alpha^\lambda_{\kappa}) = \partial_{\nu} \text{tr} \log (\alpha^\lambda_{\kappa}) = \text{tr} \left( \alpha^{-1} \partial_{\nu} \alpha \right) = \alpha^\lambda_{\kappa} \partial_{\nu} \alpha^\kappa_{\lambda} = 0. \tag{12.163}$$

These are $D$ differential equations which determine $D$ new coordinate functions $\bar{x}(x)$.

Note the difference with respect to harmonic coordinates which have $\Gamma^\mu = \Gamma^\lambda_{\lambda\mu} = 0$, i.e., the first two indices contracted, on the general connection, while the present condition has the first (or the second index) contracted with the third, on the Christoffel symbol.

12.6.5 Orthogonal Coordinates

For many calculations, it is useful to employ orthogonal coordinates, in which $g_{\mu\nu}$ has only diagonal elements. Then many entries of the Christoffel symbols vanish

$$\bar{\Gamma}_{\mu\lambda,\kappa} = 0, \quad \bar{\Gamma}_{\mu\lambda}^{\kappa} = 0, \quad \mu \neq \lambda, \kappa \neq \mu, \kappa \neq \lambda, \tag{12.164}$$

and the calculation of the others is greatly simplified. In a symmetric space, we may use formula (12.1) for the Riemann tensor $\bar{R}_{\mu\nu\lambda}^{\kappa}$ and find that it vanishes whenever all its indices are different. The non-vanishing elements can be calculated as follows:
\[ R_{\nu\kappa\lambda\mu} = -\frac{1}{2} \left( \partial_\nu \partial_\kappa g_{\lambda\mu} + \partial_\mu \partial_\nu g_{\kappa\lambda} - \partial_\lambda \partial_\nu g_{\mu\kappa} \right) \]

\[ + \frac{1}{2} \partial_\mu \left( \log \sqrt{g_{\nu\nu}} \right) \partial_\nu g_{\kappa\lambda} - \frac{1}{2} \partial_\nu \left( \log \sqrt{g_{\mu\nu}} \right) \partial_\lambda g_{\mu\kappa} \]

\[ - \frac{1}{2} \partial_\lambda \left( \log \sqrt{g_{\nu\nu}} \right) \left( \partial_\nu g_{\mu\kappa} - \partial_\kappa g_{\mu\nu} \right) - \frac{1}{2} \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right) \left( \partial_\nu g_{\kappa\mu} - \partial_\mu g_{\kappa\nu} \right) + \frac{1}{2} \Gamma_{\kappa\lambda}^\rho \partial_\rho g_{\nu\mu}, \]

\[ \nu \neq \kappa, \lambda \neq \mu, \nu \neq \lambda \]

\[ R_{\nu\kappa\lambda\mu} = -\frac{1}{2} \left( \partial_\nu \partial_\kappa g_{\lambda\mu} - \partial_\lambda \partial_\nu g_{\mu\kappa} \right) \]

\[ + \frac{1}{2} \partial_\mu \left( \log \sqrt{g_{\nu\nu}} \right) \partial_\nu g_{\kappa\lambda} - \frac{1}{2} \partial_\lambda \left( \log \sqrt{g_{\mu\nu}} \right) \partial_\nu g_{\mu\kappa} \]

\[ - \frac{1}{2} \partial_\nu \left( \log \sqrt{g_{\mu\mu}} \right) \left( \partial_\nu g_{\mu\kappa} - \partial_\kappa g_{\mu\nu} \right) - \frac{1}{2} \partial_\nu \log \sqrt{g_{\lambda\lambda}} \left( \partial_\nu g_{\kappa\mu} - \partial_\mu g_{\kappa\nu} \right), \]

\[ \nu \neq \kappa, \lambda \neq \mu, \nu \neq \lambda, \nu \neq \mu, \kappa \neq \lambda \]

\[ R_{\nu\kappa\lambda\mu} = \frac{1}{2} \partial_\nu \partial_\kappa g_{\lambda\mu} - \frac{1}{2} \partial_\lambda \left( \log \sqrt{g_{\nu\nu}} \right) \partial_\nu g_{\mu\kappa} - \frac{1}{2} \partial_\lambda \left( \log \sqrt{g_{\mu\nu}} \right) \partial_\nu g_{\mu\kappa}. \]

The Ricci tensor becomes

\[ R_{\mu\nu} = \sum_\lambda \frac{1}{g_{\lambda\lambda}} \bar{R}^{\mu\lambda\lambda\nu}, \]

giving the off-diagonal elements

\[ \mu \neq \nu \]

\[ R_{\mu\nu} = \sum_{\lambda \neq \mu, \lambda \neq \nu} \left[ \partial_\mu \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right) - \partial_\mu \left( \log \sqrt{g_{\mu\nu}} \right) \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right) \right. \]

\[ \left. - \partial_\mu \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\nu \left( \log \sqrt{g_{\mu\mu}} \right) + \partial_\mu \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\mu \left( \log \sqrt{g_{\mu\mu}} \right) \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right) \right] \]

\[ = -\partial_\mu \partial_\nu \log \sqrt{-g} + \partial_\mu \left( \log \sqrt{g_{\mu\mu}} \right) \partial_\nu \left( \log \sqrt{-g} \right) + \partial_\mu \log \sqrt{-g} \partial_\nu \log \sqrt{g_{\mu\mu}} \]

\[ + \partial_\mu \partial_\nu \log \sqrt{g_{\mu\mu}} \partial_\mu \left( \log \sqrt{-g} \right) - 2 \partial_\mu \left( \log \sqrt{g_{\nu\nu}} \right) \partial_\nu \log \sqrt{g_{\mu\mu}} - \sum_{\lambda=1}^{D} \partial_\mu \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\nu \left( \log \sqrt{g_{\lambda\lambda}} \right), \]

and the diagonal elements

\[ R^{\mu\mu} = -\partial^2 \log \sqrt{-g} + 2 \partial^2_\mu \left( \log \sqrt{g_{\mu\mu}} \right) - 2 \left( \partial_\mu \log \sqrt{g_{\mu\mu}} \right)^2 \]

\[ + 2 \partial_\mu \left( \log \sqrt{-g} \right) \partial_\mu \left( \log \sqrt{g_{\mu\mu}} \right) - \sum_{\lambda=1}^{D} \left( \partial_\mu \log \sqrt{g_{\lambda\lambda}} \right)^2 \]

\[ - g_{\mu\mu} \sum_{\lambda=1}^{D} \frac{1}{g_{\lambda\lambda}} \left[ \partial^2_\lambda \left( \log \sqrt{g_{\mu\mu}} \right) + \partial_\lambda \left( \log \sqrt{-g} \right) \partial_\lambda \left( \log \sqrt{g_{\mu\mu}} \right) \right. \]

\[ \left. - 2 \partial_\lambda \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\lambda \left( \log \sqrt{g_{\lambda\lambda}} \right) \right], \]

H. Kleinert, GRAVITY WITH TORSION
The curvature scalar reads
\[
\bar{R} = \sum_{\lambda} \frac{1}{g_{\lambda\lambda}} \{ 2\partial_\lambda^2 \left( \log \sqrt{-g} \right) - 2\partial_\lambda^2 \log \sqrt{g_{\lambda\lambda}} \\
+ 2 \left( \partial_\lambda \log \sqrt{g_{\lambda\lambda}} \right)^2 - 4 \partial_\lambda \left( \log \sqrt{g_{\lambda\lambda}} \right) \partial_\lambda \left( \log \sqrt{-g} \right) \\
+ \left( \partial_\lambda \log \sqrt{-g} \right)^2 + \sum_{\kappa} \left( \partial_\lambda \sqrt{g_{\kappa\kappa}} \right)^2 \}.
\]
(12.171)

12.7 Number of Independent Components of \( R_{\mu\nu\lambda}^\kappa \) and \( S_{\mu\nu}^\lambda \)

With the antisymmetry in \( \mu\nu \) and \( \lambda\kappa \), there exist at first \( N_d^{\bar{R}} = \left[ d(d-1)/2 \right]^2 \) components of \( \bar{R}_{\mu\nu\lambda\kappa} \) and \( N_d^S = d^2(d-1)/2 \) components of \( S_{\mu\nu}^\lambda \) in \( d \) dimensions. Thus \( R_{\mu\nu\lambda\kappa} \) may be viewed as a \( \frac{1}{2}d(d-1) \times \frac{1}{2}d(d-1) \) matrix \( R_{(\mu\nu)(\lambda\kappa)} \) in the index pairs. In symmetric spaces, there is in addition symmetry of \( \bar{R}_{\mu\nu\lambda\kappa} \) between the index pairs \( \mu\nu \) and \( \lambda\kappa \), so that it has
\[
\frac{1}{2} \left\{ \frac{1}{2}d(d-1) \times \left[ \frac{1}{2}d(d-1) - 1 \right] \right\} = \frac{1}{8}d(d-1)(d^2 - d + 2)
\]
(12.172)

components. Now, the fundamental identity (12.105) not only leads to the symmetry, it contains also the information that the completely antisymmetric part of \( \bar{R}_{\mu\nu\lambda\kappa} + \bar{R}_{\nu\lambda\mu\kappa} + \bar{R}_{\lambda\nu\kappa\mu} \) vanishes. This gives \( d(d-1)(d-2)(d-3)/4! \) further relations, and one is left with
\[
N_d^{\bar{R}} = \frac{1}{8}d(d-1)(d^2 - d + 2) - \frac{1}{24}d(d-1)(d-2)(d-3) = \frac{1}{12}d^2(d^2 - 1)
\]
(12.173)

independent components of \( \bar{R}_{\mu\nu\lambda\kappa} \). In four dimensions, this number is 20, in three dimensions it is 6.

12.7.1 Two Dimensions

In two dimensions, there is only one independent component, for instance \( \bar{R}_{1221} \).

Indeed, the most general tensor with the above symmetry properties is
\[
\bar{R}_{\mu\nu\lambda\kappa} = -\text{const} \times e_{\mu\nu} e_{\lambda\kappa},
\]
(12.174)

where \( e_{\mu\nu} = \sqrt{-g} e_{\mu\nu} \) is the covariant version of the Levi-Civita symbol \( \epsilon_{01} = -\epsilon_{10} = 1 \) (compare (11A.7)). Contracting this with \( g^{\nu\lambda} \) gives the Ricci tensor
\[
\bar{R}_{\mu\kappa} = -\text{const} \times g^{\nu\lambda} e_{\mu\nu} e_{\lambda\kappa} = -\text{const} \times \left( g_{\mu\kappa} - g_{\lambda\lambda} g_{\mu\kappa} \right) \\
= \text{const} \times g_{\mu\kappa},
\]
(12.175)
and the scalar curvature
\[ R = \text{const} \times g^{\mu\kappa} g_{\mu\kappa} = 2 \times \text{const}. \]  
(12.176)
Hence \( \text{const} = \frac{R}{2} \) and the single independent element of \( \tilde{R}_{\mu\nu\lambda\kappa} \) is
\[ \tilde{R}_{0110} = \tilde{R}_{1001} = g \tilde{R} \]  
(12.177)
while the full curvature tensor is given by
\[ \tilde{R}_{\mu\nu\lambda\kappa} = -\epsilon_{\mu\nu\delta} \epsilon_{\lambda\kappa\tau} \left( \tilde{R}^{\tau\delta} - \frac{1}{2} g^{\tau\delta} \tilde{R} \right) \]  
(12.179)
where \( \epsilon_{\mu\nu\delta} = \sqrt{-g} \epsilon_{\mu\nu\delta} \) is the three-dimensional covariant version of the Levi-Civita symbol in Eq. (11A.7). The proof of Eq. (12.179) follows by contraction with \( \epsilon_{\mu\nu\delta} \epsilon_{\lambda\kappa\tau} \), which gives via the identity (1A.17) for products of two Levi-Civita tensors:
\[ \tilde{R}^{\tau\delta} - \frac{1}{2} g^{\tau\delta} \tilde{R} = \frac{1}{4} \epsilon_{\mu\nu\delta} \epsilon_{\lambda\kappa\tau} \tilde{R}_{\mu\nu\lambda\kappa}. \]  
(12.180)
This equation is trivially valid due to the identity (1A.18) and the definition (11.140) of the Ricci tensor [see also (12.28)].

Since \( \tilde{R}_{\mu\nu\lambda\kappa} \) is a tensor, its \( N_d^d = d^2(d^2 - 1)/12 \) components are different in different coordinates frame. It is useful to find out how many independent invariants one can form which do not depend on the frame. In two dimensions, the scalar curvature \( \tilde{R} \) determined \( \tilde{R}_{\mu\nu\lambda\kappa} \) completely. In general, the invariants of \( \tilde{R}_{\mu\nu\lambda\kappa} \) can all be constructed by suitable contractions with \( g^{\mu\nu} \). The tensors \( \tilde{R}_{\mu\nu\lambda\kappa} \) and \( g^{\mu\nu} \) together have \( d^2(d^2 - 1)/12 + d(d + 1)/2 \) matrix elements. There are \( d^2 \) arbitrary coordinate transformations matrices \( \partial x^{\nu}/\partial x^\lambda \) which can be applied to these tensors. The number of invariants is equal to the number of independent components of both \( \tilde{R}_{\mu\nu\lambda\kappa} \) and \( g_{\mu\nu} \) have in a specific coordinate system. This number is
\[ N_d^{\text{inv}} = \frac{1}{12} d^2(d^2 - 1) + \frac{d(d + 1)}{2} - N^2 = \frac{d}{12} d(d - 1)(d - 2)(d + 3). \]  
(12.181)
This formula is valid only for \( d \neq 2 \) since for \( d = 2 \) we have seen before that there is \( N_d^{\text{inv}} = 1 \) invariant, the scalar curvature. The above counting breaks down since one of the \( N^2 \) coordinate transformations subtracted in (12.181) happens to leave \( R_{1234} \) and \( g_{\mu\nu} \) invariant. For \( d = 3, 4 \) the numbers are \( N_d^{\text{inv}} = 3, 14 \) respectively.
12.7 Number of Independent Components of $R_{\mu\nu\lambda}^\kappa$ and $S_{\mu\nu}^\lambda$

12.7.2 Three Dimensions

In three dimensions, the invariants are the eigenvalue of the characteristic equation
\[
\det \left( g^{\mu\lambda} \bar{R}_{\lambda\kappa} - \lambda \delta^\mu_\kappa \right) = \det \left( g^{-1} \bar{R} - \lambda \right) = -\lambda^3 + \lambda^2 I_1 - \lambda I_2 + I_3, \tag{12.182}
\]
where
\[
I_1 = \text{tr} \left( g^{-1} \bar{R} \right) = g^{\mu\nu} \bar{R}_{\lambda\mu} = \bar{R},
\]
\[
I_2 = \frac{1}{2} \left( \bar{R}^{\mu}_\nu \bar{R}^{\nu}_\lambda - \bar{R}^{\lambda}_\lambda \bar{R}^{\kappa}_\kappa \right),
\]
\[
I_3 = \det \left( g^{-1} \bar{R} \right) = \det \left( g^{\mu\lambda} \bar{R}_{\lambda\nu} \right) = \frac{\det \left( \bar{R}_{\mu\nu} \right)}{\det \left( g_{\mu\nu} \right)}.
\]

12.7.3 Four or More Dimensions

In four or more dimensions, relation (12.179) generalizes to the Weyl decomposition of the curvature tensor
\[
\bar{R}_{\mu\nu\lambda\kappa} = -\frac{1}{d-2} \left( g_{\mu\lambda} \bar{R}_{\nu\kappa} - g_{\nu\lambda} \bar{R}_{\mu\kappa} + g_{\nu\kappa} \bar{R}_{\mu\lambda} - g_{\mu\kappa} \bar{R}_{\nu\lambda} \right) + \frac{\bar{R}}{(d-1)(d-2)} \left( g_{\mu\lambda} g_{\nu\kappa} - g_{\nu\lambda} g_{\mu\kappa} \right) + C_{\mu\nu\lambda\kappa}, \tag{12.184}
\]
where $C_{\mu\nu\lambda\kappa}$ is called the Weyl conformal tensor, which vanishes for $d = 3$, due to (12.179). Each of the three terms in this decomposition has the same symmetry properties as $\bar{R}_{\mu\nu\lambda\kappa}$. In addition, $C_{\mu\nu\lambda\kappa}$ satisfies the $d(d+1)/2$ conditions
\[
C^\kappa_\mu = C^{\nu\kappa}_{\mu\nu} = 0, \tag{12.185}
\]
since the Ricci tensor comes entirely from the first two terms. Hence, the Weyl tensor has
\[
N_d^C = \frac{1}{12} d^2 (d^2 - 1) - \frac{1}{2} d(d + 1) = \frac{1}{12} d(d + 1)(d + 2)(d - 3) \tag{12.186}
\]
independent elements which is the same as the number of invariants of $\bar{R}_{\mu\nu\lambda\kappa}$. In many cases, this makes it possible to find all invariants by going to a coordinate frame in which $g_{\mu\nu} = g_{\mu\nu}$ and $\bar{R}_{\mu\nu} = \text{diagonal}$, the first by going to a freely falling frame, the second by performing an appropriate additional Lorentz transformation. This procedure works as long as $\bar{R}_{\mu\nu}$ does not have any degenerate eigenvalues. Otherwise the Lorentz transformations remain independent and the counting does not work [1].

The above results have interesting consequences as far as a possible geometric theory of gravitation in lower-dimensional spaces is concerned. It turns out that a 3+1-dimensional spacetime is impossible to have a theory which reduces to Newton’s
theory in the weak coupling limit. As we shall see later in Chapter 15, the crucial geometric quantity in Einstein’s theory is the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R.$$  \hspace{1cm} (12.187)

Einstein’s theory postulates this tensor to be proportional to the symmetric energy-momentum tensor of matter, $T_{\mu\nu}$ [see Eq. (15.61)]

$$\bar{G}_{\mu\nu} = \kappa T_{\mu\nu},$$  \hspace{1cm} (12.188)

with some constant $\kappa$. In the above discussion we have learned that the Ricci tensor in two spacetime dimensions can be expressed in terms of the scalar curvature as

$$\bar{R}_{\mu\nu} = g_{\mu\nu} \frac{\bar{R}}{2}.$$  \hspace{1cm} (12.189)

But this implies that in two spacetime dimensions, the Einstein tensor $\bar{G}_{\mu\nu}$ vanishes identically. Hence also the energy-momentum tensor vanishes and there is no Einstein theory of gravity. At first sight, there seems to be an escape by allowing for the presence of a so-called cosmological term, in which case the Einstein equation reads,

$$\bar{G}_{\mu\nu} = \kappa T_{\mu\nu} + \Lambda g_{\mu\nu}.$$  \hspace{1cm} (12.190)

However, even if this is added, the two-dimensional theory has a severe problem: The metric $g_{\mu\nu}$ is determined completely by the local energy-momentum tensor

$$g_{\mu\nu} = -\frac{\kappa}{\Lambda} \frac{m}{T_{\mu\nu}},$$  \hspace{1cm} (12.191)

and hence vanishes in the empty spacetime between mass points. Thus also this version of gravity is unphysical.

How about a geometric theory of gravitation in $2+1$ dimensions? Here the Ricci tensor is independent of scalar such that there does exist a nontrivial Einstein tensor $\bar{G}_{\mu\nu}$. Still, the tensor is almost trivial. In three dimensions there is no Weyl tensor $C_{\mu\nu\lambda\kappa}$ and the full curvature tensor is determined in terms of the Ricci tensor by Eq. (12.184) for $d = 3$

$$\bar{R}_{\mu\nu\lambda\kappa} = -(g_{\mu\lambda} \bar{R}_{\nu\kappa} + g_{\nu\lambda} \bar{R}_{\mu\kappa} + g_{\nu\kappa} \bar{R}_{\mu\lambda} - g_{\mu\kappa} \bar{R}_{\nu\lambda})$$

$$+ \frac{\bar{R}}{2} \left( g_{\mu\lambda} g_{\nu\kappa} - g_{\nu\lambda} g_{\mu\kappa} \right).$$  \hspace{1cm} (12.192)

Inserting

$$\bar{R}_{\mu\nu} = \bar{G}_{\mu\nu} - \frac{g_{\mu\nu}}{d-2} \bar{G},$$  \hspace{1cm} (12.193)
this becomes
\[
\bar{R}^{\mu\nu\lambda\kappa} = \left(-\left(g_{\mu\lambda}\bar{G}_{\nu\kappa} - g_{\nu\lambda}\bar{G}_{\mu\kappa} + g_{\nu\kappa}\bar{G}_{\mu\lambda} - g_{\mu\kappa}\bar{G}_{\nu\lambda}\right) + \bar{G}\left(g_{\mu\lambda}g_{\nu\kappa} - g_{\nu\lambda}g_{\mu\kappa}\right)\right). \tag{12.194}
\]

According to Einstein’s equation we have
\[
\bar{G}_{\mu\nu} = \kappa T_{\mu\nu}, \tag{12.195}
\]
and see that \(\bar{R}_{\mu\nu\lambda\kappa}\) is completely determined by the local energy distribution. In the empty spacetime between masses there is no curvature. As we shall understand later, this implies physically that interstellar dust would experience no relative acceleration (tidal forces).

### 12.7.4 Constant Curvature

For a space with constant curvature all these equations simplify. Consider a sphere of radius \(r\) embedded in \(D\) dimensions has an intrinsic dimension \(D' \equiv D - 1\) and a curvature scalar
\[
\bar{R} = \frac{(D' - 1)D'}{r^2}. \tag{12.196}
\]
This is most easily derived as follows. Consider a line element in \(D\) dimensions
\[
(dx)^2 = (dx^1)^2 + (dx^2)^2 + \ldots + (dx^D)^2 \tag{12.197}
\]
and restrict the motion to a spherical surface
\[
(x^1)^2 + (x^2)^2 + \ldots + (x^D)^2 = r^2, \tag{12.198}
\]
by eliminating \(x^D\). This brings (12.197) to the form
\[
(dx)^2 = (dx^1)^2 + (dx^2)^2 + \ldots + (dx^{D'})^2 + \left(\frac{x^1dx^1 + dx^2 + \ldots + x^{D'}dx^{D'}}{r^2 - r'^2}\right)^2, \tag{12.199}
\]
where \(r'^2 \equiv (x^1)^2 + (x^2)^2 + \ldots + (x^{D'})^2\). The metric on the \(D'\)-dimensional surface is therefore
\[
g_{\mu\nu}(x) = \delta_{\mu\nu} + \frac{x^\mu x^\nu}{r^2 - r'^2}. \tag{12.200}
\]
Since \(\bar{R}\) will be constant on the spherical surface, we may evaluate it for small \(x^\mu\) \((\mu = 1, \ldots, D')\) where \(g_{\mu\nu}(x) \approx \delta_{\mu\nu} + x^\mu x^\nu/r^2\) and the Christoffel symbols (11.22) are \(\Gamma_{\mu\nu}^\lambda \approx \Gamma_{\mu\nu\lambda} \approx \delta_{\mu\lambda}x^\nu/r^2\). Inserting this into (11.128) we obtain the curvature tensor for small \(x^\mu\):
\[
\bar{R}_{\mu\nu\lambda\kappa} \approx \frac{1}{r^2} \left(\delta_{\mu\nu}\delta_{\lambda\kappa} - \delta_{\mu\lambda}\delta_{\nu\kappa}\right). \tag{12.201}
\]
This can be extended covariantly to the full surface of the sphere by replacing \( \delta_{\mu\lambda} \) by the metric \( g_{\mu\lambda}(x) \):

\[
\bar{R}_{\mu\nu\lambda\kappa}(x) = \frac{1}{r^2} \left[ g_{\mu\kappa}(x)g_{\nu\lambda}(x) - g_{\mu\lambda}(x)g_{\nu\kappa}(x) \right],
\]

(12.202)

so that Ricci tensor is [recall (11.140)]

\[
\bar{R}_{\nu\kappa}(x) = \bar{R}_{\mu\nu\kappa}^{\phantom{\mu\nu}\mu}(x) = \frac{D' - 1}{r^2} g_{\nu\kappa}(x).
\]

(12.203)

Contracting this with \( g^\nu\kappa \) [recall (11.141)] yields indeed the curvature scalar (12.196).

**Notes and References**


See also


[2] In general relativity there have been theories based on spaces in which this is not satisfied. Then the object \( Q_{\lambda\mu\nu} \equiv -D_{\lambda} g_{\mu\nu} \) becomes a dynamical field to be determined from field equations. See T. DeDonder, *La gravitation de Weyl-Eddington-Einstein*, Gauthier-Villars, Paris, 1924;

H. Weyl, Phys. Z. 22, 473 (1921); Ann. Phys. 59, 101 (1919); 65, 541 (1921);


In such spaces, the correction is defined as

\[
\Gamma_{\mu\nu}^{\lambda} \equiv e_{a}^{\lambda} \left( \partial_{\mu} - D_{\mu} \right) e_{a}^{\nu}
\]

and can be decomposed as

\[
\Gamma_{\mu\nu}^{\lambda} \equiv e_{a}^{\lambda} \left( \partial_{\mu} - D_{\mu} \right) e_{a}^{\nu} = \Gamma_{\mu\nu}^{\lambda} - \left( S_{\mu\nu}^{\lambda} - S_{\nu\mu}^{\lambda} + S_{\nu\mu}^{\lambda} \right) + \frac{1}{2} (Q_{\mu\nu}^{\lambda} - Q_{\nu\mu}^{\lambda} - Q_{\nu\mu}^{\lambda}),
\]

where \( S_{\mu\nu}^{\lambda} \equiv \frac{1}{4} (\Gamma_{\mu\nu}^{\lambda} - \Gamma_{\nu\mu}^{\lambda}) \) as defined in (11.113).

[3] More on the counting of independent components and invariants can be found in the textbook:


[4] The physics of defects is explained in the textbook

[5] For the introduction of harmonic coordinates see:


This work was criticized on the basis of a general class of gravitational field theories with torsion by


Curvature and Torsion from Embedding

It is also possible to construct spaces with curvature and torsion by an embedding procedure in a flat spacetime [1]. This is done by imposing suitable constraints. We shall do this for spaces with only positive signatures of the metric, so we can talk about spaces instead of spacetimes everywhere.

13.1 Curvature

Instead of mappings from the space $x^a$ to $x^\mu$ with rotational defects, there is another way to obtain curvature. This is by embedding the space $x^\mu$ into a higher-dimensional “Minkowskian” spacetime $x^A$, $A = 1, \ldots, N$ with a metric $g_{AB}$ consisting only of diagonal elements $\pm 1$. The mapping $x^A(x^\mu)$ is smooth but cannot be inverted to $x^\mu(x^A)$. Thus there are $N$ basis vectors $e_A$ in the embedding space and

$$e_\lambda(x^\mu) = e_A e^A_\lambda(x^\mu) = e_A \frac{\partial x^A}{\partial x^\lambda}$$  (13.1)

form four local tangent vectors in the 4-dimensional submanifold $x^A(x^\mu)$. They induce a metric

$$g_{\lambda\pi}(x^\mu) = e_\lambda(x^\mu) e_\kappa(x^\mu)$$  (13.2)

which can be used to define

$$e^{A\lambda}(x^\mu) = g^{\lambda\nu}(x^\mu)e^A_\nu(x^\mu)$$  (13.3)

but these vectors are no longer reciprocal to $e^A_\lambda(x^\mu)$, i.e.,

$$e^{A\lambda}e_B_\lambda \neq \delta^A_B$$  (13.4)

which is obvious since there are not enough of them. They do fulfill, however,

$$e^A_\lambda e^\kappa_A = \delta^\kappa_\lambda.$$  (13.5)
Let us take an example: the surface of a sphere of radius $a$ in three dimensions with the mapping
\[ x^A = (x^1, x^2, x^3) = a (\sin \theta \cos \varphi, \sin \theta, \sin \varphi, \cos \theta) \quad (13.6) \]
and vectors
\[ e^A_1 = a (\cos \theta \cos \varphi, \cos \theta, \sin \varphi, -\sin \theta) = e_{A1} \]
\[ e^A_2 = a (-\sin \theta \sin \varphi, \sin \theta \cos \varphi, 0) = e_{A2} \quad (13.7) \]
where we have set $x^\mu=1=\theta, x^\mu=2=\varphi$. The metric is
\[ g_{\mu\nu} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}, \quad g^{\mu\nu} = \frac{1}{a^2} \begin{pmatrix} 1 & 0 \\ 0 & \sin^{-2} \theta \end{pmatrix} \quad (13.8) \]
such that
\[ e_A^1 = e_{A1} = e^A_1, \quad e_A^2 = \frac{1}{a} \left( -\sin \varphi \cos \theta, \sin \varphi, 0 \right) \quad (13.9) \]
The connection is symmetric:
\[ \Gamma_{22}^1 = e_{A1} \partial_2 e^A_2 = a e_{A1} (-\sin \theta \cos \varphi, -\sin \theta \sin \varphi, 0) \]
\[ = -a^2 \sin \theta \cos \theta = -\Gamma_{212} = -\Gamma_{122}. \quad (13.10) \]
All other elements vanish. By raising the last index we obtain
\[ \Gamma_{22}^1 = -\sin \theta \cos \theta, \quad \Gamma_{21}^2 = \cot \theta. \quad (13.11) \]
The curvature tensor becomes
\[ R_{1221} = \partial_1 \Gamma_{22}^1 - \partial_2 \Gamma_{12}^1 - \Gamma_{12}^1 \Gamma_{21}^1 - \Gamma_{12}^2 \Gamma_{22}^1 + \Gamma_{22}^1 \Gamma_{11}^1 + \Gamma_{22}^2 \Gamma_{12}^1 \]
\[ = -\cos^2 \theta + \sin^2 \theta + \cot \theta \sin \theta \cos \theta = \sin^2 \theta, \quad (13.12) \]
implicating that
\[ R_{12}^{21} = \frac{1}{a^2}. \quad (13.13) \]
All other elements can be obtained using antisymmetry of $R_{\mu\nu\lambda\kappa}$ in $\mu \rightarrow \nu, \lambda \rightarrow \kappa$ and symmetry under $\mu\nu \leftrightarrow \lambda\kappa$, which is a consequence of the symmetry of $\Gamma_{\mu\nu}^\lambda$ in $\mu\nu$ [recall the derivation of (12.106)]. Thus we can form the Ricci tensor
\[ R_{\mu\nu}^\lambda \rho = R_{\nu}^\lambda = \frac{1}{a^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (13.14) \]
and the curvature scalar
\[ R = R_{\mu}^{\rho} = \frac{2}{a^2}. \quad (13.15) \]
13 Curvature and Torsion from Embedding

Note that for non-invertible vectors $e^A_\lambda$, the curvature has to be calculated from (11.127). The formula (11.129) can no longer be used since in the derivation of this formula one would need

$$\partial_\mu e^a_\nu = \Gamma^\lambda_{\mu\nu} e^a_\lambda$$

(13.16)

which no longer follows from the correct relation

$$\Gamma^\lambda_{\mu\nu} = e^A_A \partial_\mu e^A_\nu$$

(13.17)

due to the non-invertibility

$$e^A_\lambda e^B_\lambda \neq \delta^A_B.$$  

(13.18)

In general, it is possible to generate any curved spaces by embedding in a higher-dimensional Minkowski space. If the curved space has $d$ dimension, the embedding space has to have at least $d(d+1)/2$ dimension. This is seen by looking at the metric

$$g_{\mu\nu}(x) = \frac{\partial x^A}{\partial x^\mu} \frac{\partial x^B}{\partial x^\nu} g_{AB},$$

(13.19)

and noticing that for a given $d$ dimensional matrix, this converts to $d(d+1)/2$ differential equations for the functions $x^A(x^\mu)$.

13.2 Torsion

It is also possible to construct spaces with curvature and torsion by an embedding procedure in a flat space. This is done by imposing nonholonomic constraints. Parallel transport in the embedded space is determined as an induced parallel transport on the surface of constraints.

13.2.1 Strategy

In the last section we have defined a geometry by selecting a subspace from a Euclidean space with the help of constraints. This is not the only possible procedure. We may equally well impose constraints only on the velocity space of all possible trajectories in the Euclidean space. A metric is naturally induced in the embedded space by restricting the scalar product in the space of all curves in the bigger embedding space to the smaller embedded space. A connection is induced uniquely by compatibility conditions for the embedding of the velocity space with the parallel-transport law in the embedding space. This means the following: Take a curve in the original space connecting points 1 and 2. This curve is embedded into a bigger Euclidean space by specifying the velocity vectors of the image curve. A vector from the velocity space at point 1 is parallel-transported along the curve to point 2, and then it is embedded in the bigger space. We require that the resulting vector must
be the same as the one obtained by running in the opposite direction: The vector at point 1 is first embedded into a bigger Euclidean space and then parallel-transported along the image of the curve connecting points 1 and 2. This compatibility condition ensures that the connection in the original space is uniquely determined by the embedding law.

Constraints imposed on the velocity space can be nonholonomic, and this is the source of torsion. In this context, the notion of ”holonomic” and ”nonholonomic” constraints is exactly the same as in classical mechanics. For a mechanical system, generalized velocities are elements of the velocity space of its configuration space. Let the motion be subject to constraints linear in velocities. According to the Hertz classification [2], constraints are said to be holonomic if they are integrable (i.e., equivalent to some constraints on the configuration space only), and nonholonomic if they are non-integrable. Sometimes dynamical systems with nonholonomic constraints are simply called nonholonomic systems. It is important to realize that the motion of nonholonomic systems does not occur on any submanifold of the configuration space, nonetheless it is described by a smaller number of parameters than the corresponding unconstrained motion.

Upon an embedding via nonholonomic constraints on the velocity space, any curve that parallel-transports its velocity vector along itself has an image with same property. Therefore straight lines in the Euclidean space have natural images in the embedded space which are autoparallels. Using Gauss’ principle of least constraint it is possible to show that autoparallels describe a constrained motion such that the acceleration (or the force) induced by the constraints has a least deviation from the acceleration of the corresponding unconstrained motion, while geodesics play no special role in the constrained dynamics. See Section 14.2 for this derivation.

13.2.2 Nonholonomic Embedding

We denote the coordinates of the embedding Euclidean space by $x^A, A = 1, 2, ..., $ and those of the embedded space of smaller dimension by $q^\mu, \mu = 1, 2, ..., N$. Suppose we specify a set of transformation functions $\varepsilon^A_\mu(q)$ which map curves $q^\mu(t)$ into Euclidean curves $x^A(t)$ by a relation

$$v^A(t) = \varepsilon^A_\mu(q(t))v^\mu(t), \quad v^\mu = \dot{q}^\mu(t), \quad v^A = \dot{x}^A(t) ,$$

(13.20)

This implies an integral relation for the associated curves

$$x^A(t) = x^A(0) + \int_0^t dt' q^\mu(t')\varepsilon^A_\mu(q(t)) .$$

(13.21)

For any curve $q^\mu(t)$, this determines a curve in the Euclidean space up to a global translation. Thus it determines an embedding of the space of all paths in $q$-space into the space of all paths in $x$-space.

The acceleration along a curve in Euclidean space is defined by

$$\frac{dv^A(t)}{dt} = \ddot{q}^A(t) .$$

(13.22)
By analogy with (11.108), we define the covariant acceleration in the embedded space by another mapping of the type (13.20):

\[ \frac{dv^A(t)}{dt} \equiv \varepsilon^A_{\mu}(q(t)) \frac{Dv^\mu}{dt} . \] (13.23)

Let us seek for an affine connection \( \Gamma^\lambda_{\mu\nu} \) in \( q \)-space which describes this covariant acceleration intrinsically, i.e., without referring back to the embedding Euclidean space. It is given by

\[ \frac{Dv^\mu}{dt} = \dot{v}^\mu(t) + \Gamma^\mu_{\nu\sigma}(t)v^\nu(t)v^\sigma(t) . \] (13.24)

The important observation is that the property (13.23) fixed \( \Gamma^\lambda_{\mu\nu} \) uniquely.

In order to prove this statement, let us introduce some useful notations. Any two tangent vectors \( \tilde{v}^A \), \( v^B \) in Euclidean space have a scalar product

\[ \tilde{v} \cdot v = \delta_{ij} \tilde{v}^A v^B . \] (13.25)

If the vectors \( v^A \) and \( \tilde{v}^A \) are the images of two different velocities \( v^\mu \) and \( \tilde{v}^\mu \) at the space point \( q \), then the embedding coefficients \( \varepsilon^A_{\mu}(q) \) determine an induced metric in \( q \)-space

\[ \tilde{v} \cdot v = g_{\mu\nu}(q)\tilde{v}^\mu(q)v^\nu(q) , \quad g_{\mu\nu}(q) \equiv \varepsilon^A_{\mu}(q)\varepsilon^A_{\nu}(q) . \] (13.26)

It is useful to introduce the quantity

\[ \varepsilon^{\mu} = \varepsilon^A_{\mu}(q)g^{\mu\nu}(q) , \] (13.27)

where \( g^{\mu\nu}(q)g_{\lambda\nu}(q) = \delta^\mu_\nu \). From (13.26) follows that

\[ \varepsilon^{A\mu}(q)\varepsilon^{A\nu}(q) = g^{\mu\nu}(q) , \quad \varepsilon^{i\mu}(q)\varepsilon^A_{\nu}(q) = \delta^{i\mu}_{\nu} . \] (13.28)

Using the metrics \( g_{\mu\nu}(q) \) and \( \delta_{ij} \) to lower or raise indices in \( q \) - and \( x \) -spaces, respectively, the embedding condition (13.23) can be written in a more general form

\[ \frac{dv^{ij \ldots}}{dt} = \frac{d}{dt} \left( \varepsilon^A_{\mu}d^B_{\nu} \ldots \varepsilon_{x}^{\lambda} \varepsilon_{x}^{\beta} \ldots v^{\mu_{i \ldots}} \right) = \varepsilon^A_{\mu}d^B_{\nu} \ldots \varepsilon_{x}^{\lambda} \varepsilon_{x}^{\beta} \ldots \frac{Dv^{\mu_{i \ldots}}}{dt} , \] (13.29)

where the covariant derivative reads

\[ \frac{Dv^{\mu_{i \ldots}}}{dt} = i_{\alpha\beta \ldots}^{\mu_{i \ldots}} + \left( \Gamma^{\mu_{i \ldots}}_{\lambda\alpha \beta \ldots} + \cdots - \Gamma^{\mu_{i \ldots}}_{\alpha\lambda \beta \ldots} \right) v^\lambda . \] (13.30)

Performing the time derivatives on the left-hand side of (13.29) and applying relations (13.28) we find

\[ \Gamma^{\mu_{i \ldots}}_{\nu\lambda}(q)v^\lambda = \varepsilon^{\mu_{i \ldots}}(q)\varepsilon^A_{\nu}(q) = -\varepsilon^A_{\nu}(q)\varepsilon^{\mu_{i \ldots}}(q) . \] (13.31)

This equation ensures that along any curve \( q^\mu(t) \), the fields \( \varepsilon^A_{\mu}(q(t)) \) and \( \varepsilon^{\mu_{i \ldots}}(q(t)) \) are transported parallel, as expressed by the relations

\[ \frac{D}{dt}\varepsilon^A_{\mu}(q(t)) = 0 , \quad \frac{D}{dt}\varepsilon^{\mu_{i \ldots}}(q(t)) = 0 . \] (13.32)
Applying the covariant derivative $D/dt$ to the metric (13.26) we obtain from the chain rule of differentiation

$$\frac{D}{dt} g_{\mu\nu}(q(t)) = 0. \quad (13.33)$$

Equation (13.31) can be rewritten as

$$\Gamma^\lambda_{\nu\mu}(q) v^\lambda(q) = \varepsilon^A_{\lambda \alpha}(q) \partial_\nu \varepsilon^A_{\alpha \mu}(q) g^{\alpha\lambda}(q). \quad (13.34)$$

This must hold for any velocity $v^\lambda(t)$ along the curve $q^\lambda(t)$, implying that

$$\Gamma_{\mu\nu}^\lambda(q) = \varepsilon^A_{\lambda \alpha}(q) \partial_\nu \varepsilon^A_{\alpha \mu}(q) g^{\alpha\lambda}(q). \quad (13.35)$$

Thus we have succeeded in determining metric and parallel transport by an embedding of all paths in $q$-space in the bigger Euclidean $x$-space.

### 13.2.3 Torsion

Let us now turn to the analysis of the affine connection (13.35). First of all, we observe that the torsion tensor

$$S_{\nu\kappa}^\mu = \frac{1}{2} g^{\mu\lambda} \left[ \varepsilon^A_{\lambda \alpha}(q) \partial_\kappa \varepsilon^A_{\alpha \nu}(q) - \varepsilon^A_{\lambda \nu}(q) \partial_\kappa \varepsilon^A_{\alpha \mu}(q) \right] \quad (13.36)$$

is, in general, nonzero. The torsion induced by the embedding is zero if and only if

$$\varepsilon^A_{\nu,\mu}(q) = \varepsilon^A_{\mu,\nu}(q). \quad (13.37)$$

If this integrability condition is satisfied, the matrix elements $\varepsilon^A_{\lambda}(q)$ are the derivatives of functions $\varepsilon^A(q)$

$$\varepsilon^A_{\mu}(q) = \partial_\mu \varepsilon^A(q), \quad (13.38)$$

and the integral in Eq. (13.21) can be performed trivially to yield a point-to-point embedding

$$x^A = \varepsilon^A(q). \quad (13.39)$$

The metric tensor $g_{\mu\nu}$ has $N(N + 1)/2$ independent components. The torsion tensor $S_{\nu\kappa}^\mu$ has $N^2(N - 1)/2$ independent components. To embed a general metric space with torsion, the number $NM$, being the number of independent embedding coefficients $\varepsilon^A_{\mu}$, should be greater or equal to $N(N^2 + 1)/2$. This leads to the relation between the dimensions of $q$- and $x$-spaces:

$$(\dim[q])^2 + 1 \leq 2 \dim[x]. \quad (13.40)$$

### Notes and References

Nonholonomic Mapping Principle

The multivalued, nonholonomic mappings from flat to curved spacetime with torsion enable us to reformulate Einstein’s equivalence principle in a new more powerful way. Whereas Einstein postulated, that equations written down in flat spacetime with curvilinear coordinates remain valid in curved spacetime, we may sharpen this postulate to the form:

*The physical laws in curved spacetime are the direct images the flat-spacetime laws under multivalued mappings.*

If the space has only curvature and no torsion, this *new equivalence principle* leads to the same conclusions as Einstein’s.

If the space has torsion, however, the new equivalence principle makes new predictions, and it is interesting to investigate these.

14.1 Motion of Point Particle

The derivation of the geodesic trajectories of point particles in curved space was performed in Section 11.2. Only minor modifications will be necessary to follow the new equivalence principle. As observed when going from Eq. (11.14) to (11.15), we simplify the discussion by considering the nonrelativistic action \( \bar{A} \) of Eq. (11.16) if we use the proper time \( \tau \) to parameterize the paths.

14.1.1 Classical Action Principle for Spaces with Curvature

Instead of performing an ordinary coordinate transformation in flat space from Minkowski coordinates \( x^a \) to curvilinear coordinates \( x^\mu \) via Eq. (11.7), we perform a multivalued coordinate transformation

\[
dx^a = e^a_\mu(x)dx^\mu,
\]

where the basis vectors \( e^a_\mu \) describe coordinate transformations in which

\[
\partial_\mu e^a_\nu(x) - \partial_\nu e^a_\mu(x) \neq 0.
\]

379
This implies that second derivatives in front of the multivalued functions $x^a(x^\mu)$ do not commute as in Eq. (11.31):

$$\left(\partial_\lambda \partial_\kappa - \partial_\kappa \partial_\lambda\right) x^a(x) \neq 0,$$

thus violating the Schwarz integrability criterion. If the space has also torsion, then the functions $e^a_\nu(x)$ have also no commuting derivatives [recall (11.32)]:

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) e^a_\lambda(x) \neq 0.$$  

In either case, the metric in the image space has the same form as in Eq. (11.9), and the derivation of the extremum of the action seems, at first, to follow the same pattern as in Section 11.2, leading to the equation of motion (11.25) for geodesic trajectories. The nonholonomically transformed action (11.2) is independent of the torsion fields $S^\lambda_{\mu\nu}$, and for this reason also the equation of motion (11.25) is different to the presence of torsion.

This result would be perfectly acceptable, were it not for an apparent inconsistency with another result obtained by applying the new variational principle. Instead of transforming the action and varying it in the usual way, we may transform the equation of motion of a free particle (11.1) in flat space nonholonomically into a space with curvature and torsion.

### 14.1.2 Autoparallel Trajectories in Spaces with Torsion

In the absence of external forces, the equation of motion (11.1) in flat space states that the second derivative of $q^i(\tau)$ vanish. In spacetime, the free equations of motion read $\ddot{q}^a(\tau) = 0$, where the dot denotes the derivative with respect to the invariant length $s$ as parametrizing the straight lines, as in Eq. (11.25). These are transformed nonholonomically by Eq. (14.1) as follows:

$$\frac{d^2 q^a}{d\tau^2} = \frac{d}{d\tau} \left( e^a_\mu \dot{q}^\mu \right) = e^a_\mu \ddot{q}^\mu + e^a_{\mu,\nu} \dot{q}^\mu \dot{q}^\nu = 0,$$

or as

$$\ddot{q}^a + e^a_\mu e^\mu_{\kappa,\lambda} \dot{q}^\kappa \dot{q}^\lambda = 0. \tag{14.5}$$

The subscript $\lambda$ separated by a comma denotes the partial derivative: $f_{,\lambda}(x) \equiv \partial_\lambda f(x)$. The quantity in front of $\dot{q}^a \dot{q}^\lambda$ is the affine connection (11.92). Thus we arrive at the transformed flat-space equation of motion

$$\ddot{q}^\mu + \Gamma^\mu_{\kappa\lambda} \dot{q}^\kappa \dot{q}^\lambda = 0. \tag{14.6}$$

The solutions of this equation are called *autoparallel* trajectories. They differ from the geodesic trajectories described by (11.25) by an extra torsion term. Inserting the decomposition (11.119) and using the antisymmetry of $S^\lambda_{\mu\nu}$ in the first two indices, we may rewrite (14.7) as

$$\ddot{q}^\mu + \Gamma^\mu_{\kappa\lambda} \dot{q}^\kappa \dot{q}^\lambda - 2S^\mu_{\kappa\lambda} \dot{q}^\kappa \dot{q}^\lambda = 0. \tag{14.8}$$
14.1 Motion of Point Particle

Note the index positions of the torsion tensor, which may be written more explicitly as $S^\mu_{\kappa\lambda} \equiv g^{\mu\rho} g_{\kappa\sigma} S_{\sigma\kappa}$. This is not antisymmetric in the last two indices so that it possesses a symmetric part which contributes to Eq. (14.7).

How can we reconcile this result with an application of the new equivalence principle applied to the action. Since the transformed action is independent of the torsion and carries only information on the Riemann part of the space geometry, torsion can enter the equations of motion only via some overlooked feature of the variation procedure. Indeed, a moment’s thought convinces us that this was applied incorrectly in the previous section. According to the new equivalence principle we must also transform the variational procedure nonholonomically to spaces with curvature and torsion. We must find the image of the flat-space variations $\delta q^a(\tau)$ under the multivalued mapping

$$\dot{q}^\mu = c^\mu_a(q) \dot{q}^a.$$ (14.9)

The images are quite different from ordinary variations as illustrated in Fig. 14.1(a). The variations of the Cartesian coordinates $\delta q^a(\tau)$ are done at fixed endpoints of the paths. Thus they form closed paths in the $x$-space. Their images, however, lie in a space with defects and thus possess a closure failure indicating the amount of torsion introduced by the mapping. This property will be emphasized by writing the images $\delta^S q^\mu(\tau)$ and calling them nonholonomic variations. The superscript indicates the special feature caused by torsion.

Let us calculate them explicitly. The paths in the two spaces are related by the integral equation

$$q^\mu(\tau) = q^\mu(\tau_a) + \int_{\tau_a}^{\tau} d\tau' e^\mu_a(q(\tau')) \dot{q}^a(\tau').$$ (14.10)

For two neighboring paths in $x$-space differing from each other by a variation $\delta q^a(\tau)$, equation (14.10) determines the nonholonomic variation $\delta^S q^\mu(\tau)$:

$$\delta^S q^\mu(\tau) = \int_{\tau_a}^{\tau} d\tau' \delta^S [e^\mu_a(q(\tau')) \dot{q}^a(\tau')].$$ (14.11)

A comparison with (14.9) shows that the variation $\delta^S$ and the derivatives with respect to the parameter $s$ of $q^\mu(\tau)$ commute with each other:

$$\delta^S \dot{q}^\mu(\tau) = \frac{d}{d\tau} \delta^S q^\mu(\tau),$$ (14.12)

just as for ordinary variations $\delta q^a$ in Eq. (2.7):

$$\dot{\delta q}^a(\tau) = \frac{d}{d\tau} \delta q^a(\tau).$$ (14.13)

Let us also introduce auxiliary nonholonomic variations of the paths $q^\mu(\tau)$ in $x^\mu$-space:

$$\delta q^\mu \equiv e^\mu_a(q) \delta q^a.$$ (14.14)
In contrast to $\delta S_{q\mu}(\tau)$, these vanish at the endpoints,

$$\delta q(\tau_a) = \delta q(\tau_b) = 0,$$

just as the usual variations $\delta q^a(\tau)$, i.e., they form closed paths with the unvaried orbits.

Using (14.12), (14.13), and the fact that $\delta S_{q\mu}(\tau) \equiv \delta q^a(\tau)$, by definition, we derive from (14.11) the relation

$$d\frac{d}{d\tau} \delta S_{q\mu}(\tau) = \delta S_{e\mu}(q(\tau))q^a(\tau) + e^\mu_{\nu}(q(\tau))\frac{d}{d\tau} \delta q^a(\tau)$$

$$= \delta S_{e\mu}(q(\tau))q^a(\tau) + e^\mu_{\nu}(q(\tau))\frac{d}{d\tau}[e^a_{\nu}(\tau) \delta q^\nu(\tau)].$$

(14.16)

After inserting

$$\delta S_{e\mu}(q) = -\Gamma^\nu_{\lambda\nu} \delta S_{q\lambda}(q) e^a_{\mu}, \quad \frac{d}{d\tau} e^a_{\nu}(q) = \Gamma^\mu_{\lambda\nu} q^\lambda e^a_{\mu},$$

(14.17)

this becomes

$$d\frac{d}{d\tau} \delta S_{q^\mu}(\tau) = -\Gamma^\nu_{\lambda\nu} \delta S_{q^\lambda}(q) q^\nu + \Gamma^\mu_{\lambda\nu} q^\lambda \delta q^\nu + \frac{d}{d\tau} \delta q^\mu.$$ 

(14.18)

It is useful to introduce the difference between the nonholonomic variation $\delta S_{q^\mu}$ and an auxiliary closed nonholonomic variation $\delta q^\mu$:

$$\delta S b^\mu \equiv \delta S_{q^\mu} - \delta q^\mu.$$ 

(14.19)

Then we can rewrite (14.18) as a first-order differential equation for $\delta S b^\mu$:

$$d\frac{d}{d\tau} \delta S b^\mu = -\Gamma^\nu_{\lambda\nu} \delta S b^\lambda q^\nu + 2S^\mu_{\lambda\nu} q^\lambda \delta q^\nu.$$ 

(14.20)

After introducing the matrices

$$G^\mu_{\lambda}(\tau) \equiv \Gamma^\mu_{\lambda\nu}(q(\tau))q^\nu(\tau)$$

(14.21)

and

$$\Sigma^\mu_{\nu}(\tau) \equiv 2S^\mu_{\lambda\nu}(q(\tau))q^\lambda(\tau),$$

(14.22)

equation (14.20) can be written as a vector differential equation:

$$d\frac{d}{d\tau} \delta S b = -G \delta S b + \Sigma(\tau) \delta q^\nu(\tau).$$ 

(14.23)

Although not necessary for the further development, we solve this equation by

$$\delta S b(\tau) = \int^{\tau}_{\tau_a} d\tau' U(\tau, \tau') \Sigma(\tau') \delta q(\tau'),$$

(14.24)
14.1 Motion of Point Particle

Figure 14.1 Images under holonomic and nonholonomic mapping of fundamental $\delta$-function path variation. In the holonomic case, the paths $q(\tau)$ and $q(\tau) + \delta q(\tau)$ in (a) turn into the paths $q(\tau)$ and $q(\tau) + \delta q(\tau)$ in (b). In the nonholonomic case with $S_{\lambda\mu\nu} \neq 0$, they go over into $q(\tau)$ and $q(\tau) + \delta S q(\tau)$ shown in (c) with a closure failure $b^\mu$ analogous to the Burgers vector $b^\mu$ in a solid with dislocations.

with the matrix

$$
U(\tau, \tau') = \hat{T}_s \exp \left[ -\int_{\tau}^{\tau'} d\tau'' G(\tau'') \right],
$$

(14.25)

where $\hat{T}_s$ denotes the time-ordering operator for the parameter $s$. In the absence of torsion, $\Sigma(\tau)$ vanishes identically and $\delta S b(\tau) \equiv 0$, and the variations $\delta S q^\mu(\tau)$ coincide with the auxiliary closed nonholonomic variations $\delta q^\mu(\tau)$ [see Fig. 14.1(b)]. In a space with torsion, the variations $\delta S q^\mu(\tau)$ and $\delta q^\mu(\tau)$ are different from each other [see Fig. 14.1(c)].

We now calculate the variation of the action (11.11) under an arbitrary nonholonomic variation $\delta^S q^\mu(\tau) = \delta q^\mu + \delta S b^\mu$. Since $s$ is the invariant path length, we may just as well use the auxiliary action (11.16) to calculate this quantity (it differs only by a trivial factor 2):

$$
\delta^S \bar{A} = M \int_{\tau_a}^{\tau_b} d\tau \left( g_{\mu\nu} \dot{q}^\mu \delta^S \dot{q}^\nu + \frac{1}{2} \partial^\mu g_{\lambda\kappa} \delta^S q^\mu \dot{q}^\lambda \dot{q}^\kappa \right).
$$

(14.26)
After a partial integration of the $\delta q$-term we use (14.15), (14.12), and the identity 
$\partial_\mu g_{\mu \lambda} \equiv \Gamma^\mu_{\mu \lambda} + \Gamma^\mu_{\mu \nu}$, which follows directly form the definitions $g_{\mu \nu} \equiv e^a_\mu e^a_\nu$ and $\Gamma^\mu_{\mu \nu} \equiv e^a_\mu \partial_\mu e^a_\nu$, we obtain

$$
\delta^S \tilde{A} = M \int_{t_a}^{t_b} dt \left[ -g_{\mu \nu} \left( \ddot{q}^\nu + \bar{\Gamma}_{\lambda \kappa}^{\nu \lambda \kappa} \dot{q}^\lambda \dot{q}^\kappa \right) \delta q^\mu + \left( g_{\mu \nu} \ddot{q}^\nu \frac{d}{dt} \delta^S b^\mu + \Gamma^\mu_{\mu \lambda} \delta^S b^\lambda \dot{q}^\kappa \right) \right].
$$

(14.27)

To derive the equation of motion we first vary the action in a space without torsion. Then $\delta^S b^\mu (\tau) \equiv 0$, and (14.27) becomes

$$
\delta^S \tilde{A} = -M \int_{t_a}^{t_b} dt g_{\mu \nu} \left( \ddot{q}^\nu + \bar{\Gamma}_{\lambda \kappa}^{\nu \lambda \kappa} \dot{q}^\lambda \dot{q}^\kappa \right) \delta q^\mu.
$$

(14.28)

Thus, the action principle $\delta^S \tilde{A} = 0$ produces the equation for the geodesics (11.25), which are the correct particle trajectories in the absence of torsion.

In the presence of torsion, $\delta^S b^\mu$ is nonzero, and the equation of motion receives a contribution from the second parentheses in (14.27). After inserting (14.20), the nonlocal terms proportional to $\delta^S b^\mu$ cancel and the total nonholonomic variation of the action becomes

$$
\delta^S \tilde{A} = -M \int_{t_a}^{t_b} dt g_{\mu \nu} \left( \ddot{q}^\nu + \bar{\Gamma}_{\lambda \kappa}^{\nu \lambda \kappa} \dot{q}^\lambda \dot{q}^\kappa \right) \delta q^\mu.
$$

(14.29)

The second line follows from the first after using the identity $\bar{\Gamma}_{\lambda \kappa}^{\nu \lambda \kappa} = \tilde{\Gamma}_{\lambda \kappa}^{\nu \lambda \kappa} + 2S^\nu_{\{\lambda \kappa\}}$. The curly brackets indicate the symmetrization of the enclosed indices. Setting $\delta^S \tilde{A} = 0$ and inserting for $\delta q(\tau)$ the image under (14.14) of an arbitrary $\delta$-function variation $\delta q^a(\tau) \propto \epsilon^a \delta (\tau - s_0)$ gives the autoparallel equations of motions (14.7), which is what we wanted to show.

The above variational treatment of the action is still somewhat complicated and calls for a simpler procedure [1, 2]. The extra term arising from the second parenthesis in the variation (14.27) can be traced to a simple property of the auxiliary closed nonholonomic variations (14.14). To find this we form the time derivative $d_t \equiv d/dt$ of the defining equation (14.14) and find

$$
d_t \delta q^\mu (\tau) = \partial_\nu e^\mu_a (q(\tau)) \dot{q}^\nu (\tau) \delta q^a (\tau) + e^\mu_a (q(\tau)) d_s \delta q^a (\tau).
$$

(14.30)

Let us now perform variation $\delta$ and $s$-derivative in the opposite order and calculate $d_s \delta q^\mu (\tau)$. From (14.9) and (11.41) we have the relation

$$
d_s q^\lambda (\tau) = e^\lambda_i (q(\tau)) d_s q^i (\tau).
$$

(14.31)

Varying this gives

$$
\delta d_s q^\mu (\tau) = \partial_\nu e^\mu_a (q(\tau)) \delta q^\nu d_s q^a (\tau) + e^\mu_a (q(\tau)) \delta d_s q^a.
$$

(14.32)
14.1 Motion of Point Particle

Since the variation in $x^a$-space commute with the $s$-derivatives [recall (14.13)], we obtain

$$
\delta d_\tau q^\mu(\tau) - d_\tau \delta q^\mu(\tau) = \partial_\nu e^\mu_a(q(\tau)) \delta q^\nu d_\tau q^a(\tau) - \partial_\nu e^\mu_a(q(\tau)) \dot{q}^\nu(\tau) \delta q^a(\tau). \quad (14.33)
$$

After re-expressing $\delta q^\mu(\tau)$ and $d_\tau q^a(\tau)$ back in terms of $\delta q^\mu(\tau)$ and $d_\tau q^a(\tau) = \dot{q}^\mu(\tau)$, and using (11.92), this becomes

$$
\delta d_\tau q^\mu(\tau) - d_\tau \delta q^\mu(\tau) = 2S^\mu_{\nu\lambda} \dot{q}^\nu(\tau) \delta q^\lambda(\tau). \quad (14.34)
$$

Thus, due to the closure failure in spaces with torsion, the operations $d_\tau$ and $\delta$ do not commute in front of the path $q^\mu(\tau)$. In other words, in contrast to the open variations $\delta q^\mu(\tau)$ (and of course the usual $\delta q^\mu(\tau)$), the auxiliary closed nonholonomic variations $\delta$ of velocities $\dot{q}^\mu(\tau)$ no longer coincide with the velocities of variations.

This property is responsible for shifting the trajectory from geodesics to autoparallels. Indeed, let us vary an action

$$
\bar{A} = \int_{\tau_a}^{\tau_b} d\tau L(q^\mu(\tau), \dot{q}^\mu(\tau)) \quad (14.35)
$$

directly by $\delta q^\mu(\tau)$ and impose (14.34), we find

$$
\delta \bar{A} = \int_{\tau_a}^{\tau_b} d\tau \left\{ \frac{\partial L}{\partial q^\mu} \delta q^\mu + \frac{dL}{dq^\mu} \frac{d}{d\tau} \delta q^\mu + 2S^\mu_{\nu\lambda} \frac{\partial L}{\partial \dot{q}^\nu} \delta q^\lambda \right\}. \quad (14.36)
$$

After a partial integration of the second term using the vanishing $\delta q^\mu(\tau)$ at the endpoints, we obtain the Euler-Lagrange equation

$$
\frac{\delta \bar{A}}{\delta q^\mu} = \frac{\partial L}{\partial q^\mu} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\mu} = \frac{\delta \bar{A}}{\delta q^\mu} - 2S^\mu_{\nu\lambda} \dot{q}^\nu \frac{\partial L}{\partial \dot{q}^\lambda} = -2S^\mu_{\nu\lambda} \dot{q}^\nu \frac{\partial L}{\partial \dot{q}^\lambda}. \quad (14.37)
$$

This differs from the standard Euler-Lagrange equation by the additional torsion force. For the action (11.11), we thus obtain the equation of motion

$$
\ddot{q}^\mu + \left[ g^{\mu\kappa}(\partial_\nu g_{\lambda\kappa} - \frac{1}{2} \partial_\kappaos g_{\nu\lambda}) + 2S^\mu_{\nu\lambda} \dot{q}^\nu \dot{q}^\lambda \right] = 0, \quad (14.38)
$$

which is once more the equation (14.7) for autoparallels.

Thus a consistent application of the new equivalence principle yields consistently autoparallel trajectories for point particles in space with curvature and torsion.

14.1.3 Special Properties of Gradient Torsion

A torsion tensor which consists of an antisymmetric combination of gradients of a scalar field $\theta(q)$ as follows:

$$
S_{\mu\nu}^\lambda(q) = \frac{1}{2} \left[ \delta^\lambda_{\nu} \partial_\mu \theta(q) - \delta^\lambda_{\mu} \partial_\nu \theta(q) \right], \quad (14.39)
$$
is called \textit{gradient torsion}. If spacetime possesses only gradient torsion, its effect upon the equations of motion of a point particle can be simulated in a purely Riemannian spacetime by adding the scalar field $\theta(q)$ in a suitable way to the action. Then gradient torsion appears as a nongeometric external field. By extremizing the transformed action in the usual way, the resulting equation of motion coincides with the autoparallel equation derived in the initial spacetime with torsion from the modified variational principle in Eqs. (14.8):

\[
\ddot{q}^\mu + \Gamma^\mu_{\kappa\lambda} \dot{q}^\kappa \dot{q}^\lambda - 2S^\mu_{\kappa\lambda} \dot{q}^\kappa \dot{q}^\lambda = 0.
\] (14.40)

For a pure gradient torsion (14.39), this becomes

\[
\ddot{q}^\lambda(s) + \Gamma^\lambda_{\mu\nu}(q(s)) \dot{q}^\mu(s) \dot{q}^\nu(s) = -\dot{\theta}(q(s)) \dot{q}^\lambda(s) + g^{\kappa\lambda}(q(s)) \partial_\kappa \theta(q(s)),
\] (14.41)

with the extra terms on the right-hand side reflecting the closure failure of parallelograms caused by the torsion.

The same trajectory is found from the following alternative action in a purely Riemannian spacetime

\[
\mathcal{A} = -mc \int_{\sigma_n}^{\sigma_f} d\sigma \ e^{\theta(q)} \left[ g_{\mu\nu}(q(\sigma)) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma) \right]^{\frac{1}{2}}.
\] (14.42)

The extra factor $e^{\theta(q)}$ has precisely the same effect in a Riemannian spacetime as the gradient torsion (14.39) in a Riemann-Cartan spacetime. Indeed, the extremum of this action can be derived from the geodesic trajectory without the $\sigma$-field by introducing, for a moment, an auxiliary metric

\[
\tilde{g}_{\mu\nu}(q) \equiv e^{2\theta(q)} \equiv g_{\mu\nu}(q).
\] (14.43)

The invariant line element remains, of course,

\[
ds = \left[ g_{\mu\nu}(q(\sigma)) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma) \right]^{\frac{1}{2}} = e^{-\theta(q)} \left[ \tilde{g}_{\mu\nu}(q(\sigma)) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma) \right]^{\frac{1}{2}} = e^{-\theta(q)} d\tilde{s}.
\] (14.44)

By varying the action as in Eqs. (11.14)–(11.19), we obtain the modified equation of motion (11.21):

\[
\tilde{g}_{\lambda\nu} \frac{d^2 q^\nu(\sigma)}{d\sigma^2} + \left( \partial_\mu \tilde{g}_{\lambda\nu} - \frac{1}{2} \partial_\lambda \tilde{g}_{\mu\nu} \right) \frac{dq^\mu(\sigma)}{d\sigma} \frac{dq^\nu(\sigma)}{d\sigma} = 0.
\] (14.45)

Inserting (14.43) and (14.44), this becomes

\[
g_{\lambda\nu} \left( \frac{d^2 q^\nu(\sigma)}{d\sigma^2} - \dot{\theta} \frac{dq^\nu(\sigma)}{d\sigma} \right) + \left( \partial_\mu g_{\lambda\nu} - \frac{1}{2} \partial_\lambda g_{\mu\nu} \right) \frac{dq^\mu(\sigma)}{d\sigma} \frac{dq^\nu(\sigma)}{d\sigma} + 2\dot{\theta}(q) \frac{dq^\nu(\sigma)}{d\sigma} - \partial_\lambda \theta(q) g_{\mu\nu} \frac{dq^\mu(\sigma)}{d\sigma} \frac{dq^\nu(\sigma)}{d\sigma} = 0.
\] (14.46)

This coincides with the autoparallel trajectory (14.41).
14.2 Autoparallel Trajectories from Embedding

There exists another way of deriving autoparallel trajectories. Instead of using multivalued mappings to carry physical laws from flat space to spaces with curvature and torsion, we may use the embedding of Section 13.1 to do so.

14.2.1 Special Role of Autoparallels

Let us first remark that apart from extremizing a length between two fixed endpoints, geodesics in a Riemann space can be obtained by embedding the Riemann space in a Euclidean space of a higher dimension. This is done by imposing certain constraints on the Cartesian coordinates spanning the Euclidean space. The points on the constraint surface constitute the embedded Riemann space. Straight lines in the Euclidean space, which are geodesic and autoparallel and also determine a free motion in that space, become geodesics when the motion is restricted to the constraint surface. The restriction of the free motion to the constraint surface is done in a conventional way, i.e., by adding the equations of constraints to the equations of motion. When the constraint force is removed, geodesic trajectories turn into straight lines in the embedding space.

For curved space with torsion the embedding procedure was described in Chapter 13. The consequences for the trajectories were worked out in Ref. [3]. It turns out that also from this point of view, autoparallel curves are specially favored geometric curves in the embedded space. They are the images of straight lines in the embedding space by Eq. (13.23).

14.2.2 Gauss Principle of Least Constraint

There is also a mechanical argument favoring autoparallel over geodesic motion. This is intrinsically linked with the concept of inertia. Inertia favors trajectories whose acceleration deviates minimally from the acceleration of the corresponding unconstrained motion. This property can be formulated mathematically by means of Gauss’ principle of least constraint [4, 5].

Consider a Lagrangian system in the space \([x]\) with a Lagrangian \(L = L(x, v)\). At each moment of time, a state of the system can be labeled by a point in phase space \((x^i(t), v^i(t))\). Let \(H_{ij}(x^i(t), v^i(t)) = \partial^2 L / \partial v^i \partial v^B\) be the Hessian matrix of the system. Consider two paths \(x^i_1(t)\) and \(x^i_2(t)\). Gauss’ deviation function (sometimes also called Gauss’ constraint) for two paths reads

\[
G = \frac{1}{2} \left( \dot{v}^i_1 - \dot{v}^i_2 \right) H_{ij} \left( \dot{v}^B_1 - \dot{v}^B_2 \right). \tag{14.47}
\]

It measures the deviation of two possible motions from one another [4, 5].

Now, let the motion in \(x\)-space be subject to constraints. All paths \(x^i(t)\) allowed by the constraints are called conceivable motions. A path \(\bar{x}^i(t)\) is called released motion if it satisfies the Euler-Lagrange equations for the Lagrangian \(L\) without
constraint. Gauss’ principle of least constraint says that the deviation of conceivable motions from a released motion takes a stationary value for the actual motion.

In our case, the released motion is a free motion with zero acceleration \( \ddot{x} = 0 \). Accelerations of the conceivable motions satisfy

\[
\dot{v}^i = \varepsilon^i_{\mu} \dot{v}^\mu + \varepsilon^i_{\mu,\nu} v^\mu v^\nu.
\]  (14.48)

Since \( H_{ij} = \delta_{ij} \) for the Euclidean Lagrangian, Gauss’ deviation function (14.47) assumes the form

\[
G = \frac{1}{2} [\dot{v}^i]^2 = \frac{1}{2} [\dot{v}^\mu + \varepsilon^\mu_{\nu} \dot{v}^\nu]^2,
\]  (14.49)

where an infinitesimal factor \( d\tau^2 \) has been removed. Remembering Eq. (13.31), we may also write

\[
G = \frac{1}{2} \left[ \frac{Dv^\mu}{dt} \right]^2.
\]  (14.50)

This has a minimum at \( G \geq 0 \), which is reached for trajectories satisfying the autoparallel equation of motion

\[
\frac{Dv^\mu}{d\tau} = 0.
\]  (14.51)

Another derivation of the autoparallel equation (14.51) rests on the d’Alembert-Lagrange principle [4, 5]. In theoretical mechanics, one defines a Lagrange derivative

\[
[L]_A \equiv \frac{d}{d\tau} \frac{\partial L}{\partial v^A} - \frac{\partial L}{\partial x^A}.
\]  (14.52)

The d’Alembert-Lagrange principle asserts that motion of a system with the Lagrangian \( L \) proceeds such that

\[
v^A [L]_A = 0
\]  (14.53)

for all velocities allowed by the constraints. Taking the free Lagrangian \( L = v^A v_A/2 \) with \( [L]_A = \delta_{AB} \dot{v}^B \), and the constraint (13.20) we find that only the autoparallel equation (14.51) satisfies this principle.

Finally we point out that the motion of a holonomic system is completely determined by the restriction of the Lagrangian to the constraining surface [4]. Thus, holonomic constrained systems are indistinguishable from ordinary unconstrained Lagrangian systems. This is not true for nonholonomic systems, meaning that the Euler-Lagrange equations for the Lagrangian restricted on the constraining surface do not coincide with the original equations for the constrained motion. This difficulty prevents us from applying a conventional Hamiltonian formalism to the autoparallel motion, and subjecting it to a canonical quantization. In other words, Dirac’s method of quantizing constrained systems [6] does not apply to nonholonomic systems since these do not follow the conventional Lagrange formalism [4].
14.3 Maxwell-Lorentz Orbits as Autoparallel Trajectories

It is rather straightforward to set up the Maxwell-Lorentz equations for the motion of a charged particle in curved space. We rewrite the flat-spacetime equation of motion (1.165) as

\[
\frac{d^2 q^a(\tau)}{d\tau^2} = \frac{e}{c} F^a_{\ b}(q(\tau)) \frac{d}{d\tau} q^b(\tau),
\]

and subject this to a multivalued mapping. This adds the tidal forces to the acceleration term and leads to

\[
\ddot{q}^\lambda(\tau) + \Gamma^\lambda_{\mu\nu} \dot{q}^\mu \dot{q}^\nu = \frac{e}{cm} F^a_{\ b}(q(\tau)) q^b(\tau).
\]

It is now interesting to observe that this equation of motion may be viewed as an autoparallel motion in an affine geometry with torsion. The torsion is created on the orbit of the particle from the equation [7]

\[
S^\mu_{\ \nu\lambda} = \frac{e}{mc^3} F^\mu_{\ \nu} \dot{q}^\lambda.
\]

Indeed, if we insert this torsion into the autoparallel equation (14.8), we obtain the Maxwell-Lorentz equation (14.55) in curved spacetime.

14.4 Bargmann-Michel-Telegdi Equation from Torsion

Interestingly enough, also the spin precession equation (1.318) can be understood as a purely geometric equation in a space with torsion. If we transform the flat-spacetime equation (1.318) to curved spacetime, it becomes

\[
\frac{\bar{D}}{d\tau} S^a = \frac{e}{2mc} \left[ g F^a_{\ b} S^b + \frac{g - 2}{m^2 c^2} p^a F^\rho\kappa S^\rho p_\kappa \right] = 0.
\]

For a classical particle for which \( g = 1 \), this equation is the same as for a spin vector undergoing a parallel transport along the trajectory \( q^\mu(\tau) \) according to the law (11.111):

\[
\frac{DS^\mu}{d\tau} = \frac{\bar{D}S^\mu}{d\tau} + S^\mu_{\ \nu\lambda} \dot{S}^\nu - S^\mu_{\ \nu\lambda} \dot{S}^\nu + S^\lambda_{\ \mu\nu} \dot{S}^\nu\dot{q}^\lambda.
\]

Inserting (14.56) yields

\[
\frac{DS^\mu}{d\tau} = \frac{\bar{D}S^\mu}{d\tau} + \frac{e}{mc^3} \left( F^\mu_{\ \nu} \dot{q}_\lambda S^\nu - F^\mu_{\ \nu} \dot{q}_\lambda S^\nu + F^\nu_{\ \lambda} \dot{q}_\nu S^\mu \dot{q}^\lambda \right).
\]

Recalling the transversality (1.299) of the spin vector, the last term vanishes, and we arrive at

\[
\frac{DS^\mu}{d\tau} = \frac{\bar{D}S^\mu}{d\tau} + \frac{e}{mc^3} \left( F^\mu_{\ \nu} S^\nu - \frac{1}{c^2} \dot{q}^\mu S^\nu F_{\nu\lambda} \dot{q}^\lambda \right),
\]

which is indeed the same as (14.57) for \( g = 1 \).
Notes and References


Field Equations of Gravitation

In the previous chapter, we have shown that a particle subject to a gravitational field follows equations of motion which look precisely the same as those of a particle in Minkowski space if these are expressed in curvilinear coordinates. The information on the gravitational field is contained in certain properties of the metric. We may now ask the question how to calculate such a metric associated with a gravitational massive object. For this, the ten components of the metric tensor $g_{\mu\nu}(x)$ have to be considered as dynamical variables and we need an action principle for it [1, 2, 3, 4].

15.1 Invariant Action

The equation of motion for $g_{\mu\nu}(x)$ must be independent of the general coordinates employed. This is guaranteed if the action is invariant under Einstein transformations $x^\mu \to x'^\mu(x^\mu)$, whose infinitesimal form is

$$dx^\mu \to dx'^\mu = \alpha'^{\mu}_{\mu}(x) dx^\mu.$$  \hfill (15.1)

We want to set up a local action for the gravitational field, which by definition in Subsection 2.3.1 should have the form of an integral over a Lagrangian density

$$\mathcal{A} = \int d^4x \mathcal{L}(x),$$ \hfill (15.2)

where $\mathcal{L}(x)$ is a function of the metric and its derivatives, depending at most quadratically on the derivatives $\partial_{\lambda}g_{\mu\nu}(x)$ (after a possible integration by parts). Under the coordinate transformations (15.1), the volume element transforms as

$$d^4x \to d^4x' = d^4x \det \alpha.$$ \hfill (15.3)

The simplest Lagrangian density $\mathcal{L}(x)$ which leaves the action (15.2) invariant can be formed from the determinant of the metric

$$g = \det(g_{\mu\nu}).$$ \hfill (15.4)
Since the metric changes under (15.1) to 
\[ g'_{\mu\nu}(x') = g_{\mu\nu}(\alpha^{-1})^\mu_{\mu'}(\alpha^{-1})^\nu_{\nu'}, \]
we see that 
\[ g \rightarrow g' = g \det \alpha^{-2}, \tag{15.5} \]
implying that an action 
\[ \mathcal{A}_c = \frac{\Lambda}{\kappa} \int d^4x \sqrt{-g} \tag{15.6} \]
is invariant under coordinate transformations (15.1). Such an action by itself is, however, not capable of giving equations of motion for the gravitational field since it contains no derivatives of \( g_{\mu\nu} \). The \( g_{\mu\nu} \)-field cannot propagate. We must find some scalar Lagrangian density \( L \) containing \( g_{\mu\nu} \) and \( \partial_\lambda g_{\mu\nu} \) in a local way. Then 
\[ \mathcal{A} = \int d^4x \sqrt{-g} L(g, \partial g) \tag{15.7} \]
would be an invariant action which could govern the gravitational forces.

Now, the only fundamental scalar quantity which occurred in the previous geometric analysis and which involves the derivatives \( \partial_\lambda g_{\mu\nu} \) is \( R \), the scalar curvature. Therefore, Einstein postulated the following gravitational field action 
\[ \mathcal{A} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \tilde{R}. \tag{15.8} \]
Here \( \kappa \) is a constant related to Newton’s gravitational coupling \( G_N \approx 6.673 \cdot 10^{-8} \) \( \text{cm}^3 \text{g}^{-1} \text{s}^{-2} \) of Eq. (1.3) by 
\[ \frac{1}{\kappa} = \frac{c^3}{8\pi G_N}. \tag{15.9} \]
It can be expressed in terms of the Planck length (28.39) as 
\[ \frac{1}{\kappa} = \frac{\hbar}{8\pi l_P^2}. \tag{15.10} \]
For a system consisting of a set of mass points \( m_1, \ldots, m_N \), we add the particle action (11.2) and obtain a total action 
\[ \mathcal{A} = \mathcal{A} - \sum_{n=1}^{N} m_n c \int ds_n = \mathcal{A} + m. \tag{15.11} \]
In the following formulas it will be convenient to set \( \kappa = 1 \) since \( \kappa \) can always be reintroduced as a relative factor between field and matter parts in all field equations to be derived.

Variation of the particle paths \( x_n(s_n) \) at fixed \( g_{\mu\nu}(x) \) gives the equations of motion of a point particle in an external gravitational field as discussed in the beginning.
addition, the action (15.11) permits to find out which gravitational field generated by the presence of these points. They are obtained from the variational equation

\[ \delta A = 0. \] (15.12)

There are 10 independent components of \( g_{\mu \nu} \). Four of them are unphysical, representing merely reparametrization degrees of freedom.

Equations (15.12) are not sufficient to determine the geometry of spacetime. The curvature tensor \( R_{\mu \nu \lambda}^\kappa \) also contains torsion tensors \( S_{\mu \nu}^\lambda \) combined to a contortion tensor \( K_{\mu \nu}^\lambda \). It has 24 independent components, which are determined by the equation of motion

\[ \delta A = 0. \] (15.13)

Einstein avoided this problem by considering only symmetric (Riemannian) spaces from the outset. For spinning matter, however, this may not be sufficient, and a determination of torsion fields from the spin densities may be necessary for a complete dynamical theory.

### 15.2 Energy-Momentum Tensor and Spin Density

It is useful to study separately the derivatives of the different pieces of the action with respect to \( g_{\mu \nu} \) and \( K_{\mu \nu}^\lambda \) separately. In view of the physical interpretations to be given later we introduce

\[ \delta m A = - \frac{1}{2} \sqrt{-g} T_{\mu \nu}^m, \] (15.14)

\[ \delta f A = - \frac{1}{2} \sqrt{-g} T_{\mu \nu}^f, \] (15.15)

respectively, as the symmetric energy-momentum tensors of matter and field, and

\[ \delta m A = - \frac{1}{2} \sqrt{-g} \sum_{\nu \lambda} T_{\mu \nu}^{m}, \] (15.16)

\[ \delta f A = - \frac{1}{2} \sqrt{-g} \sum_{\nu \lambda} T_{\mu \nu}^{f}, \] (15.17)

as the spin current density of matter and field, respectively.

We have remarked before that the identity (11A.25) implies a change of sign if we calculate the energy-momentum tensors from a variation \( \delta g_{\mu \nu} \) rather than \( \delta g_{\mu \nu} \) so that Eqs. (15.14) and (15.15) go over into

\[ \delta m A = \frac{1}{2} \sqrt{-g} T_{\mu \nu}^m, \] (15.18)
\[ \frac{\delta f}{\delta g^{\mu \nu}} \equiv \frac{1}{2} \sqrt{-g} f_{\mu \nu}. \]  

(15.19)

Let us calculate these quantities for a point particle. For a specific world line \( q^\mu(\sigma) \) parameterized by an arbitrary timelike variable \( \sigma \), the action reads [recall (11.11), (11.12)]

\[ \mathcal{A} = -mc \int ds = -mc^2 \int d\sigma \sqrt{g_{\mu \nu}(q(\sigma)) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma)} \]

\[ = -mc\sqrt{-g} \int d\sigma \int d^4x \sqrt{-g} g_{\mu \nu}(x) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma) \delta^{(4)}(x - q(\sigma)). \]  

(15.20)

Variation with respect to \( g_{\mu \nu}(x) \) and \( K_{\mu \nu}^\lambda(x) \) gives

\[ \frac{\delta \mathcal{A}}{\delta g_{\mu \nu}(x)} \equiv -\frac{1}{2} \sqrt{-g} mc \int d\sigma \frac{1}{\sqrt{g_{\mu \nu}(q(\sigma)) \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma)}} \dot{q}^\mu(\sigma) \dot{q}^\nu(\sigma) \delta^{(4)}(x - q(\sigma)) \]  

(15.21)

\[ = -\frac{1}{2} \sqrt{-g} m \int d\tau \dot{q}^\mu(\tau) \dot{q}^\nu(\tau) \delta^{(4)}(x - q(\tau)), \]

where \( \tau = s/c \) is the proper time (11.4). The functional derivative with respect to \( K_{\mu \nu}^\lambda(x) \) vanishes identically:

\[ \frac{\delta \mathcal{A}}{\delta K_{\mu \nu}^\lambda(x)} \equiv 0. \]  

(15.22)

Thus we identify energy-momentum tensor and spin current densities

\[ \mathcal{T}^{\mu \nu}(x) \equiv m \int d\tau \dot{q}^\mu(\tau) \dot{q}^\nu(\tau) \delta^{(4)}(x - q(\tau)), \]  

(15.23)

\[ \Sigma^\nu_{\lambda \mu}(x) \equiv 0. \]  

(15.24)

We now determine these quantities for the gravitational field in the action (15.8). First we perform the variation of \( \sqrt{-g} \) with respect to \( \delta g_{\mu \nu} \). For this we write

\[ \delta \sqrt{-g} = -\frac{1}{2\sqrt{-g}} \delta g, \]  

(15.25)

and recall Eq. (11A.26) to express this as

\[ \delta \sqrt{-g} = \frac{1}{2} \sqrt{-g} g^{\mu \nu} \delta g_{\mu \nu} = -\frac{1}{2} \sqrt{-g} g_{\mu \nu} \delta g^{\mu \nu}. \]  

(15.26)

After writing the action (15.8) as

\[ \mathcal{A} = -\frac{1}{2} \int d^4x \sqrt{-g} g^{\mu \nu} R_{\mu \nu}, \]

H. Kleinert, GRAVITY WITH TORSION
we find
\[
\delta f_A = -\frac{1}{2} \int d^4x \sqrt{-g} \left\{ -\frac{1}{2} g_{\mu\nu} \delta g^{\mu\nu} + \delta g^{\mu\nu} R_{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} \right\}
\]
\[
= -\frac{1}{2} \int d^4x \sqrt{-g} \left[ \delta g^{\mu\nu} (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) + g^{\mu\nu} \delta R_{\mu\nu} \right].
\]
(15.27)
The factor accompanying $\delta g^{\mu\nu}$ is known as the Einstein tensor
\[
G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R.
\]
(15.28)
Note that this tensor is symmetric only in symmetric spaces. The variation in $\delta g^{\mu\nu}$, however, picks out only the symmetrized part of it.

Consider now the variation of the Ricci tensor in (15.27)
\[
\delta R_{\mu\nu} = \partial_\kappa \delta \Gamma^\kappa_{\mu\nu} - \partial_\mu \delta \Gamma_{\kappa\nu}^\kappa - \delta \Gamma^\kappa_{\mu\nu} \Gamma^\tau_{\kappa\tau} - \Gamma^\kappa_{\mu\nu} \delta \Gamma^\tau_{\kappa\tau} + \delta \Gamma^\tau_{\mu\nu} \Gamma^\kappa_{\tau\kappa} + \Gamma^\kappa_{\mu\nu} \delta \Gamma^\tau_{\kappa\tau}.
\]
(15.29)
The left-hand side is a tensor. It is then useful to express also the right-hand side via covariant forms. For this we observe that contrary to the affine connection $\Gamma_{\mu\nu}^\kappa$ itself, the variation $\delta \Gamma_{\mu\nu}^\kappa$ is a tensor.\(^1\) This follows directly from the transformation law (11.105). The last, non-holonomic piece $\partial_\mu \partial_\nu \xi^\kappa$ in $\Gamma_{\mu\nu}^\kappa$ disappear in $\delta \Gamma_{\mu\nu}^\kappa$ since it is the same for $\Gamma_{\mu\nu}^\kappa$ and $\Gamma_{\mu\nu}^\kappa + \delta \Gamma_{\mu\nu}^\kappa$. We therefore rewrite (15.29) in terms of covariant derivatives as
\[
\delta R_{\mu\nu} = D_\kappa \delta \Gamma^\kappa_{\mu\nu} - D_\mu \delta \Gamma_{\kappa\nu}^\kappa + 2 S_{\kappa\mu\tau} \delta \Gamma^\tau_{\tau\nu}^\kappa.
\]
(15.30)
Indeed, this is equal to
\[
\delta R_{\mu\nu} = -\partial_\kappa \delta \Gamma^\kappa_{\mu\nu} + \delta \Gamma^\kappa_{\mu\nu} - \Gamma^\tau_{\kappa\mu} \delta \Gamma^\tau_{\kappa\nu} - \Gamma^\tau_{\kappa\nu} \delta \Gamma^\tau_{\kappa\mu} + \Gamma^\kappa_{\mu\nu} \delta \Gamma^\tau_{\tau\nu} - \Gamma^\kappa_{\mu\nu} \delta \Gamma^\tau_{\tau\nu} + \Gamma^\tau_{\mu\nu} \delta \Gamma^\kappa_{\tau\kappa} + \Gamma^\tau_{\mu\nu} \delta \Gamma^\kappa_{\tau\kappa} + 2 S_{\kappa\mu\tau} \delta \Gamma^\tau_{\tau\nu}^\kappa,
\]
(15.31)
which is the same as (15.29). In symmetric spaces, the covariant relation (15.29) was first used by Palatini.

We now have to express $\delta R_{\mu\nu}$ in terms of $\delta g^{\mu\nu}$ and $\delta K_{\mu\nu}^\lambda$. It is useful to perform all operations underneath the integral in (15.27):
\[
-\frac{1}{2} \int d^4x \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu}.
\]
(15.32)
Due to the tensor nature of $\delta \Gamma_{\mu\nu}^\kappa$ we can take $g^{\mu\nu}$ through the covariant derivative and write (15.32) as
\[
-\frac{1}{2} \int d^4x \sqrt{-g} \left( \partial_\kappa \delta \Gamma^\kappa_{\mu\nu} - D_\mu \delta \Gamma^\kappa_{\kappa\nu} + 2 S_{\kappa\mu\tau} \delta \Gamma^\tau_{\kappa\nu} \right).
\]
(15.33)
\(^1\)In contrast, the difference $\delta \Gamma_{\mu\nu\kappa} \equiv \Gamma_{\mu\nu\kappa}^{g+\delta g}$ now is not a tensor.
The covariant derivatives can now be removed by a partial integration. In a space with torsion, partial integration has some particular features which requires a special discussion.

Take any tensors $U^{\mu...\nu}, V_{...\nu...}$ and consider an invariant volume integral

$$
\int d^4x \sqrt{-g} U^{\mu...\nu...} D_\mu V_{...\nu...}. \tag{15.34}
$$

A partial integration gives

$$
- \int d^4x \left[ \partial_\mu \sqrt{-g} U^{\mu...\nu...} V_{...\nu...} + \sum_i U^{\mu...\nu_i...} \Gamma_{\mu\nu_i}^{\lambda_i} V_{...\lambda_i...} \right] + \text{surface terms}, \tag{15.35}
$$

where the sum over $i$ runs over all indices of $V_{...\lambda_i...}$, linking them via the affine connection with the corresponding indices of $U^{\mu...\nu_i...}$. Now we use the relation

$$
\partial_\mu \sqrt{-g} = \sqrt{-g} \bar{\Gamma}_\mu^\kappa = \sqrt{-g} \Gamma_\mu^\kappa = \sqrt{-g} \left( 2S_\mu + \Gamma_\kappa^\mu \right) \tag{15.36}
$$

and (15.35) becomes

$$
- \int d^4x \sqrt{-g} \left[ \left( \partial_\mu U^{\mu...\lambda_i...} - \Gamma_{\mu\kappa}^{\lambda_i} U^{\mu...\lambda_i...} + \sum_i \Gamma_{\mu\nu_i}^{\lambda} U^{\mu...\nu_i...} \right) V_{...\lambda_i...}

+ 2S_\mu \sum_i U^{\mu...\lambda_i...} V_{...\lambda_i...} \right] + \text{surface terms}. \tag{15.37}
$$

Now, the terms in parentheses are just the covariant derivative of $U^{\mu...\nu_i...}$ such that we arrive at the rule

$$
\int d^4x \sqrt{-g} U^{\mu...\nu...} D_\mu V_{...\nu...} = - \int d^4x \sqrt{-g} D^*_\mu U^{\mu...\nu...} V_{...\nu...} + \text{surface terms}, \tag{15.38}
$$

where $D^*_\mu$ is defined as

$$
D^*_\mu \equiv D_\mu + 2S_\mu, \tag{15.39}
$$

where we have abbreviated:

$$
S_\kappa \equiv S_{\kappa\lambda}^{\lambda}, \quad S^\kappa \equiv S_{\lambda}^{\lambda \lambda}. \tag{15.40}
$$

It is easy to show that the operators $D_\mu$ and $D^*_\mu$ can be interchanged, i.e., there is also the rule

$$
\int d^4x \sqrt{-g} V_{...\nu...} D_\mu U^{\mu...\nu...} = - \int d^4x \sqrt{-g} U^{\mu...\nu...} D^*_\mu V_{...\nu...} + \text{surface terms}. \tag{15.41}
$$

For the particular case that $V_{...\nu...}$ is equal to 1, the second rule yields

$$
\int d^4x \sqrt{-g} D_\mu U^\mu = - \int d^4x \sqrt{-g} 2S_\mu U^\mu + \text{surface terms}. \tag{15.42}
$$
With the help of the last rule we can replace the covariant derivatives $D_\mu$ in Eq. (15.33) by $-2S_\mu$, and obtain
\[
\frac{1}{2} \int d^4 x \sqrt{-g} g^{\mu \nu} \delta R_{\mu \nu} = \frac{1}{2} \int d^4 x \sqrt{-g} \left(-2S_\kappa \delta \Gamma^{\nu \kappa \nu} + 2S_\mu \delta \Gamma^{\mu \kappa}_\kappa + 2S_\nu^{\nu \tau} \delta \Gamma^{\tau \nu}_\nu \right).
\]
(15.43)
The result can also be stated as follows:
\[
\frac{1}{2} \int d^4 x \sqrt{-g} g^{\mu \nu} \delta R_{\mu \nu} = \frac{1}{2} \int d^4 x \sqrt{-g} S_\mu^{\nu \tau} \delta \Gamma^{\nu \tau}_\mu
\]
where $S_\mu^{\nu \tau}$ is the following combination of torsion tensors:
\[
\frac{1}{2} S_\mu^{\nu \tau} \equiv S_\mu^{\nu \tau} + \delta^{\nu \tau}_\mu S_\kappa - \delta^{\tau \kappa}_\mu S_\nu.
\]
(15.44)
This tensor is referred to as the Palatini tensor. The relation can be inverted to
\[
S_{\mu \nu \lambda} = \frac{1}{2} \left( S_{\mu \nu \lambda} + \frac{1}{2} g_{\mu \lambda} S^{\nu \kappa} - \frac{1}{2} g_{\nu \lambda} S^{\mu \kappa} \right).
\]
(15.46)
We now proceed to express $\delta \Gamma^{\nu \tau}_\mu$ in terms of $\delta g_{\mu \nu}$ and $\delta K_{\mu \nu \lambda}$. For this purpose we note that the varied metric $g_{\mu \rho} + \delta g_{\mu \rho}$ certainly satisfies the identity (11.96),
\[
D^{\tau + \delta \Gamma}_{\tau} \left( g_{\mu \rho} + \delta g_{\mu \rho} \right) = 0,
\]
(15.47)
where $D^{\tau + \delta \Gamma}_{\tau}$ is the covariant derivative formed with the varied connection $\Gamma^{\lambda}_{\mu \nu} + \delta \Gamma^{\lambda}_{\mu \nu}$. For variations $\delta g_{\mu \rho}$ this implies
\[
D^{\tau}_{\tau} \delta g_{\mu \rho} = \delta \Gamma^{\tau}_{\mu \rho} + \delta \Gamma^{\tau}_{\tau \rho},
\]
(15.48)
where we have introduced
\[
\delta \Gamma^{\mu \tau \rho} \equiv g_{\rho \lambda} \delta \Gamma^{\lambda}_{\mu \tau}.
\]
(15.49)
This gives
\[
\frac{1}{2} \left( \delta D_{\tau} g_{\mu \rho} + \delta D_{\mu} g_{\tau \rho} - \delta D_{\rho} g_{\tau \mu} \right) = \delta \Gamma^{\tau}_{\mu \rho} - \delta S^{\tau}_{\tau \mu \rho} + \delta S^{\tau}_{\mu \rho \tau} - \delta S^{\tau}_{\rho \tau \mu} = \delta \Gamma^{\tau}_{\mu \rho} - \delta K_{\tau \mu \rho},
\]
(15.50)
where
\[
\delta S^{\tau}_{\tau \mu \rho} \equiv g_{\rho \lambda} \delta S^{\lambda}_{\tau \mu} \equiv g_{\rho \lambda} \frac{1}{2} \left( \Gamma^{\lambda}_{\tau \mu} - \Gamma^{\lambda}_{\mu \tau} \right)
\]
(15.51)
and
\[
\delta K_{\tau \mu \rho} \equiv \delta S^{\tau}_{\tau \mu \rho} - \delta S^{\tau}_{\mu \rho \tau} - \delta S^{\tau}_{\rho \tau \mu} + \delta S^{\tau}_{\rho \tau \mu},
\]
(15.52)
are the results of a variation of \( S_{\mu\nu}^\lambda \) at fixed \( g_{\mu\nu} \). Note that even though \( \bar{\Gamma}_{\mu\nu}^\lambda = \Gamma_{\mu\nu}^\lambda - K_{\mu\nu}^\lambda \), the left-hand side of (15.50) cannot be identified with \( g_{\rho\kappa} \delta \bar{\Gamma}_{\tau\mu}^\kappa \) since \( \delta K_{\mu\nu}^\lambda \) contains contribution from \( \delta S_{\mu\nu}^\lambda \) at fixed \( \delta g_{\mu\nu} \) and from \( \delta g_{\mu\nu} \) at fixed \( S_{\mu\nu}^\lambda \).

The first term in (15.50) is, in fact, equal to \( g_{\rho\kappa} \delta \bar{\Gamma}_{\tau\mu}^\kappa + \frac{1}{2} \left( D_\tau \delta g_{\mu\kappa} + D_\mu \delta g_{\tau\kappa} - D_\kappa \delta g_{\tau\mu} \right) \). Using (15.50), we rewrite (15.44) as

\[
-\frac{1}{2} \int d^4x \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} = \frac{1}{2} \int d^4x \sqrt{-g} S^{\mu\nu;\tau} \left[ \delta K_{\tau\mu\kappa} + \frac{1}{2} \left( D_\tau \delta g_{\mu\kappa} + D_\mu \delta g_{\tau\kappa} - D_\kappa \delta g_{\tau\mu} \right) \right].
\]

The first term shows that the Palatini tensor \( S_{\rho\kappa;\tau}^\mu \) plays the role of the spin current of the gravitational field [recall the definition (15.17) up to a factor \( 1/\kappa \)]

\[
\Sigma_{\rho;\kappa;\tau} = -\frac{1}{\kappa} S_{\rho;\kappa;\tau}.
\]  

(15.54)

The second term can now be partially integrated, leading to

\[
\frac{1}{4} \int d^4x \sqrt{-g} \left\{ D_\mu^* S_{\mu\rho;\kappa} \delta g_{\rho\mu} + D_\rho^* S_{\rho\mu;\kappa} \delta g_{\mu\rho} - D_\kappa^* S_{\kappa\mu;\rho} \delta g_{\mu\kappa} \right\} + \text{surface term}.
\]

(15.55)

After relabeling the indices in (15.50), we arrive at the following variation of the action with respect to \( \delta g_{\mu\nu} \), using the identity \( \delta g^{\mu\nu} G_{\mu\nu} = -\delta g_{\mu\nu} G^{\mu\nu} \) following from (11A.25),

\[
-\frac{1}{2} \int d^4x \sqrt{-g} \left[ G^{\mu\nu} - \frac{1}{2} D_\lambda^* \left( S_{\mu\lambda;\nu} - S_{\nu\lambda;\mu} + S_{\lambda\mu;\nu} \right) \right] \delta g_{\mu\nu},
\]

so that the complete energy-momentum tensor of the field reads

\[
T^{\mu\nu} = -\frac{1}{\kappa} \left[ G^{\mu\nu} - \frac{1}{2} D_\lambda^* \left( S_{\mu\lambda;\nu} - S_{\nu\lambda;\mu} + S_{\lambda\mu;\nu} \right) \right].
\]

(15.57)

Actually, the variation \( \delta g^{\mu\nu} \) yields only the symmetrized part of \( T^{\mu\nu} \). This specification is, however, unnecessary. We shall demonstrate later that total angular momentum conservation [see Eq. (18.13)] makes \( T^{\mu\nu} \) symmetric as it stands (even though \( G^{\mu\nu} \) is not).

Thus we arrive at the following field equations

\[
-\kappa \Sigma^{\mu\kappa;\tau} = S_{\mu\kappa;\tau} = \kappa T^{m\mu\kappa;\tau},
\]

(15.58)

\[
-\kappa T^{\mu\nu} = G^{\mu\nu} - \frac{1}{2} D_\lambda^* \left( S_{\mu\lambda;\nu} - S_{\nu\lambda;\mu} + S_{\lambda\mu;\nu} \right) = \kappa T^{m\mu\nu},
\]

(15.59)

which for a set of spinless point particles reduce to

\[
S_{\mu\kappa;\tau} = 0,
\]

(15.60)

\[
G^{\mu\nu} = \kappa T^{m\mu\nu}.
\]

(15.61)
15.3 Total Energy-Momentum Tensor and Defect Density

In defect physics, the total energy-momentum tensor obtained in (15.57) has a direct physical interpretation. In three Euclidean dimensions, the linearized version of (15.57) reads

$$-\kappa T^{ij} = G_{ij} - \frac{1}{2} \partial_k \left( S_{ij,k} - S_{jk,i} - S_{ki,j} \right),$$

(15.62)

with the spin density (15.45)

$$\frac{1}{2} \kappa \hat{\Sigma}_{ij,k} = S_{ij,k} = S_{ijk} + \delta_{ik} S_j - \delta_{jk} S_i.$$

(15.63)

Let us insert the dislocation density according to

$$S_{ijk} = \frac{1}{2} \left( \partial_i \partial_j - \partial_j \partial_i \right) u_k = \frac{1}{2} \epsilon_{ijl} \alpha_{lk}. $$

(15.64)

Then the spin density reads

$$S_{ij,k} = \epsilon_{ijl} \alpha_{lk} + \delta_{ik} \epsilon_{jpl} \alpha_{lp} - \delta_{jk} \epsilon_{ipl} \alpha_{lp}.$$ 

(15.65)

Since both sides are antisymmetric in $ij$, we can contract them with $\epsilon_{ijn}$,

$$\epsilon_{ijn} S_{ij,k} = 2 \alpha_{nk} + \epsilon_{kjm} \epsilon_{jpl} \alpha_{lp} - \epsilon_{ikn} \epsilon_{ipl} \alpha_{lp} = 2 \alpha_{nk} - 2 \left( \delta_{kp} \delta_{nl} - \delta_{kl} \delta_{np} \right) \alpha_{lp} = 2 \alpha_{kn},$$

(15.66)

and see that $S_{ij,k}$ becomes simply

$$S_{ij,k} = \epsilon_{ijl} \alpha_{kl}.$$ 

(15.67)

Thus the spin density is equal to the dislocation density.

The spin density has a vanishing divergence

$$\partial_k S_{ij,k} = \epsilon_{ijl} \partial_k \alpha_{kl} = 0.$$ 

(15.68)

In terms of the derivatives of the displacement field $u_i(x)$, the spin density reads

$$S_{ij,k} = \epsilon_{ijl} \epsilon_{kmn} \partial_m \partial_n u_l.$$ 

(15.69)

In this expression, the conservation law (15.68) is trivially fulfilled.

Let us now form the three combinations of $ij, k$ appearing in (15.62)

$$\frac{1}{2} \left( S_{ij,k} - S_{jk,i} + S_{ki,j} \right) = \frac{1}{2} \left( \epsilon_{ijl} \alpha_{kl} - \epsilon_{jkl} \alpha_{ij} + \epsilon_{kul} \alpha_{jl} \right).$$ 

(15.70)

By contracting the identity

$$\epsilon_{ijl} \delta_{km} + \epsilon_{jkl} \delta_{im} + \epsilon_{kul} \delta_{jm} = \epsilon_{ijk} \delta_{lm}$$

(15.71)
with \( \alpha_{ml} \), we find

\[
\epsilon_{ijkl} \alpha_{kl} + \epsilon_{jklm} \alpha_{il} + \epsilon_{kilm} \alpha_{jl} = \epsilon_{ijkl} \alpha_{il}
\]

so that

\[
\frac{1}{2} \left( S_{ij,k} - S_{jk,i} + S_{ki,j} \right) = -\epsilon_{jkl} \alpha_{il} + \frac{1}{2} \epsilon_{ijkl} \alpha_{il}.
\]

The right-hand side is recognized to be

\[
\epsilon_{jkl} K_{li}
\]

where

\[
K_{ij} = -\alpha_{ij} + \frac{1}{2} \delta_{ij} K_{kk}
\]

is Nye’s contortion tensor. With this notation, equation (15.62) becomes

\[
-\kappa f T_{ij} = G_{ij} - \epsilon_{jhl} \partial_n K_{li}.
\]

Now we recall that the Einstein tensor \( G_{ij} \) for a metric \( g_{ij} = \delta_{ij} + \partial_i u_j + \delta_j u_i \) coincides with the disclination density \( \Theta_{ji} \). But then, comparison with Eq. (12.42) shows that the total energy-momentum tensor limes \(-\kappa\) is nothing but the total defect density \( \eta_{ij} \):

\[
-\kappa f T_{ij} = \eta_{ij}
\]

### Notes and References


Minimally Coupled Fields of Integer Spin

So far we have discussed the gravitational field interacting with classical relativistic massive point particles. If we want to include quantum effects, we must describe these particles by the Klein-Gordon equation (2.59). For a consistent interpretation of the negative-energy wave functions, this field has to be quantized, so that incoming negative-energy wave functions describe outgoing antiparticles (recall p. 55).

Photons are described by Maxwell’s equations [recall (1.19) and (2.86)]. For electrons, muons, and neutrinos, and other particles with spin 1/2, the fields follow Dirac’s equation.

All these equations can be coupled minimally to gravity by means of the multivalued mapping principle. We simply transform the flat-spacetime action to spacetimes with curvature and torsion by means of a multivalued coordinate transformation.

In this text we shall not discuss the quantum aspect of the relativistic fields and treat only their classical limit.

16.1 Scalar Fields in Riemann-Cartan Space

The action (2.25) of a charged scalar field in flat spacetime is transformed nonholonomically to a general affine spacetime. The partial derivative $\partial_a$ is equal via Eq. (14.1) to

$$\partial_a = e^\mu_a (x) \partial_\mu,$$

and the volume element $d^4x^a$ in flat spacetime becomes

$$d^4 x^a = d^4 x^\mu | \det e^a_\mu (x) |.$$

Since $e^\mu_a (x)$ is the square root of the metric $g_\mu\nu (x)$ by Eq. (11.39), the determinants are also related by a square root, so that (16.2) without the indices amounts to the replacement rule for the flat-spacetime volume:

$$d^4 x \to d^4 x \sqrt{-g}.$$ 

401
With these rules, the action (2.25) becomes
\[
\mathcal{A} = \int d^4x \sqrt{-g} \left[ \hbar^2 e^{\alpha\mu}(x) \partial_\mu \phi^*(x) e^{\alpha\nu}(x) \partial_\nu \phi(x) - M^2 c^2 \phi^*(x) \phi(x) \right].
\] (16.4)

This expression cannot yet be used for field-theoretic calculations since the fields \( e^{\alpha\nu}(x) \) are multivalued. We can, however, use Eq. (11.42) to rewrite the action as
\[
\mathcal{A} = \int d^4x \sqrt{-g} \left[ \bar{h}^2 \bar{g}^{\mu\nu}(x) \partial_\mu \phi^*(x) \partial_\nu \phi(x) - M^2 c^2 \phi^*(x) \phi(x) \right].
\] (16.5)

This expression contains only the single-valued metric tensor. The equation of motion can be derived most simply by applying an integration by parts to the gradient term. Ignoring a boundary term we obtain
\[
\mathcal{A} = \int d^4x \sqrt{-g} \left[ -\hbar^2 \phi^*(x) \Delta \phi(x) - M^2 c^2 \phi^*(x) \phi(x) \right],
\] (16.6)

where
\[
\Delta \equiv \frac{1}{\sqrt{-g}} \left( \partial_\mu \sqrt{-gg^{\mu\nu}} \partial_\nu \right)
\] (16.7)
is the well-known Laplace-Beltrami differential operator. From the action (16.6) we obtain directly the equation of motion as in (2.38):
\[
\frac{\delta \mathcal{A}}{\delta \phi^*(x)} = \int d^4x' \sqrt{-g'} \left[ -\hbar^2 \phi^*(x') \Delta' \phi(x') - M^2 c^2 \delta^{(4)}(x' - x) \phi(x') \right]
\]
\[
= (-\hbar^2 \Delta - M^2 c^2) \phi(x) = 0.
\] (16.8)

This equation of motion contains an important prediction. There is no extra \( R \)-term in the wave equation, which would be allowed by covariance. In many textbooks [1], the Klein-Gordon equation is therefore written as
\[
(-\hbar^2 \Delta - \xi \hbar^2 R - M^2 c^2) \phi(x) = 0,
\] (16.9)

with a parameter \( \xi \) for which several numbers have been proposed in the literature: \( 1/6, 1/12, 1/8 \). If present, the same \( R \)-term would of course appear in the nonrelativistic limit of (16.9). This is obtained by setting \( \phi(x) = e^{-iMc\hbar \psi(x)} \) and letting \( c \to \infty \). Assuming that \( g_{00} = 0 \) and choosing \( g_{00} = 1 \), this leads to the Schrödinger equation:
\[
\left( -\frac{1}{2M} \hbar^2 \Delta - \xi \hbar^2 R^s \right) \psi(x) = i\hbar \partial_t \psi(x),
\] (16.10)

where \( R^s \) is curvature scalar of space. On a sphere of radius \( r \) in \( D \) dimensions it is equal to \( (D - 1)(D - 2)/r^2 \).

The number \( \xi = (D - 2)/4(D - 1) \), makes the massless equation (16.9) conformally invariant in \( D \) spacetime dimensions [2], so that \( \xi = 1/6 \) is a preferred value in some theories. When DeWitt set up a time-sliced path integral in curved space, he found \( \xi = 1/6 \) from his particular slicing assumptions [3]. A slightly different slicing
16.1 Scalar Fields in Riemann-Cartan Space

led to $\xi = 1/12$ [4]. In more recent work, DeWitt prefers the value $\xi = 1/8$ [5]. The value $\xi = 0$ was deduced from the multivalued mapping principle in Ref. [6].

So far there is no direct experimental confirmation of this prediction. There is however, indirect evidence. To see this we must assume that the extra $R$-term is universal, i.e., that the number $\xi$ holds for all point particles and irrespective of the coordinates in which the Schrödinger equation is expressed. This is a reasonable physical assumption for otherwise each particle would carry two gravitational parameters, mass and something else. The point is now that the hydrogen atom in momentum space is equivalent to a particle on a sphere in four dimensions of radius $p_E = \sqrt{-2ME}$. The spectrum of the Schrödinger equation without an $R$-term would be given by the eigenvalue equation

$$p_{En}/Mc = \alpha, \quad n = 1, 2, 3, \ldots,$$  \hspace{1cm} (16.11)

where $\alpha \approx 1/137$ is the fine structure constant. This equation yields Rydberg levels

$$E = -\frac{Mc^2\alpha^2}{2n^2}. \hspace{1cm} (16.12)$$

If there was an $R$-term as in (16.10), this spectrum would be determined by

$$p_E(n + 6\xi)/Mc = \alpha, \quad n = 1, 2, 3, \ldots.$$ \hspace{1cm} (16.13)

The parameter $\xi$ would directly appear in the denominator of (16.12) as [7]

$$E = -\frac{Mc^2\alpha^2}{2(n + 6\xi)^2}. \hspace{1cm} (16.14)$$

Such a distortion of the Rydberg spectrum would certainly have been detected for all the above candidates for $\xi$.

It is sometimes useful to express the Laplace-Beltrami operator (16.7) in terms of the Riemann connection $\bar{\Gamma}_{\mu\lambda}^{\kappa}$. For this we use Eq. (11A.24) to calculate the derivative

$$\frac{1}{\sqrt{-g}} \left( \partial_{\mu} \sqrt{-g} \right) = \frac{1}{2} g^{\lambda\kappa} (\partial_{\mu} g_{\lambda\kappa}) = \bar{\Gamma}_{\mu\lambda}^{\kappa}. \hspace{1cm} (16.15)$$

By differentiating the inverse metric (11.42) we have, furthermore,

$$\partial_{\mu} g_{\mu\nu} = -\bar{\Gamma}_{\mu}^{\mu\nu} - \bar{\Gamma}_{\nu}^{\nu\mu}. \hspace{1cm} (16.16)$$

Using the chain rule of differentiation on the right-hand side of (16.7), we then find the alternative expression

$$\Delta = g^{\mu\nu} \partial_{\mu} \partial_{\nu} - \bar{\Gamma}_{\mu}^{\mu\nu} \partial_{\nu}. \hspace{1cm} (16.17)$$
16.2 Electromagnetism in Riemann-Cartan Space

Let us go through the same procedure for the electromagnetic action (16.20). The volume element is again mapped according to the rule (16.3). The covariant curl is treated as follows. First we introduce vector fields transforming like the coordinate differentials $dx_\mu$

$$A_\mu(x) = e^a_\mu(x)A_a(x), \quad j_\mu(x) = e^a_\mu(x)j_a(x), \quad (16.18)$$

and rewrite the field strengths with the help of (11.85) as

$$F_{ab}(x) = \partial_a A_b(x) - \partial_b A_a(x) = e^\mu_a(x)\partial_\mu e^\nu_b(x)A_\nu(x) - e^\mu_b(x)\partial_\mu e^\nu_a(x)A_\nu(x)$$

$$= e^\mu_a(x)e^\nu_b(x)\left[D_\mu A_\nu(x) - D_\nu A_\mu(x)\right] \equiv e^\mu_a(x)e^\nu_b(x)F_{\mu\nu}(x). \quad (16.19)$$

Then the action (16.20) becomes

$$A_{em} = \int d^4x \sqrt{-g} \, \mathcal{L}_{em}(x) = \int d^4x \sqrt{-g} \left[ -\frac{1}{4c}F_{\mu\nu}(x)F_{\mu\nu}(x) - \frac{1}{c^2}j^\mu(x)A_\mu(x) \right]. \quad (16.20)$$

Expressing the covariant derivatives as in (11.86), the field strength takes the form

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) - 2S_{\mu\nu\lambda}A_\lambda(x). \quad (16.21)$$

The last term destroys gauge invariance as noted first by Schrödinger [8] (see the remarks in Preface). This is why he derived upper bounds for the photon mass from experimental observations. The present upper bound is

$$m_\gamma < 3 \times 10^{-27} \text{eV}. \quad (16.22)$$

This estimate comes from observations of the range of magnetic fields emanating into spacetime from pulsars. The range corresponding to the above number is the Compton wavelength

$$l_\gamma = \frac{\hbar}{m_\gamma c} > 6952 \text{ light years}. \quad (16.23)$$

In order to be invariant under the usual electromagnetic gauge transformations

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda, \quad (16.24)$$

the electromagnetic action

$$A_{em} = -\frac{1}{4} \int d^4x \sqrt{-g} F_{\mu\nu}F^{\mu\nu}, \quad (16.25)$$

must contain the same field strengths in spacetimes with curvature and torsion as in flat spacetime:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (16.26)$$
This covariant curl is compatible with invariance under general coordinate transformation since it can be rewritten as a covariant curl involving Riemann covariant derivatives $\bar{D}_\nu$:

$$\bar{D}_\mu A_\nu - \bar{D}_\nu A_\mu = \partial_\mu A_\nu - \bar{\Gamma}^{\lambda}_{\mu\nu} A_\lambda - \partial_\nu A_\mu + \bar{\Gamma}^{\lambda}_{\nu\mu} A_\lambda = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (16.27)$$

Thus the field strength (25.45) is gauge invariant under electromagnetic and Einstein transformations. For this reason, many authors have postulated that the photon does not couple to torsion. Later we shall argue that this is not really necessary: torsion will be seen to be a nonpropagating field, so that empty space where photons are observed to propagate with light velocity cannot carry any torsion.

Notes and References


17

Particles with Half-Integer Spin

Let us now see how electrons and other particles of half-integer spin are coupled to gravity [1].

17.1 Local Lorentz Invariance and Anholonomic Coordinates

Spin is defined in Lorentz-invariant theories as the total angular momentum in the rest frame of the particle. To measure the spin $s$ of a particle moving with velocity $\mathbf{v}$, we go to a comoving frame by a local Lorentz transformation. Then its quantum mechanical description requires $2s + 1$ states $|s, s_3\rangle$ with $s_3 = -s, \ldots, s$ which, upon rotations, transform according to an irreducible representation of the rotation group with angular momentum $s$.

For a particle of spin $s = 1/2$ such as an electron, a muon, or any other massive lepton in Minkowski spacetime, this transformation property is automatically accounted for by the states that can be created by a quantized Dirac field $\psi_\alpha(x)$ with the action (2.140):

$$\mathcal{A} = \int d^4x \bar{\psi}(x^a) (i\gamma^a \partial_a - m) \psi(x^a),$$  

where the matrices $\gamma^a$ satisfy the Dirac algebra (1.218):

$$\{\gamma^a, \gamma^b\} = 2g^{ab}. \quad \text{(17.2)}$$

The Dirac equation is obtained by extremizing this action as in Eq. (2.142):

$$\frac{\delta \mathcal{A}}{\delta \psi(x^a)} = (i\gamma^a \partial_a - m) \psi(x^a) = 0.$$  

\[406\]
17.1 Local Lorentz Invariance and Anholonomic Coordinates

17.1.1 Nonholonomic Image of Dirac Action

By complete analogy with the treatment of the action of a scalar field in Section 16.1 we can immediately write down the action in a spacetime with curvature and torsion:

\[ m \mathcal{A} = \int d^4x \sqrt{-g} \bar{\psi}(x) \left[ i \gamma^a e_a^\mu(x) \partial_\mu - m \right] \psi(x), \quad (17.4) \]

where \( x \) are the physical coordinates \( x^\mu \). In contrast to the scalar case, however, this transformed action contains the multivalued tetrad fields for which the field-theoretic formalism is invalid. We must find a way of transforming away the multivalued content in \( e_a^\mu(x) \). This is done by the introduction, at each point \( x^\mu \), of infinitesimal coordinates \( dx^\alpha \) associated with a freely falling Lorentz frame. We may simply imagine infinitesimal freely falling elevators inside of which there is no gravity. The removal of the gravitational force holds only at the center of mass of a body. At any distance away from it there are tidal forces where either the centrifugal force or the gravitational attraction becomes dominant. At the center of mass, the coordinates \( dx^\alpha \) are Minkowskian, but the affine connections are nonzero and have in general a nonzero curvature which cause the tidal forces.

Intermediate Theory

We proceed as in Section 4.5 and observe that the modified Dirac Lagrangian density

\[ \bar{\psi}(x^\alpha) \left\{ i \gamma^\alpha \left[ \partial_\alpha - D(\Lambda^a(x^\alpha))^{-1} \partial_\alpha D(\Lambda(x^\alpha)) \right] - m \right\} \psi(x^\alpha) \quad (17.5) \]

describes electron just as well as the original one, where \( \Lambda(x^\alpha) \) is an arbitrary local set of Lorentz transformations which connects the flat-spacetime coordinates \( x^a \) in (17.1) with new coordinates \( x^\alpha \):

\[ dx^a = \Lambda^a_{\;\alpha}(x^\alpha)dx^\alpha, \quad dx^a = (\Lambda^{-1})^a_{\;\alpha}(x)dx^\alpha \equiv \Lambda^a_{\;\alpha}(x)dx^\alpha. \quad (17.6) \]

Here and in the sequel, we shall suppress the superscript of \( x \) in the arguments of \( \Lambda(x) \) whenever it is not necessary to be explicit. The same will be done for \( \psi(x) \). The metrics in the two coordinate systems are Minkowskian for any choice of \( \Lambda^a_{\;\alpha}(x)dx^\alpha \):

\[ g_{\alpha\beta}(x) = \Lambda^a_{\;\alpha}(x)\Lambda^b_{\;\beta}(x)g_{ab} = (\Lambda^T)_\alpha^\alpha(x) g_{ab} \Lambda^b_{\;\beta}(x) \equiv g_{ab}, \quad (17.7) \]

in accordance with Eq. (1.28).

The solutions of the associated Dirac equation \( \psi'(x) \) are obtained from the original solutions \( \psi(x) \) by the spinor representation of the local Lorentz transformation: \( \psi'(x) = D(\Lambda(x))\psi(x) \). This reflects the freedom of solving Dirac’s anticommutation rules (1.218) by the \( x \)-dependent \( \gamma \)-matrices \( D(\Lambda(x))^{-1}_\gamma\alpha D(\Lambda(x)) \) [recall (1.229) and (1.28)]:

\[ \{ D(\Lambda(x))^{-1}_\gamma\alpha D(\Lambda(x)), D(\Lambda(x))^{-1}_{\gamma\beta} D(\Lambda(x)) \} = \{ \gamma^a, \gamma^b \} \Lambda^a_{\;\alpha}(x)\Lambda^b_{\;\beta}(x) = g^{ab}\Lambda^a_{\;\alpha}(x)\Lambda^b_{\;\beta}(x) = g^{\alpha\beta}. \quad (17.8) \]
We now recall Eq. (1.254) according to which, in a slightly different notation,

\[ D(\Lambda(x))^{-1} \partial_\alpha D(\Lambda(x)) = -\frac{i}{2} \omega_{\alpha;\delta\sigma}(x) \left( \Sigma^{\delta\sigma} \right)_B^C. \]  

(17.9)

The right-hand side may be defined as the spin connection for Dirac fields:

\[ \Gamma^{\alpha B C}(x) \equiv \frac{1}{2} \omega_{\alpha;\delta\sigma}(x) \left( \Sigma^{\delta\sigma} \right)_B^C. \]  

(17.10)

Here \( \omega_{\alpha;\beta}^\gamma \) are the generalized angular velocities obtained by relations of the type (1.251) from the tensor parameters \( \omega_{\beta}^\gamma \) of the local Lorentz transformations \( \Lambda(x) = e^{-i\omega_{\beta}^\gamma(x)\Sigma^\beta\gamma} \).

According to Eq. (1.253), the generalized angular velocities \( \omega_{\alpha;\beta}^\gamma \) appear also in the derivatives of the local Lorentz matrices \( \Lambda^\alpha_a(x) \) as

\[ \Lambda^{-1\gamma}_a(x)\partial_\alpha \Lambda^a_\beta(x) = \omega_{\alpha;\beta}^\gamma(x) = -\omega_{\alpha;\beta}^\gamma(x). \]  

(17.11)

Thus, if we define

\[ \Lambda^{\alpha\beta}_a \equiv \Lambda^\alpha_a \partial_\alpha \Lambda^a_\beta = -\Lambda^a_\beta \partial_\alpha \Lambda^\alpha_a, \]  

(17.12)

we can write the Dirac spin connection as

\[ \Gamma^{\alpha B C}(x) \equiv -i \frac{1}{2} \Lambda^{\alpha\beta}_a \left( \Sigma^\delta\sigma \right)_B^C. \]  

(17.13)

The transformation has produced a Lagrangian density

\[ \mathcal{L} = \bar{\psi}(x) \left( i\gamma^\alpha D_\alpha - m \right) \psi(x), \]  

(17.14)

with the covariant derivative matrix

\[ (D_\alpha)_B^C = \delta_B^C \partial_\alpha - \Gamma^{\alpha B C}_a(x), \]  

(17.15)

which is completely equivalent to the original Dirac Lagrangian density in Eq. (17.4), as long as the spin connection is given by (17.12) with single-valued Lorentz transformations \( \Lambda^\alpha_a \). This is the analog of the Schrödinger Lagrangian (4.80) which was the starting point for the introduction of electromagnetism by multivalued gauge transformations.

We may now proceed in the same way as before by postulating the local parameters the tensor fields \( \omega_{\beta\gamma}(x) \) to be multivalued. Then the components of the spin connection are no longer generalized angular velocities (17.11) but new independent fields, which cannot be calculated from \( \omega_{\beta\gamma}(x) \) by gradient equations like (1.251). Then the Lagrangian density (17.14) with the covariant derivative (17.15) describes a nontrivial theory coupled to torsion. The affine connection (17.12) will be seen in Eq. (17.66) to coincide with the contortion in the locally Minkowskian coordinates \( dx^\alpha \). The intermediate spacetime \( dx^\alpha \) has no Riemannian curvature, so that the Riemann-Cartan curvature tensor is determined completely by the contortion
17.1 Local Lorentz Invariance and Anholonomic Coordinates

The covariant derivative formed with the Christoffel symbols allows for the definition of parallel vector fields over any distance. This theory is a counterpart of the famous teleparallel theory developed by Einstein since 1928 and influenced by a famous letter exchange with Cartan (recall Preface). There the situation is the opposite: the Riemann-Cartan curvature tensor vanishes identically, and the Riemann curvature is given via Eq. (11.145) by

\[- \bar{R}_{\mu \nu \lambda}^\kappa = \bar{D}_\mu K_{\nu \lambda}^\kappa - \bar{D}_\nu K_{\mu \lambda}^\kappa + \left( K_{\mu \lambda}^\rho K_{\nu \rho}^\kappa - K_{\nu \lambda}^\rho K_{\mu \rho}^\kappa \right), \quad (17.16)\]

17.1.2 Vierbein Fields

In order to describe the correct gravitational forces, we must go one step further. Following the standard procedure of Section 4.5 we first perform the analog of a single-valued gauge transformation which is here an ordinary coordinate transformation from \( x^\alpha \) to \( x^\mu \):

\[dx^\alpha = dx^\mu h_\alpha^\mu(x). \quad (17.17)\]

The transformation has an inverse

\[dx^\mu = dx^\alpha h_\alpha^\mu(x), \quad (17.18)\]

and the matrix elements \( h_\alpha^\mu(x) \) and \( h_\alpha^\mu(x) \) satisfy, at each \( x \), the orthonormality and completeness relations

\[h_\alpha^\mu(x)h_\beta^\mu(x) = \delta_\alpha^\beta, \quad h_\alpha^\mu(x)h_\alpha^\nu(x) = \delta_\mu^\nu. \quad (17.19)\]

The 4 \( \times \) 4 transformation matrices \( h_\alpha^\mu(x) \) and \( h_\alpha^\mu(x) \) are called vierbein fields and reciprocal vierbein fields, respectively. As in the case of the multivalued basis tetrads \( e^\alpha_a(x), e_\alpha^\mu(x) \) we shall freely raise and lower the indices \( \alpha, \beta, \gamma, \ldots \) with the metric \( g_{\alpha \beta} \) and its inverse \( g^{\alpha \beta} \):

\[h^{\alpha \mu}(x) \equiv g^{\alpha \beta} h_\beta^\mu(x), \quad h_{\alpha \beta}(x) \equiv g_{\alpha \beta} h_\beta^\mu(x). \quad (17.20)\]

Since the transformation functions are, for the moment, single-valued, they satisfy

\[\partial_\mu h_\alpha^\nu(x) - \partial_\nu h_\alpha^\mu(x) = 0, \quad \partial_\mu h_\alpha^\nu(x) - \partial_\nu h_\alpha^\mu(x) = 0. \quad (17.21)\]

For spinor fields as functions depending on the final, physical coordinates \( x^\mu \), the covariant derivative (17.15) becomes

\[(D_\alpha)_B^C = \delta_B^C h_\alpha^\mu(x)\partial_\mu - \bar{D}_{\alpha \beta}^C(x). \quad (17.22)\]

The flat spacetime \( x^a \) coordinates are now related to the physical coordinates \( x^\mu \) by the equation

\[dx^a = \Lambda_a^\alpha(x)dx^\alpha = \Lambda_a^\alpha(x)h_\alpha^\mu(x)dx^\mu, \quad (17.23)\]

where the Lorentz transformation \( \Lambda(x) \) are multivalued, but the action

\[\tilde{\mathcal{A}} = \int d^4x\sqrt{-g} \bar{\psi}(x) \left(i\gamma^\alpha h_\alpha^\mu(x)D_\mu - m\right) \psi(x), \quad (17.24)\]

contains only the single-valued geometric fields \( h_\alpha^\mu(x) \) and \( A_{\alpha \beta \sigma}(x) \).
17.1.3 Local Inertial Frames

This is now the place where we can easily introduce curvature by allowing the coordinates $x^\alpha$ to be multivalued functions of the physical coordinates of $x^\mu$. Then the vierbein fields satisfy no longer the relation (17.21). Since they, themselves, are single-valued functions the action is perfectly suited to describe electrons and other elementary spin-1/2 Dirac particles in spacetimes with curvature and torsion.

Since $dx^\alpha$ is related to $dx^\mu$ by a Lorentz transformation (17.23), the length of $dx^\alpha$ is measured by of the nonholonomic coordinates $dx^\alpha$ is Minkowskian at the point $x$:

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta,$$

(17.25)

where

$$g_{\alpha\beta} = \Lambda^a_\alpha(x)\Lambda^b_\beta(x)g_{ab} \equiv \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}_{\alpha\beta}$$

(17.26)

is the Minkowski metric, due to (1.28).

Combining (17.25) with (17.17), we see that the vierbein fields $h_{\alpha}^\mu(x)$ transform the metric $g_{\mu\nu}(x)$ to the constant Minkowski metric

$$g_{\alpha\beta} = h_{\alpha}^\mu(x)h_{\beta}^\nu(x)g_{\mu\nu}(x)$$

(17.27)

at each $x$. The inverse of this relation shows that the metric $g_{\mu\nu}(x)$ is the square of the matrices $h_{\alpha}^\mu(x)$,

$$g_{\mu\nu}(x) = h_{\alpha}^\mu(x)h_{\beta}^\nu(x)g_{\alpha\beta} \equiv h_{\alpha}^\mu(x)h_{\beta}^\nu(x),$$

(17.28)

just as it was the square of the multivalued basis tetrads $e_{a}^\nu(x)$ in Eq. (11.39).

There is a simple physical relation between the physical coordinates $x^\mu$, and the infinitesimal coordinates $dx^\alpha$. The latter are associated with small freely falling Lorentz frames at each $x^\mu$. Such frames are also called inertial frames. They may be imagined as small freely falling elevators in which there is no gravity. The removal of the gravitational force holds only at the center of mass of the elevators. At any distance away from it there are tidal forces where either the centrifugal force or the gravitational attraction becomes dominant. Let us verify this explicitly. In a small neighborhood of an arbitrary point $X^\mu$ we solve the differential equation (17.18) by the functions

$$x^\alpha(X; x) = a^\alpha + h_{\mu}^\alpha(X)(x^\mu - X^\mu) + \frac{1}{2}h_{\lambda}^\alpha(X)\Gamma^\lambda_{\mu\nu}(X)(x^\mu - X^\mu)(x^\nu - X^\nu) + \ldots$$

(17.29)

The derivatives

$$\frac{\partial x^\alpha(X; x)}{\partial x^\mu} = h_{\mu}^\alpha(X) + h_{\lambda}^\alpha(X)\Gamma^\lambda_{\mu\nu}(X)(x^\nu - X^\nu) \equiv h(X; x)$$

(17.30)
fulfill (17.18) at \( x = X \). Consider now a point particle satisfying the equation of motion (14.7). In the coordinates (17.29), the trajectory satisfies the equation
\[
\dot{q}^\alpha = h^\alpha_{\mu}(X)\dot{q}^\mu + h^\alpha_{\lambda}(X)\Gamma^\lambda_{\mu\nu}(X)\dot{q}^\mu(q^\nu - X^\nu) + \ldots,
\]
and
\[
\ddot{q}^\alpha = h^\alpha_{\mu}(X)\dot{q}^\mu + h^\alpha_{\lambda}(X)\Gamma^\lambda_{\mu\nu}(X)\ddot{q}^\mu(q^\nu - X^\nu) + h^\alpha_{\lambda}(X)\Gamma^\lambda_{\mu\nu}(X)\dot{q}^\mu\dot{q}^\nu + \ldots.
\]
Inserting here (14.7), the first and third terms cancel each other, and the trajectory experiences no acceleration at the point \( X \). In the neighborhood, there are tidal forces. Thus the infinitesimal constitute an inertial frame in an infinitesimal neighborhood of the point \( X \).

The metric in the coordinates \( x^\alpha(X; x) \) is
\[
g_{\alpha\beta}(X; x) = \frac{\partial x^\alpha(X; x)}{\partial x^\mu} \frac{\partial x^\beta(X; x)}{\partial x^\nu}g^{\mu\nu}(x) = h^\alpha_{\mu}(X)h^\beta_{\nu}(X)g^{\mu\nu}(x)
+ \left[ h^\alpha_{\lambda}(X)h^\beta_{\nu}(X)\Gamma^\lambda_{\mu\kappa}(X)(x^\kappa - X^\kappa) + (\alpha \leftrightarrow \beta) \right] g^{\mu\nu}(x).
\]
We now expand the metric \( g^{\mu\nu}(x) \) in the neighborhood of \( X \) as
\[
g^{\mu\nu}(x) = g^{\mu\nu}(X) + \partial_\lambda g^{\mu\nu}(X)(x^\lambda - X^\lambda) + \ldots
= g^{\mu\nu}(X) - \left[ g^{\mu\lambda}\Gamma^\lambda_{\kappa\nu}(X) + g^{\kappa\lambda}\Gamma^\lambda_{\mu\nu}(X) \right] (x^\kappa - X^\kappa) + \ldots.
\]
Inserting this into (17.33) and using (17.27), we obtain
\[
g_{\alpha\beta}(X; x) = g^{\alpha\beta} + \mathcal{O}(x - X)^2.
\]
This ensures, that the affine connection formed from \( g_{\alpha\beta}(X; x) \) vanished at \( x = X \), so that there are no forces at this point. In any neighborhood of \( X \), however, there will be tidal forces.

In the coordinates \( dx^\alpha \), there are no tidal forces at all. This is possible everywhere only due to defects which make \( \Lambda^\alpha_\alpha(x) \) multivalued.

The coordinates \( x^\alpha(X; x) \) are functions of \( x \) which depend on \( X \). There exists no single function \( x^\alpha(x) \), so that derivatives in front of \( x^\alpha(x) \) do not commute:
\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) x^\alpha(x) \neq 0,
\]
implying that
\[
\partial_\mu h^\alpha_{\nu}(x) - \partial_\nu h^\alpha_{\mu}(x) \neq 0.
\]
The functions \( h^\alpha_{\mu}(x) \) and \( h^\alpha_{\nu}(x) \), however, which describe the transformation to the freely falling elevators are single-valued. They obey the integrability condition
\[
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) h^\alpha_{\lambda}(x) = 0,
\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) h^\lambda_{\alpha}(x) = 0.
\]
This condition has the consequence that if we construct a tensor \( \tilde{R}_{\mu\nu\lambda}^\kappa(x) \) from the transformation matrices \( h^\mu_{\alpha}(x) \) in the same way as \( R_{\mu\nu\lambda}^\kappa(x) \) was made from \( e_{\alpha\nu}(x) \) in Eq. (11.129), we find an identically vanishing result:
\[
\tilde{R}_{\mu\nu\lambda}^\kappa = h^\alpha_{\kappa}(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu) h^\alpha_{\lambda} \equiv 0.
\]
17.1.4 Difference between Vierbein and Multivalued Tetrads Fields

Note that \( h^\alpha_\mu(x) \) and \( e^a_\mu(x) \) are completely different mathematical objects with different integrability properties. While \( h^\alpha_\lambda(x) \) satisfies (17.38), \( e^a_\chi(x) \) does not, since the commutator of the derivatives in front of \( e^a_\lambda(x) \) determine the curvature tensor via Eq. (11.129).

Both, the multivalued basis tetrads \( e^a_\mu(x) \) and the vierbein fields \( h^\alpha_\mu(x) \) are “square roots” of the metric \( g_{\mu\nu}(x) \). They differ by a local Lorentz transformation \( \Lambda^\mu_\alpha(x) \). This is precisely the freedom one has in defining such a “square root”. By introducing Lorentz transformations with rotational defects \( dx^\alpha = dx^a\Lambda^\alpha_a(x) \), where \( \Lambda^\alpha_a(x) \) are non-integrable functions of \( x \), we have achieved that the vierbein fields have commuting derivatives [recall (17.38)].

From Eq. (17.23) we see that the relation between the vierbein and multivalued basis tetrad fields is given by

\[
e^a_\mu(x) = \Lambda^a_\alpha(x)h^\alpha_\mu(x), \tag{17.40}
\]

whose inverse is

\[
\Lambda^a_\alpha(x) \equiv e^a_\mu(x)h^\alpha_\mu(x), \tag{17.41}
\]

Since \( h^\alpha_\mu \) and \( h^a_\mu \) are single-valued functions with commuting derivatives, the curvature tensor \( R_{\mu\nu\lambda}^\kappa \) in (11.129) may be expressed completely in terms of the noncommuting derivatives of the local Lorentz transformations \( \Lambda^\alpha_a(x) \) and \( \Lambda^a_\alpha(x) \). To see this we insert Eq. (17.40) into (11.129), and use (17.38) to find for the curvature tensor the alternative expression

\[
R_{\mu\nu\lambda}^\sigma = h^\sigma_\gamma \left[ \Lambda^\gamma_\alpha \left( \partial^\gamma_\mu \partial^\nu_\rho - \partial^\nu_\rho \partial^\mu_\gamma \right) \Lambda^\alpha_\lambda \right] h^\alpha_\lambda \equiv h^\gamma_\gamma R_{\mu\rho\lambda}^\gamma h^\alpha_\lambda. \tag{17.42}
\]

From the defect point of view, the single-valued matrices \( h^\alpha_\mu(x) \) create an intermediate coordinate system \( dx^\alpha \) which, by the integrability condition (17.38), has the same disclination content as the coordinates \( x^\mu \), but is completely free of dislocations. The metric in the new coordinate system \( x^\alpha \) is Minkowski-like at each point in spacetime. Still, the coordinates \( x^\alpha \) do not form a Minkowski spacetime since they differ from the inertial coordinates \( dx^\alpha \) by the presence of disclinations, i.e., there are wedge-like pieces missing with respect to an ideal reference crystal. The coordinates \( x^\alpha \) cannot be defined globally from \( x^\mu \). Only the differentials \( dx^\alpha \) are uniquely related to \( dx^\mu \) at each spacetime point by Eqs. (17.17) and (17.18). The local Lorentz transformations \( \Lambda^a_\alpha(x) \) have noncommuting derivatives on account of the disclinations residing in the coordinates \( dx^\alpha \). The coordinate system \( dx^\alpha \) can only be used to specify derivatives with respect to \( x^\alpha \), or the directions of vectors (and tensors) with respect to the intermediate local axes

\[
e_\alpha(x) \equiv e_\mu(x)\frac{\partial x^\mu}{\partial x^\alpha} = e_\mu(x)h^\mu_\alpha(x) \\
\equiv e_\alpha e^a_\mu(x)h^\mu_\alpha(x) \equiv e_\alpha \Lambda^a_\alpha(x). \tag{17.43}
\]
17.1 Local Lorentz Invariance and Anholonomic Coordinates

We can go back to the local basis via the reciprocal vierbein fields

\[ e_\mu(x) = e_\alpha(x) \partial_{x\alpha} = e_\alpha(x) h^{\alpha}_{\mu}(x). \]  

(17.44)

Thus, an arbitrary vector may be transformed as follows,

\[ v(x) \equiv e_\alpha v^\alpha(x) = e_\alpha \Lambda^\alpha_{\alpha}(x) h^{\alpha}_{\mu}(x) v^\mu(x) = e_\alpha \Lambda^\alpha_{\alpha}(x) v^\alpha(x) = e_\alpha \Lambda_{\alpha \alpha}(x) v_\alpha(x), \]  

(17.45)

where we have introduced the co- and contravariant components

\[ v_\alpha(x) \equiv v_\mu(x) h^{\mu}_{\alpha}(x), \quad v^\alpha(x) \equiv v^\mu(x) h^{\alpha}_{\mu}(x). \]  

(17.46)

The orthogonality relations (17.19) imply the inverse relations

\[ v_\mu(x) = v_\alpha(x) h^{\alpha}_{\mu}(x), \quad v^\mu(x) = v^\alpha(x) h^{\alpha}_{\mu}(x). \]  

(17.47)

For vector fields \( v_\beta, v^\beta \) whose components refer to the intermediate basis \( e_\alpha(x) \), the covariant derivatives are

\[ D_\alpha v_\beta = \partial_\alpha v_\beta - \Lambda^\alpha_{\alpha \beta} \Gamma^\gamma_{\alpha \gamma} v_\gamma, \quad D_\alpha v^\beta = \partial_\alpha v^\beta + \Lambda^\alpha_{\alpha \beta} \Gamma^\gamma_{\alpha \gamma} v_\gamma, \]  

(17.48)

where \( \Gamma^\alpha_{\alpha \beta} \) is the spin connection for vector fields to be calculated from the local multivalued Lorentz transformations \( \Lambda^\alpha_{\beta}(x) \) of Eq. (17.41) rather than from \( e_\alpha^\mu \) [compare also (11.92)]:

\[ \Gamma^\alpha_{\alpha \beta}(x) = \Lambda^\alpha_{\alpha \beta}(x) h^{\alpha}_{\mu}(x) \equiv \Lambda^\alpha_{\alpha \beta}(x) \Gamma^\gamma_{\alpha \gamma}. \]  

(17.49)

Thus, the spin connection for local Lorentz vector fields is precisely the object found in the covariant derivatives of the spinor fields (17.22).

By analogy with the pure gradient (4.50) of a single-valued gauge function in the abelian gauge theory of magnetism, the spin connection (17.49) reduces to a trivial gauge field for single-valued local Lorentz transformations \( \Lambda(x) \). Indeed, it is easily verified that the field strength associated with this gauge field, the covariant curl

\[ F_{\mu \nu \alpha \beta} \equiv \partial_\mu \Gamma^\alpha_{\nu \alpha \beta} - \partial_\nu \Gamma^\alpha_{\mu \alpha \beta} - [\Gamma^\alpha_{\mu \alpha \beta}, \Gamma^\gamma_{\nu \gamma}], \]  

(17.50)

vanishes. As in Eqs. (11.128), (11.125), the commutator is defined by considering \( \Gamma^\alpha_{\nu \alpha} \) as matrices \( (\Gamma^\alpha_{\nu})_\alpha \). For multivalued Lorentz transformations, the covariant curl is in general nonzero, and \( \Gamma^\alpha_{\nu \alpha} \) is a nonabelian gauge field, whose field strength \( F_{\mu \nu \alpha \beta} \equiv \) is a tensor under single-valued local Lorentz transformations.

For multivalued transformations \( \Lambda(x) \), the a nonabelian gauge field (17.49) has nonzero field strengths (17.50).
From Eq. (17.49) it follows that \( \Lambda^a \alpha (x) \) and \( \Lambda_\alpha ^a (x) \) satisfy identities like (11.93), (11.94):

\[
D_a \Lambda^a = 0, \quad D_\alpha \Lambda_\alpha = 0.
\] (17.51)

It is instructive to rewrite the spin connection in matrix notation using the notation (17.17) for the local Lorentz transformations. Then the relation (17.45) shows that

\[
v^\alpha (x) = \Lambda^a \alpha v^a (x), \quad v_\alpha (x) = \Lambda_\alpha ^a v^a (x) = (g \Lambda g)_a \alpha v^a (x) = \left( \Lambda^T \right)_a ^a v^a (x),
\] (17.52)

and therefore

\[
\partial_\alpha v^\alpha (x) = \Lambda^a \beta D_a v^\beta (x) = \Lambda^a \beta \left[ -\partial_\alpha \delta^\beta _\alpha + (\Lambda^{-1} \partial_\alpha \Lambda)^\beta _\gamma \right] v^\gamma (x).
\] (17.53)

\[
\partial_\alpha v_\alpha (x) = \Lambda_\alpha ^a D_a v^\beta (x) = \Lambda_\alpha ^a \left[ \partial_\alpha \delta^\beta _\gamma + (\Lambda^T \partial_\alpha \Lambda^{-1})^\beta _\gamma \right] v^\gamma (x)
= \Lambda_\alpha ^a \left[ \partial_\alpha \delta^\beta _\gamma - (\Lambda^{-1} \partial_\alpha \Lambda)^\gamma _\beta \right] v^\gamma (x).
\] (17.54)

From this we identify

\[
\Gamma^\alpha _{\alpha \beta} = (\Lambda^{-1} \partial_\alpha \Lambda)^\beta _\gamma = - (\Lambda^{-1} \partial_\alpha \Lambda)^\gamma _\beta,
\] (17.55)

which is the same as (17.49).

If a field has several local Lorentz indices \( \alpha, \beta, \gamma, \ldots \), each index receives an own contribution proportional to the gauge field \( A_{\alpha \beta} \). If it has, in addition, Einstein indices \( \mu, \nu, \lambda, \ldots \), there are also additional terms proportional to the affine connection \( \Gamma^\lambda _{\mu \nu} \). As an example, the covariant derivatives of the fields \( v^\mu _\beta \) and \( v^\mu _\nu \) with respect to the nonholonomic coordinates \( dx^\alpha \) are from (17.48) and (12.67), (12.68):

\[
D_\alpha v^\mu _\beta = \partial_\alpha v^\mu _\beta - \Lambda^\alpha _{\alpha \beta} \gamma v^\gamma _\gamma + h^a \Gamma^a _{\alpha \beta} \mu v^\nu _\beta.
\] (17.56)

\[
D_\alpha v^\beta _\mu = \partial_\alpha v^\beta _\mu + \Lambda^\alpha _{\alpha \gamma} \beta v^\gamma _\gamma - h^a \Gamma^a _{\alpha \mu} \nu v^\nu _\nu.
\] (17.57)

The covariant derivatives with respect to the physical coordinates \( x^\lambda \) are

\[
D^\lambda v^\mu _\beta = \partial^\lambda v^\mu _\beta - h^a \Lambda^\lambda _{\alpha \beta} \mu v^\gamma _\gamma + \Gamma^\mu _{\lambda \nu} \mu v^\nu _\beta.
\] (17.58)

\[
D^\lambda v^\beta _\mu = \partial^\lambda v^\beta _\mu + h^a \Lambda^\lambda _{\alpha \gamma} \beta v^\gamma _\nu - \Gamma^\nu _{\lambda \mu} \nu v^\beta _\nu.
\] (17.59)

Let us express the spin connection (17.49) for vector fields in terms of \( e^a _\mu \) and \( h^a _\mu \). With the help of (17.41), we calculate

\[
\Gamma^\alpha _{\alpha \beta} = e^a _\lambda h^\gamma _{\lambda \alpha} h^\beta _{\mu \beta} (e^a _\nu h^\nu _{\beta \beta}) = h^\gamma _{\lambda \alpha} h^\mu _{\mu \beta} (\Gamma^\mu _{\mu \nu} + h^\gamma _{\lambda \alpha} \delta^\lambda _{\mu} \partial^\nu \mu h^\nu _{\beta \beta}).
\] (17.60)
Employing the covariant derivatives (17.56) and (17.57), this equation can be recast as
\[ D_\alpha h_\beta^\mu = 0, \quad D_\alpha h_\beta^\mu = 0, \] (17.61)
so that \( h_\alpha^\mu \) satisfies similar identities as \( e_a^\mu \) in (11.94) and as \( \Lambda_a^\alpha \) in (17.51).

At this place it is useful to introduce the symbols
\[ h_{\Gamma}^{\mu\nu\lambda} \equiv h_\alpha^\lambda \partial_\mu h_\alpha^\nu \equiv -h_\alpha^\nu \partial_\mu h_\alpha^\lambda. \] (17.62)

They are defined in terms of \( h_\alpha^\mu \) in the same way as \( \Gamma^{\mu\nu\lambda} \) is defined in terms of \( e_a^\mu \) in Eq. (11.92). Then we may rewrite Eq. (17.60) as
\[ \Lambda^\alpha_{\Gamma}^{\mu\nu} = h_\gamma^\lambda h_\alpha^\mu (\Gamma^{\mu\nu\lambda} - h_\delta^\lambda \partial_\mu h_\delta^\nu) = h_\gamma^\lambda h_\alpha^\mu h_\beta^\nu (\Gamma^{\mu\nu\lambda} - \Lambda_{\Gamma}^{\mu\nu\lambda}). \] (17.63)

If we now decompose the two connections on the right-hand side into Christoffel parts and contortion tensors in the same way as in Eqs. (11.114)–(11.116), we realize that due to the identity
\[ g_{\mu\nu}(x) = e_a^\mu(x)e_b^\nu(x)g_{ab} \equiv h_\alpha^\mu(x)h_\beta^\nu(x)g_{\alpha\beta}, \] (17.64)
the two Christoffel parts in \( \Gamma_{\mu\nu}^{\lambda} \) and \( h_{\Gamma}^{\mu\nu\lambda} \) are the same:
\[ \bar{\Gamma}_{\mu\nu}^{\lambda} \equiv \frac{h}{\Gamma}_{\mu\nu}^{\lambda}. \] (17.65)

As a consequence, \( \Lambda^\alpha_{\Gamma}^{\mu\nu\gamma} \) becomes
\[ \Lambda^\alpha_{\Gamma}^{\mu\nu\gamma} = h_\gamma^\lambda h_\alpha^\mu h_\beta^\nu (\Gamma^{\mu\nu\lambda} - K_{\mu\nu}^{\lambda} - \bar{h}_{\Gamma}^{\mu\nu\lambda} - h_{\Gamma}^{\mu\nu\lambda}), \] (17.66)
where \( K_{\mu\nu}^{\lambda} \) is the contortion tensor (11.118), and \( h_{\Gamma}^{\mu\nu\lambda} \) denotes the expression (11.116) with \( e_a^\mu \), \( e_a^\mu \) replaced by \( h_\alpha^\mu \), \( h_\alpha^\mu \). Explicitly, these tensors are
\[ K_{\mu\nu}^{\lambda} = S_{\mu\nu}^{\lambda} - S_{\nu\mu}^{\lambda} + S_{\mu\nu}^{\lambda}, \] (17.67)
\[ h_{\Gamma}^{\mu\nu\lambda} = S_{\mu\nu}^{\lambda} - \bar{h}_{\Gamma}^{\mu\nu\lambda} + h_{\Gamma}^{\mu\nu\lambda}, \] (17.68)
where
\[ h_{\Gamma}^{\mu\nu\lambda} \equiv \frac{1}{2} \left( h_\alpha^\lambda \partial_\mu h_\alpha^\nu - h_\alpha^\lambda \partial_\nu h_\alpha^\mu \right). \] (17.69)
This is the so-called object of anholonomy, often denoted by \( \Omega_{\mu\nu}^{\lambda} \). They are antisymmetric in the first two indices, which makes their combinations (17.68) antisymmetric in the last two indices, if the last index is lowered by a contraction with \( g_{\lambda\kappa} \).
The tensors (17.69) and (17.68) have therefore the same symmetry properties as the contortion and torsion tensors. The spin connection (17.66) in which the last index is lowered by a contraction with the Minkowski metric \( g_{\alpha\beta} \) [recall (17.27)] is then antisymmetric in the last two indices.

Note that the antisymmetric part of the spin connection (17.66) is

\[
\Lambda S_{\alpha\beta}^{\gamma} \equiv \frac{1}{2} \left( \Gamma^{\gamma}_{\alpha\beta} - \Gamma^{\gamma}_{\beta\alpha} \right) = h^{\gamma\lambda} h_{\alpha\mu} h_{\beta\nu} (S_{\mu\nu}^\lambda - h^{\mu\nu} S_{\mu\nu}^\lambda). \tag{17.70}
\]

It will be helpful to freely use \( h_{\alpha\mu}^\gamma \) for changing indices \( \alpha \) into \( \mu \), for instance,

\[
K_{\alpha\beta}^{\gamma} \equiv h^{\gamma\lambda} h_{\alpha\mu} h_{\beta\nu} K_{\mu\nu}^\lambda, \tag{17.71}
\]

\[
\tilde{K}_{\alpha\beta}^{\lambda} = h^{\gamma\lambda} h_{\alpha\mu} h_{\beta\nu} \tilde{K}_{\mu\nu}^\lambda. \tag{17.72}
\]

### 17.2 Dirac Action in Riemann-Cartan Space

Inserting (17.66) into (17.22) we finally obtain single-valued image of the flat-spacetime action (17.1):

\[
\mathcal{A} = \int d^4x \sqrt{-g}(x) \left[ i \gamma^\alpha h_{\alpha\mu}(x) D_{\mu} - m \right] \psi(x), \tag{17.73}
\]

with the covariant derivative

\[
D_{\mu} \equiv \partial_{\mu} - i \frac{1}{2} h^{\gamma\lambda} h_{\alpha\mu} h_{\beta\nu} (K_{\mu\nu}^\lambda - h_{\mu\nu}^\lambda \tilde{K}_{\mu\nu}^\lambda) \Sigma_{\beta}^{\gamma}. \tag{17.74}
\]

Note that with the help of the gauge fields in (17.49), the curvature tensor \( R_{\mu\nu\alpha}^{\gamma} \) defined in Eq. (17.42) can be rewritten as

\[
R_{\mu\nu\alpha}^{\gamma} = \Lambda_{\alpha}^{\gamma} \left( \partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu} \right) \Lambda_{\alpha}^{\mu} = \partial_{\mu} \hat{\Gamma}_{\nu\alpha}^{\gamma} - \partial_{\nu} \hat{\Gamma}_{\mu\alpha}^{\gamma} - \hat{\Gamma}_{\mu\nu}^{\lambda} \delta_{\alpha}^{\lambda} \hat{\Gamma}_{\gamma}^{\lambda} + \hat{\Gamma}_{\nu\alpha}^{\lambda} \delta_{\mu}^{\lambda} \hat{\Gamma}_{\gamma}^{\lambda}. \tag{17.75}
\]

This follows directly by performing the derivatives successively and inserting (17.49), while using the pseudo-orthogonality of the Lorentz matrices \( \Lambda(x) \). On the right-hand side we recognize the standard covariant curl formed from the nonabelian gauge field \( \hat{\Gamma}_{\nu\alpha}^{\gamma} \) in the same way as in Eq. (17.50). Thus we shall denote the right-hand side of (17.75) by

\[
F_{\mu\nu\alpha}^{\gamma} \equiv \partial_{\mu} \hat{\Gamma}_{\nu\alpha}^{\gamma} - \partial_{\nu} \hat{\Gamma}_{\mu\alpha}^{\gamma} - \hat{\Gamma}_{\mu\nu}^{\lambda} \delta_{\alpha}^{\lambda} \hat{\Gamma}_{\gamma}^{\lambda} + \hat{\Gamma}_{\nu\alpha}^{\lambda} \delta_{\mu}^{\lambda} \hat{\Gamma}_{\gamma}^{\lambda} = R_{\mu\nu\alpha}^{\gamma} = h_{\alpha}^{\lambda} R_{\mu\nu\lambda}^{\kappa} h_{\beta}^{\kappa}. \tag{17.76}
\]

It is instructive to prove this equality in another way using Eq. (17.63). This leads to the complicated expression

\[
F_{\mu\nu\beta}^{\gamma} = \left\{ \partial_{\mu} \left[ (\Gamma - \hat{\Gamma}_{\nu\alpha}^{\lambda} h_{\beta}^{\lambda} h_{\kappa}^{\kappa}) \right] - (\mu \leftrightarrow \nu) \right\} - \left\{ (\Gamma - \hat{\Gamma}_{\mu\alpha}^{\lambda} (\Gamma - \hat{\Gamma}_{\nu\beta}^{\lambda} h_{\alpha}^{\alpha} h_{\gamma}^{\gamma} - (\mu \leftrightarrow \nu) \right\}, \tag{17.77}
\]
which may be regrouped to
\[
\left[ \partial_\mu \Gamma^\kappa_\nu_\lambda - (\Gamma^\lambda_\mu_\nu)\kappa - (\mu \leftrightarrow \nu) \right] h^\lambda_\beta h^\gamma_\kappa \\
+ \left\{ \Gamma^\kappa_\nu_\lambda \partial_\mu \left( h^\lambda_\beta h^\gamma_\kappa \right) - \partial_\mu \left( \Gamma^\lambda_\mu_\nu \kappa h^\beta_\gamma_\kappa \right) - (\mu \leftrightarrow \nu) \right\} \\
+ \left\{ \left( \Gamma^h_\mu_\nu + \Gamma^h_\mu_\nu - \Gamma^h_\mu_\nu \right) \kappa h^\lambda_\beta h^\gamma_\kappa - (\mu \leftrightarrow \nu) \right\}. \quad (17.78)
\]

Recalling (11.129) we see that the equality (17.76) is verified if we demonstrate the vanishing of the terms in curly brackets. The first term inside these brackets is
\[
\Gamma^\kappa_\nu_\lambda \partial_\mu h^\lambda_\beta h^\gamma_\kappa + \Gamma^\lambda_\mu_\nu \partial_\mu h^\beta_\gamma_\kappa - (\mu \leftrightarrow \nu) = -\Gamma^\gamma_\nu_\lambda h^h_\mu_\beta + \Gamma^\lambda_\nu_\beta \Gamma^\kappa_\nu_\beta h^h_\mu_\kappa + (\mu \leftrightarrow \nu),
\]
and the second contributes [using (17.62)]
\[
-\partial_\mu \left( h^\lambda_\beta \partial_\nu h^\gamma_\lambda \right) - (\mu \leftrightarrow \nu) = h^h_\mu_\beta \lambda \Gamma^{\lambda \mu \gamma}_\nu \beta - (\mu \leftrightarrow \nu) - h^\lambda_\beta \left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) h^\gamma_\lambda.
\]

Thus we find indeed, recalling (17.38) and (17.39),
\[
F^\gamma_\mu_\nu_\beta = \left( R^\gamma_\mu_\nu_\beta - h^h_\mu_\nu \kappa \right) h^\lambda_\beta h^\gamma_\kappa = R^\gamma_\mu_\nu_\lambda \kappa h^\beta_\gamma_\kappa. \quad (17.79)
\]

### 17.3 Ricci Identity

The equality of the covariant curls of $F^\gamma_\mu_\nu_\beta$ and $R^\kappa_\mu_\nu_\lambda$ up to a coordinate transformation of the last two indices is related to a fundamental algebraic property of covariant derivatives. Consider a vector field $v_\lambda$ and apply a commutator of covariant derivatives to it, yielding
\[
\left[ D^\mu_\nu, D^\nu_\mu \right] v_\lambda = \partial_\mu \left( \partial_\nu v_\lambda - \Gamma^\kappa_\nu_\lambda \kappa v_\kappa \right) - \Gamma^\kappa_\mu_\nu \tau D^\tau v_\lambda - \Gamma^\kappa_\mu_\nu \tau \left( \Gamma^\lambda_\nu_\tau \kappa v_\kappa - (\mu \leftrightarrow \nu) \right). \quad (17.80)
\]

For a single-valued vector field satisfying $\left( \partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) v_\lambda = 0$, we obtain the so-called Ricci identity
\[
\left[ D^\mu_\nu, D^\nu_\mu \right] v_\lambda = -R^\kappa_\mu_\nu_\lambda \kappa v_\kappa - 2S^\tau_\mu_\nu D^\tau v_\lambda. \quad (17.81)
\]

For a general tensor, $R^\kappa_\mu_\nu_\lambda$ and $S^\tau_\mu_\nu$ act separately on each index. Now, a similar relation may be calculated for the components of the vector in the nonholonomic basis $e^\alpha_\beta$:
\[
\left[ D^\mu_\nu, D^\nu_\mu \right] v_\beta = \partial_\mu \left( \partial_\nu v_\beta - \Gamma^\lambda_\nu_\beta \tau v_\gamma \right) - \Gamma^\lambda_\mu_\nu \tau D^\tau v_\beta - \Gamma^\lambda_\mu_\nu \tau \left( \Gamma^\lambda_\nu_\tau \delta v_\delta - (\mu \leftrightarrow \nu) \right) - \Gamma^\lambda_\mu_\nu \tau \left( \Gamma^\lambda_\nu_\tau \delta v_\delta - (\mu \leftrightarrow \nu) \right) - (\mu \leftrightarrow \nu)
= -F^\gamma_\mu_\nu_\beta \gamma v_\gamma - 2S^\tau_\mu_\nu D^\tau v_\beta. \quad (17.82)
\]
For a field of arbitrary spin, this generalizes to

\[
\left[ D_\mu, D_\nu \right] \psi = \frac{i}{2} F_{\mu\nu}^\beta \gamma^\beta \psi - 2 S_{\mu\nu}^\tau D_\tau \psi.
\] (17.83)

Due to the complete covariance of (17.80) and (17.82), we may multiply (17.80) by \( h_\beta^\lambda \) and pass this factor through the covariant derivatives (which, in this process, change their connection since they are applied to different objects before and after the passage). The \( R \)-term in (17.81) and the \( F \) term in (17.83) remain simply related by (17.75).

17.4 Alternative Form of Coupling

Let us compare the above derived minimal coupling of the vierbein field \( h_\alpha^{\mu}(x) \) to a spinning particle with the theory of Weyl, Fock, and Iwanenko [2, 3, 4] in Riemann spacetimes. These authors proposed to search for spacetime-dependent Dirac matrices \( \gamma^\mu(x) \) solving the Dirac algebra

\[
\{ \gamma^\mu(x), \gamma^\nu(x) \} = g^{\mu\nu}(x),
\] (17.84)

which can obviously be expressed in terms of our vierbein fields as

\[
\gamma^\mu(x) = \gamma^\alpha h_\alpha^\mu(x).
\] (17.85)

In terms of these, the Dirac action can be written as

\[
\mathcal{A} = \int d^4 x \sqrt{-g} \bar{\psi}(x) \left\{ i \gamma^\mu(x) D_\mu - m \right\} \psi(x),
\] (17.86)

where \( D_\mu \) is the covariant derivative (omitting the Dirac spin indices)

\[
D_\mu = \partial_\mu \delta_\mu - \Gamma_\mu(x),
\] (17.87)

with the Dirac spin connection [compare (17.13)]

\[
\Gamma_\mu(x) \equiv -\frac{1}{4} \gamma_\lambda(x) D_\mu \gamma^\lambda(x) = -\frac{1}{4} \gamma_\lambda(x) \left[ \partial_\mu \gamma^\lambda(x) + \Gamma_{\mu\nu}^\lambda(x) \gamma^\nu(x) \right].
\] (17.88)

Let us verify that the action (17.86) is equivalent to the previous one in (17.73) if there is no torsion. Inserting (17.85) into (17.88) we find

\[
\Gamma_\mu(x) = -\left( h_\alpha^\lambda \partial_\mu h^\beta_\lambda + h_\alpha^\lambda h^\nu_\beta \Gamma_{\mu\nu}^\lambda \right) \frac{1}{4} \gamma^\alpha \gamma^\beta.
\] (17.89)

We now use the analog of (11.87) for \( h_\alpha^{\nu}(x) \) [which follows from the completeness relation (17.19)]

\[
\partial_\mu h^\lambda_\beta = -h^\nu_\beta \left( h^\lambda \partial_\mu h^\gamma_\nu \right).
\] (17.90)
Then we can rewrite (17.89) as

\[ \Gamma_{\mu} = \left( h_{\Gamma_{\mu\nu}^{\lambda}} - \Gamma_{\mu\nu}^{\lambda} \right) \frac{1}{4} \gamma^{\lambda} \gamma^{\nu}, \]  

(17.91)

where we have used the definition (17.62). Comparison with (17.63) shows that

\[ \Gamma_{\mu} = - \Gamma_{\mu\beta}^{\lambda} \frac{i}{4} \gamma^{\alpha} \gamma^{\beta}. \]  

(17.92)

Since \( \Gamma_{\mu\alpha\beta}^{\lambda} \) is antisymmetric in \( \alpha\beta \), this is the same as [recall (1.222)]

\[ \Gamma_{\mu} = - \frac{i}{2} \Gamma_{\mu\alpha\beta}^{\lambda} \Sigma^{\alpha\beta}, \]  

(17.93)

in agreement with (17.15), if the Dirac indices are added.

### 17.5 Invariant Action for Vector Fields

Any theory which is invariant under general coordinate transformations can be recast in such a way that its derivatives refer to the nonholonomic coordinates \( dx^\alpha \). Since the metric in these coordinates is \( g^{\alpha\beta} \), the action has the same form as those in a flat spacetime, except that derivatives of vector and tensor fields are replaced by covariant ones, for example

\[ \partial_{\alpha} v_{\beta} \rightarrow D_{\alpha} v_{\beta} = \partial_{\alpha} v_{\beta} - \Gamma_{\alpha\beta}^{\lambda} \gamma_{\gamma}. \]  

(17.94)

For example,

\[ A = \int d^4x^{\alpha} D^\alpha v_{\beta}(x) D_\alpha v^\beta(x) \]  

(17.95)

is the nonholonomic form of a generally covariant action. As we said in the beginning, the specification of spacetime points must be made with the \( x^\mu \) coordinates. For this reason the action is preferably written as

\[ A = \int d^4x^{\mu} \sqrt{-g} D_\alpha v_{\beta}(x^{\mu}) D_\alpha v^\beta(x^{\mu}). \]  

(17.96)

Under a general coordinate transformation à la Einstein, \( dx^\mu \rightarrow dx'^{\mu'} = dx^{\mu} \alpha_\mu^{\mu'} \), the indices \( \alpha \) are inert. For instance, \( h_\alpha^{\mu'}(x) \) itself transforms as

\[ h_\alpha^{\mu'}(x) \rightarrow h_\alpha^{\mu'}(x') = h_\alpha^{\mu}(x) \alpha_\mu^{\mu'}. \]  

(17.97)

Vectors and tensors with indices \( \alpha, \beta, \ldots \), experience only changes of their arguments \( x \rightarrow x - \xi \) so that their infinitesimal substantial changes are

\[ \delta_E v_\alpha(x) = \xi^\lambda \partial_\lambda v_\alpha(x) \]  

(17.98)

\[ \delta_E D_\alpha v_{\beta}(x) = \xi^\lambda \partial_\lambda D_\alpha v_{\beta}(x). \]  

(17.99)
The freedom in choosing $h_\alpha^\mu(x)$ up to a local Lorentz transformation, when taking the “square root” of $g_{\mu\nu}(x)$ in (2.50), implies that the theory should be invariant under

$$
\delta_L dx^\alpha = \omega^\alpha_\beta(x) dx^\beta, \quad (17.100)
$$

$$
\delta_L h_\alpha^\mu(x) = \omega^\alpha_\beta(x) h_\beta^\mu(x). \quad (17.101)
$$

Here $\omega^\alpha_\alpha(x)$ are the local versions of the infinitesimal angles introduced in (11.58) and (11.59).

Indeed the action (17.96) is automatically invariant if every index $\alpha$ is transformed accordingly.

$$
\delta_L v_\alpha(x) = \omega^\alpha_\alpha(x) v_\alpha(x), \quad (17.102)
$$

$$
\delta_L D_\alpha v_\beta(x) = \omega^\alpha_\alpha(x) D_\alpha v_\beta(x) + \omega^\beta_\alpha(x) D_\alpha v_\beta(x). \quad (17.103)
$$

The variables $x^\mu$ are unchanged since the local Lorentz transformations (17.100) affect only the intermediate local directions defined by the differentials $dx^\alpha$. They leave the physical coordinate $x^\mu$ unchanged.

It is useful to verify explicitly how the covariant derivatives guarantee local Lorentz invariance. Consider

$$
\delta_L v_\alpha = \omega^\alpha_\alpha(x) v_\alpha(x), \quad \delta_L v_\alpha = \omega^\alpha_\alpha(x) v_\alpha. \quad (17.104)
$$

Then the derivative $\partial_\alpha v_\beta$ transforms as

$$
\delta_L \partial_\alpha v_\beta = (\delta_L \partial_\alpha) v_\beta + \partial_\alpha (\delta_L v_\beta)
= \omega^\alpha_\alpha \partial_\alpha v_\beta + \partial_\alpha (\omega^\beta_\alpha v_\beta)
= \omega^\alpha_\alpha \partial_\alpha v_\beta + \omega^\beta_\alpha \partial_\alpha v_\beta + (\partial_\alpha \omega^\beta_\alpha) v_\beta. \quad (17.105)
$$

The spin connection behaves as follows: Due to the factors $h_\alpha^\lambda h_\alpha^\mu h_\beta^\nu$ in (17.63), the first term in $\Gamma^{\lambda}_{\alpha\beta}$, call it $\Gamma^{\lambda(1)}_{\alpha\beta}$, transforms like a local Lorentz tensor:

$$
\delta_L \Gamma^{\lambda(1)}_{\alpha\beta} = \omega^\lambda_\alpha \Gamma^{\lambda(1)}_{\alpha\beta} + \omega^\beta_\alpha \Gamma^{\lambda(1)}_{\alpha\beta} + \omega^\gamma_\alpha \Gamma^{\lambda(1)}_{\alpha\beta}. \quad (17.106)
$$

Only the substantial variation of the second term $\Gamma^{\lambda}_{\mu\nu}$ in (17.63) contains a nontensorial derivative contribution:

$$
\delta_L \Gamma^{\lambda}_{\mu\nu} = (\delta h^\lambda_\delta) \partial^\delta_\nu + h^\lambda_\delta \partial_\mu (\delta h^\delta_\nu)
= \omega^\delta_\delta h^\lambda_\delta \partial^\delta_\nu + h^\lambda_\delta \partial_\mu (\omega^\delta_\delta h^\delta_\nu)
= \omega^\delta_\delta h^\lambda_\delta \partial^\delta_\nu + \omega^\delta_\delta h^\lambda_\delta \partial^\delta_\nu + \partial_\mu \omega^\delta_\delta (h^\lambda_\delta h^\delta_\nu)
= \partial_\mu \omega^\delta_\delta h^\lambda_\delta h^\delta_\nu. \quad (17.107)
$$

H. Kleinert, GRA VITY WITH TORSION
the cancellation in the third line being due to the antisymmetry of \( \omega^\delta_\delta = -\omega^\delta_\delta. \)
Thus we arrive at
\[
\delta_L \Gamma^\mu_{\rho\nu} = \partial^\mu \omega^\rho_\delta h^\nu_\delta', \quad \delta_L \Gamma^\Lambda_{\alpha\beta} = \delta_L \Gamma^\Lambda_{\alpha\beta} + \partial^\Lambda \omega^\gamma_\beta.
\] (17.108)
The last term is precisely what is required to cancel the last non-tensorial part in (17.105) when transforming \( D^\alpha v^\beta, \) so that we indeed obtain the covariant transformation law (17.103).

### 17.6 Verifying Local Lorentz Invariance

Let us study the invariance under local Lorentz transformations in more detail. These serve to go at an arbitrary point \( x^\mu \) from one freely falling elevator to another. A spinor field \( \psi(x) \) transforms under them as follows:
\[
\delta_L \psi(x) = -\frac{i}{2} \omega^{\alpha\beta}(x) \Sigma_{\alpha\beta} \psi(x).
\] (17.109)
Here \( \Sigma_{\alpha\beta} \) are the spin representation matrices of the local Lorentz group. They are antisymmetric in \( \alpha, \beta \) and satisfy the commutation relations
\[
[\Sigma_{\alpha\beta}, \Sigma_{\alpha\gamma}] = -ig_{\alpha\alpha'} \Sigma_{\beta\gamma'}.
\] (17.110)
Recall that for vectors, the representation matrices were given by (1.51):
\[
(L_{\alpha\beta})_{\alpha'\beta'} = i \left[ g_{\alpha\alpha'} g_{\beta\beta'} - (\alpha \leftrightarrow \beta) \right],
\] (17.111)
by which the analog of the transformation law (17.109) reduces to (17.102):
\[
\delta_L v^\alpha = -\frac{i}{2} \omega^{\gamma\delta} i \left( g_{\gamma\alpha} \delta^\beta_\delta - g_{\delta\alpha} \delta^\beta_\gamma \right) v^\beta = \omega^\alpha_\beta v^\beta.
\] (17.112)
For Dirac spinors, the Lorentz generators \( L_{\alpha\beta} \) are replaced by \( \Sigma_{\alpha\beta} \) of Eq. (1.221):
\[
\Sigma_{\alpha\beta} = \frac{i}{4} [\gamma^\alpha, \gamma^\beta].
\] (17.113)
The infinitesimal Lorentz transformation of the derivative of \( \psi \) is
\[
\delta_L \partial^\alpha \psi = \omega_{\alpha}^{\alpha'} \partial^\alpha \psi + \partial^\alpha \delta_L \psi
= \omega_{\alpha}^{\alpha'} \partial^\alpha \psi - \frac{i}{2} \partial^\alpha (\omega^{\beta\gamma} \Sigma_{\beta\gamma}) \psi
= \omega_{\alpha}^{\alpha'} \partial^\alpha \psi - \frac{i}{2} \omega^{\beta\gamma} \Sigma_{\beta\gamma} \partial^\alpha \psi - \frac{i}{2} (\partial^\alpha \omega^{\beta\gamma}) \Sigma_{\beta\gamma} \psi.
\] (17.114)
The first two terms describe the usual Lorentz behavior of \( \partial^\alpha \psi. \) The last term is due to the dependence of the angles \( \omega^{\beta\gamma}(x) \) on \( x. \) It can be removed with the help of the
spin connection $\Gamma_{\alpha\beta}^\gamma$ and the covariant derivative (17.15) with the spin connection (17.13)

$$D_\alpha \psi(x) \equiv \partial_\alpha \psi(x) + \frac{i}{2} \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \psi(x).$$  \hspace{1cm} (17.115)

Indeed, if we calculate the variation of the second term in $D_\alpha \psi(x) \equiv \partial_\alpha \psi(x)$:

$$\delta L \frac{i}{2} \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \psi(x),$$  \hspace{1cm} (17.116)

we obtain two terms. There is a term with the regular Lorentz transformation property

$$\delta_L \frac{i}{2} \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \psi = -\frac{i}{2} \omega^{\sigma\tau} \Sigma_{\sigma\tau} \left( \frac{i}{2} \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \psi \right).$$  \hspace{1cm} (17.117)

This follows from

$$\frac{i}{2} \delta L \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \psi + \frac{i}{2} \Gamma_{\alpha\beta}^\gamma \Sigma^\beta_\gamma \delta_L \psi,$$  \hspace{1cm} (17.118)

and an application of the commutation rule (17.110). A second term arises from $\partial_\alpha \omega_\beta^\gamma$, which is

$$\frac{i}{2} \partial_\alpha \omega_\beta^\gamma \Sigma^\beta_\gamma \psi$$  \hspace{1cm} (17.119)

and cancels against the last term in (17.114). Thus $D_\alpha \psi$ behaves like

$$\delta_L D_\alpha \psi = \omega_\alpha^{\alpha'}(x) D_{\alpha'} \psi - \frac{i}{2} \omega^{\beta\gamma}(x) \Sigma_{\beta\gamma} D_\alpha \psi,$$  \hspace{1cm} (17.120)

and represents, therefore, a proper covariant derivative which generalizes the standard Lorentz transformation behavior to the case of local transformations $\omega_\alpha^\beta(x)$.

An important observation is the following. The covariant derivative (17.15) does not need the contortion field $K_{\mu\nu}^\lambda$ to be covariant under local Lorentz transformations. For this, the term $h_{\mu\nu}^\lambda$ is completely sufficient. It supplies the compensating nontensorial term to make the derivative in front of a Dirac field a vector. Thus a consistent theory which is invariant under local Lorentz transformations exists in a Riemann spacetime. Torsion is a pure luxury of the theory.

### 17.7 Field Equations with Gravitational Spinning Matter

Consider the action of a spin–1/2 field interacting with a gravitational field:

$$A[h, K, \psi] = -\frac{1}{2K} \int d^4 x \sqrt{-g} R + \frac{1}{2} \int d^4 x \sqrt{-g} \bar{\psi} \gamma^\alpha D_\alpha \psi(x) + \text{h.c.}$$

$$= \int A[h, K] + m \int A[h, K, \psi].$$  \hspace{1cm} (17.121)

It is a functional of the vierbein field $h_\alpha^\mu$, the contortion $K_{\mu\nu}^\lambda$ and the Dirac field $\psi(x)$. Varying $A$ with respect to $\bar{\psi}$ we obtain the equation of motion

$$\frac{\delta A}{\delta \bar{\psi}} = \sqrt{-g} (\gamma^\alpha D_\alpha - m) \psi(x) = 0$$  \hspace{1cm} (17.122)
of a Dirac particle in a general affine spacetime.

To obtain the gravitational field equations we again define the spin current density, just as we did in (15.17), by differentiating with respect to \( K_{\mu \nu}^\lambda \) at fixed \( h_{\alpha \mu} \), and find for the gravitational field

\[
\frac{\delta f}{\delta K_{\mu \nu}^\lambda} = -\frac{1}{2} \sqrt{-g} \Sigma_{\lambda \mu}^{\nu}, \tag{17.123}
\]
as given in (15.63).

From the matter action (17.15) we obtain

\[
\sqrt{-g} \Sigma_{\lambda \mu}^{\nu} = 2 \frac{\delta m}{\delta K_{\mu \nu}} = \sqrt{-g} \left[ -\frac{i}{2} \bar{\psi}(x) \gamma^\mu(x) \gamma^\nu(x) \Sigma_{\lambda \mu}^{\nu}(x) \psi(x) + \text{h.c.} \right] = h_{\gamma \lambda} h_{\alpha \mu} h_{\beta \nu} \sqrt{-g} \left[ -\frac{i}{2} \bar{\psi}(x) \gamma^{\alpha \beta}(x) \psi + \text{h.c.} \right] = h_{\gamma \lambda} h_{\alpha \mu} h_{\beta \nu} \sqrt{-g} \Sigma_{\gamma \beta}^{\alpha \mu}.
\tag{17.124}
\]
The expression \( h_{\gamma \lambda}^{\beta \alpha} \) is recognized as the canonical spin current of a Dirac particle in Minkowski spacetime and \( h_{\gamma \lambda}^{\beta \alpha} \) is its generally covariant analog. Thus, for the spin-1/2 field, the definition (15.16) of the spin current density is consistent with the canonical definition,

\[
\Sigma_{\lambda \mu}^{\nu} \equiv -i \sum_i \pi_i^{\lambda \mu} \varphi_i = -i \sum_i \frac{\partial L}{\partial D_{\mu} \varphi_i} \Sigma_{\lambda \mu}^{\nu} \varphi_i, \tag{17.125}
\]
where the sum over \( i \) covers all independent matter fields of the system. This is also true, in general, by the fact that the general Einstein invariant matter action has the functional form [compare (17.15)]

\[
m = m \left[ h, K, \varphi_i \right] = \int d^4x \sqrt{-g} \Sigma_{\lambda \mu}^{\nu} \left( h_{\alpha \mu}^{\nu}, \varphi_i, D_{\mu} \varphi_i \right) \tag{17.126}
\]
so that indeed, for fixed \( h_{\alpha \mu} \),

\[
\left. 2 \frac{\delta m}{\delta K_{\mu \nu}} \right|_{h_{\alpha \mu}} = 2 \sqrt{-g} \sum_i \frac{\partial L}{\partial D_{\mu} \varphi_i} \frac{i}{2} \Sigma_{\lambda \mu}^{\nu} \varphi_i = i \sqrt{-g} \sum_i \pi_i^{\lambda \mu} \varphi_i = -\Sigma_{\lambda \mu}^{\nu} \varphi_i. \tag{17.127}
\]
The field equations associated with \( \delta K_{\mu \nu}^\lambda \) are therefore

\[
-\kappa \Sigma_{\lambda \mu}^{\nu} = \kappa \Sigma_{\lambda \mu}^{\nu}. \tag{17.128}
\]
thus extending the field equation Eq. (15.58) to systems with spinning matter. Together with Eq. (15.54), this determines the Palatini tensor (15.45):

$$S_{\mu\nu,\lambda} = -\kappa \sum^{m}_{\mu\nu,\lambda},$$

(17.129)

and thus, via Eq. (15.46), the torsion of spacetime by the field equation:

$$S_{\mu\nu\lambda} = \kappa \left( \sum^{m}_{\mu\nu,\lambda} \right)^{2} + \frac{1}{2} \sum^{m}_{\mu\nu,\lambda} - \frac{1}{2} \sum^{m}_{\mu\nu,\lambda}. \right),$$

(17.130)

Let us now turn to the field equations arising from extremization with respect to $h_{\alpha}^{\mu}$. We define the total energy-momentum tensor as

$$\sqrt{-g} T_{\mu}^{\alpha}(x) \equiv \frac{\delta A}{\delta h_{\alpha}^{\mu}(x)} \bigg|_{S_{\mu\nu,\lambda}},$$

(17.131)

with the derivative formed at fixed $S_{\mu\nu,\lambda}$. Due to the relation (17.90), we may use the chain rule of differentiation to write alternatively

$$\sqrt{-g} T_{\alpha}^{\mu}(x) = -\frac{\delta A}{\delta h_{\alpha}^{\mu}(x)} \bigg|_{S_{\mu\nu,\lambda}}.$$

(17.132)

For the pure gravitational action which depends only on $g^{\mu\nu} = h^{\alpha\mu} h_{\alpha}^{\nu}$ and $K_{\mu\nu,\lambda}$, this definition leads trivially to the same symmetric energy-momentum tensor as that introduced earlier in (15.15), except that one index has that $\alpha$-form. This follows from the chain rule of differentiation together with (15.15):

$$\sqrt{-g} f_{\mu}^{\alpha} = \frac{\delta A}{\delta h_{\alpha}^{\mu}} = \frac{\delta A}{\delta g^{\alpha\mu} \partial h_{\alpha}^{\mu}} = \sqrt{-g} f_{\mu}^{\alpha} h_{\alpha}^{\nu}.$$

(17.133)

There is, of course, a similar rule involving the derivative with respect to $h_{\alpha}^{\mu}$ as in (17.132):

$$\sqrt{-g} f_{\alpha}^{\mu} = -\frac{\delta A}{\delta h_{\alpha}^{\mu}} = -\frac{\delta A}{\delta g^{\lambda\mu} \partial h_{\alpha}^{\mu}} = \sqrt{-g} f_{\alpha}^{\mu} h_{\alpha}^{\nu}.$$

(17.134)

For matter, the actual calculation of the symmetric energy-momentum tensor is most conveniently performed in two steps. Take, for instance, the Dirac field. As a first step we differentiate $\sqrt{-g}$ and $\gamma^{\alpha} h_{\alpha}^{\mu} \partial_{\mu}$ with respect to $h_{\alpha}^{\mu}$ while keeping, for the moment, $D_{\mu} = \text{const}$. The result is the so-called canonical energy-momentum tensor:

$$\sqrt{-g} \Theta_{\mu}^{\alpha} = \frac{1}{g^{2}} \left( \bar{\psi} \gamma_{i}^{\alpha} i D_{\mu} \psi - h_{\alpha}^{\mu} L \right) + \text{h.c.}$$

(17.135)

This is a general feature of the formalism: The derivative of (17.126) with respect to the $h_{\alpha}^{\mu}$ fields contained in the covariant derivative $D_{\mu} \bar{\psi}_{i} = h_{\alpha}^{\mu} D_{\alpha} \bar{\psi}_{i}$ gives

$$\frac{\delta A}{\delta h_{\alpha}^{\mu}} \rightarrow \sqrt{-g} \sum_{i} \frac{\partial L}{\partial D_{\mu} \bar{\psi}_{i}} D_{\mu} \bar{\psi}_{i} h_{\alpha}^{\mu}.$$

(17.136)
The derivative of (17.126) with respect to the \( h_\alpha^\mu \) contained in the \( \sqrt{-g} \) term adds to this
\[
\delta \frac{m}{\delta h_\alpha^\mu} \rightarrow -\sqrt{-g} g_\mu^\nu \mathcal{L} h_\alpha^\nu. \tag{17.137}
\]
The sum of the two contributions yields
\[
m_\Theta^\alpha_\mu = \left( \sum_i \frac{\partial L}{\partial \frac{\partial \varphi_i}{\partial D^\nu}} D_\mu \varphi_i - g_\mu^\nu \mathcal{L} \right) h_\alpha^\nu, \tag{17.138}
\]
which is indeed the canonical energy-momentum tensor for an arbitrary Lagrangian containing covariant derivatives.

Applying this formalism to a pure gravitational field we can compare the first step of differentiation at fixed \( D_\mu \) with the variation (15.27) and find the symmetric part of the equation
\[
\Theta^\alpha_\mu = -\frac{1}{\kappa} G_\mu^\nu h_\alpha^\nu. \tag{17.139}
\]
We will see below that this holds, in fact, without symmetrization. Thus the canonical energy-momentum tensor of the gravitational field is equal to minus \( 1/\kappa \) times the Einstein tensor.

We now turn to the second step, the calculation of the remaining functional derivative with respect to \( h_\alpha^\mu \). This is somewhat tedious. Let us write the additional contribution to \( m_\Theta^\kappa_\delta \) as
\[
\sqrt{-g} \delta m_\Theta^\kappa_\delta = \int d^4x \frac{\delta A^m}{\delta \Gamma^\kappa_\mu_\beta_\gamma} \bigg|_{\Gamma} = \frac{1}{2} \int d^4x \sqrt{-g} \sum_\beta_\gamma \delta h_\kappa^\mu \frac{\delta \Gamma^\kappa_\mu_\beta_\gamma}{\delta h_\kappa^\mu}, \tag{17.140}
\]
and use for the spin connection the explicit form
\[
\Gamma^\kappa_\mu_\beta_\gamma = h_\gamma^\lambda h_\beta^\nu (\Gamma^\lambda_\mu_\nu - \bar{\Gamma}^\lambda_\mu_\nu) = -h_\beta^\nu D_\mu h_\gamma^\nu = h_\gamma^\nu D_\mu h_\beta^\nu \tag{17.141}
\]
where \( D_\mu \) denotes the part of the covariant derivative containing only the ordinary connection \( \Gamma_\mu_\beta^\lambda \). If we vary \( \delta h_\kappa^\mu \) and hold \( \Gamma_\mu_\beta^\lambda \) fixed we have
\[
\delta \Gamma^\kappa_\mu_\beta_\gamma \bigg|_{\Gamma} = \delta h_\gamma^\nu D_\mu h_\beta^\nu + h_\gamma^\nu D_\mu \delta h_\beta^\nu. \tag{17.142}
\]
Since \( D_\mu h_\gamma^\nu = 0 \) [recall (17.61)], we see that \( D_\mu h_\beta^\nu = \bar{\Gamma}^\kappa_\mu_\beta_\gamma h_\lambda^\nu \) and we may write
\[
\delta \Gamma^\kappa_\mu_\beta_\gamma \bigg|_{\Gamma} = h_\gamma^\nu D_\mu \delta h_\beta^\nu. \tag{17.143}
\]
Inserting this into (17.140), a partial integration gives the first contribution
\[
\Delta_1 m_\Theta^\kappa_\delta = -(1/2)D_\mu S^\kappa_\delta_\mu. \tag{17.144}
\]
We now include the contribution from $\delta \Gamma_{\mu \nu}^{\lambda}$. Using the decomposition (15.50) with $\delta S_{\mu \nu \lambda} = 0$, i.e. $\delta K_{\mu \nu \lambda} = 0$, we find

$$\Delta_2 m^{\delta}_\kappa = \frac{1}{4} \left[ D_{\mu} \left( \Sigma^{\nu \sigma, \mu} - \Sigma^{\nu} \sigma, \mu + \Sigma^{\mu, \sigma} \right) \right] \frac{\partial g_{\nu \sigma}}{\partial h^\delta_{\kappa}}.$$  \hfill (17.145)

With

$$\frac{\partial g_{\nu \sigma}}{\partial h^\delta_{\kappa}} = g_{\nu \kappa} h^\delta_{\sigma} + (\nu \leftrightarrow \sigma),$$ \hfill (17.146)

this gives, altogether,

$$\Delta m^{\delta}_\kappa(x) = - \frac{1}{2} D^*_\mu \left( \Sigma^{\mu \delta}_\kappa - \Sigma^{\mu, \kappa} \right).$$ \hfill (17.147)

This is precisely the same type of correction $\Delta \Theta^{\delta}_\kappa = \Delta \Theta^{\mu \nu \delta}_\kappa$ that was added to the canonical energy-momentum tensor $\Theta^{\mu \nu \delta}$ of the gravitational field in Eq. (15.57), in order to produce the symmetric one $T^{\mu \nu \delta}$. Here it is obtained for arbitrary spinning matter fields:

$$m^{\mu \nu \delta}_\kappa = \Theta^{\mu \nu \delta}_\kappa + \Delta \Theta^{\mu \nu \delta}_\kappa + \Delta \Theta^{\mu \nu \delta}_\kappa = - \frac{1}{2} D^*_\mu \left( \Sigma^{\mu \delta}_\kappa - \Sigma^{\mu, \kappa} \right).$$ \hfill (17.148)

For spin $\frac{1}{2}$, this is the expression (3.230) found by Belinfante in 1939. We have lowered the index $\nu$ on both sides which is permissible due to the covariant form of the equation.

In terms of $m^{\mu \nu \delta}_\kappa$, the field equations which follow from variations of the action with respect to $\delta h^\mu_{\alpha}$ have once more the simple form (15.61):

$$G^{\mu \nu} = \kappa m^{\mu \nu \delta},$$ \hfill (17.149)

with the energy-momentum tensors (17.148) of spinning matter.

Notes and References

[1] This Chapter follows largely the textbook


Covariant Conservation Law

According to Noether’s theorem derived in Chapter 3, the invariance of the action and general coordinate transformations and local Lorentz transformations must be associated with certain conservation laws. For the following considerations, we shall consider $h_\alpha^\mu(x)$ and $\bar{\Gamma}_{\mu\beta}^\alpha(x)$ as independent variables. Then, from the derivation of the canonical energy-momentum tensor in (17.133) it follows that varying the action in $h_\alpha^\nu(x)$ at fixed $A_{\mu\beta}^\gamma(x)$ gives the canonical energy-momentum tensor

$$\frac{\delta A[h_\alpha^\mu, A_{\mu\beta}^\gamma]}{\delta h_\alpha^\mu} = \sqrt{-g} \Theta_\mu^\alpha.$$  

A functional derivative with respect to $A_{\mu\beta}^\gamma = \bar{\Gamma}_{\mu\beta}^\gamma$ at fixed $h_\alpha^\mu$, on the other hand, is equivalent to a derivative with respect to $K_{\mu\beta}^\gamma$ [recall (17.66)] and produces, according to Eq. (17.123), the spin current density

$$\frac{\delta A[h_\alpha^\mu, A_{\mu\beta}^\gamma]}{\delta A_{\mu\beta}^\gamma} = -\frac{1}{2} \sqrt{-g} \Sigma_\gamma^\beta \Sigma_\beta^\alpha h_\alpha^\mu \equiv -\frac{1}{2} \sqrt{-g} \Sigma_\gamma^\beta \cdot \mu.$$  

These quantities will now be shown to satisfy covariant conservation laws.

18.1 Spin Density

Consider first local Lorentz transformations. Under these the vierbein fields $h_\alpha^\mu(x)$ ($\mu = 0, \ldots, 3$) behave like vectors in the index $\alpha$,

$$\delta_L h_\alpha^\mu(x) = \omega_\alpha^{\alpha'}(x) h_\alpha^{\alpha'}(x).$$  

Similarly, the field $A_{\mu\beta}^\gamma$ ($\mu = 0, \ldots, 3$) behaves like a tensor in $\beta, \gamma$, and receives, in addition, a typical derivative term of gauge fields [see (17.112)]

$$\delta_L A_{\mu\beta}^\gamma = \omega_\beta^{\beta'}(x) A_{\mu\beta'}^\gamma + \omega^{\gamma \gamma'}(x) A_{\mu\beta}^{\gamma'} + \partial_\mu \omega_\beta^\gamma(x).$$  

Recall that the field $A_{\mu\beta}^\gamma$ has the pure contortion form, $A_{\mu\beta}^\gamma = h_{\gamma \lambda} h_{\beta \nu}^\alpha (K_{\mu\nu}^\lambda - \bar{K}_{\mu\nu}^\lambda)$ and thus is antisymmetric in $\beta, \gamma$, as is the case with $\bar{\Gamma}_{\alpha\beta}^\gamma$. 

\[\text{427}\]
The symmetry variations \( \delta_L A \) of the action have to vanish as a consequence of the Euler-Lagrange equations. Inserting (18.3) and (18.4) into an arbitrary invariant action \( A \), we obtain

\[
\delta_L A = \int d^4x \left\{ \delta A_\alpha^\alpha(x) h_\alpha^\mu(x) + \frac{\delta A}{\delta A_{\mu\beta}^\gamma}(x) \left( \omega_\beta^\beta(x) A_{\mu\beta}^\gamma(x) + \omega_\gamma^\gamma(x) A_{\mu\beta}^\gamma(x) + \partial_\mu \omega_\beta^\gamma(x) \right) \right\}
\]

(18.5)

Partially integrating the last term gives

\[
\int d^4x \sqrt{-g} \left\{ \Theta_\mu^\alpha \omega_\alpha^\gamma h_\gamma^\mu - \frac{1}{2} \Sigma_\gamma^\beta^\mu \left( \omega_\beta^\beta A_{\mu\beta}^\gamma + \omega_\gamma^\gamma A_{\mu\beta}^\gamma + \partial_\mu \omega_\beta^\gamma \right) \right\}.
\]

(18.6)

Since \( \omega_\beta^\gamma(x') \) is an arbitrary antisymmetric function of \( x' \) it can be chosen to be zero everywhere except at some place \( x \) and we find

\[
\frac{1}{2} \left[ \Theta_\gamma^\beta - \Theta_\beta^\gamma \right] + \frac{1}{2} \Gamma_\mu^\sigma \Sigma_\gamma^\beta^\sigma = 0 \quad (18.9)
\]

where \( \Gamma_\mu^\sigma \Sigma_\gamma^\beta^\sigma \) is the covariant derivative for the local Lorentz index \( \gamma \), i.e., for a vector

\[
\frac{L}{D_\mu} v_\alpha = \partial_\mu v_\alpha - A_{\mu\alpha}^\beta v_\beta = h_\alpha^\beta D_\beta v_\alpha,
\]

(18.10)

and the derivative \( \frac{L}{D_\mu} \sigma^\beta_\gamma_\mu \) can be made completely covariant also in the Einstein index \( \mu \), by going to

\[
D_\mu \Sigma_\gamma^\beta_\mu \equiv \frac{L}{D_\mu} \Sigma_\gamma^\beta_\mu - \Gamma_\mu^\nu \Sigma_\gamma^\beta_\nu.
\]

(18.12)
If \( \mu \) is contracted with \( \nu \) in the last term of cancels part of to the middle in (18.9), we obtain with \( D^\ast_\mu \) of Eq. (15.39):

\[
\frac{1}{2} D^\ast_\mu \Sigma^{\beta\gamma,\mu} = \frac{1}{2} \left[ \Theta^{\beta\gamma} - \Theta^{\gamma\beta} \right]. \tag{18.13}
\]

Multiplying this by \( h^\lambda_\beta h^\kappa_\gamma \), and moving the vierbeins to the right of the covariante
derivative, which is possible due to relation (17.61), we obtain

\[
\frac{1}{2} h^\lambda_\beta h^\kappa_\gamma D^\ast_\mu \Sigma^{\beta\gamma,\mu} - \Theta^{[\lambda,\kappa]} = \frac{1}{2} D^\ast_\mu \Sigma^{\lambda\kappa,\mu} - \Theta^{[\lambda,\kappa]} = 0. \tag{18.14}
\]

### 18.2 Energy-Momentum Density

Let us now deduce the consequence of local Einstein invariance. In this case the
spacetime coordinates must be transformed as well and the action is invariant in the
following sense:

\[
\mathcal{A} = \int d^4x \sqrt{\text{−}g(x)} L(h(x), A(x)) = \int d^4x' \sqrt{\text{−}g'(x')} L\left(h'(x'), A'(x')\right). \tag{18.15}
\]

If we change the variables \( x' \) to \( x \) in the second integral we see that the difference

\[
\int d^4x \left\{ \sqrt{\text{−}g'(x')} L\left(h'(x'), A'(x')\right) - \sqrt{\text{−}g(x)} L(h(x), A(x)) \right\} \tag{18.16}
\]

must be concentrated in the neighborhood of the surface of the integration volume. This is because the original integrations \( \int d^4x' \int d^4x \) covered the same
volume so that, after the change of variables \( x' \rightarrow x \), the first integral runs through a slightly
different region. Infinitesimally this amounts to the statement that

\[
\delta_E \mathcal{A} = \int d^4x \delta_E \left[ \sqrt{\text{−}g(x)} L(h(x), A(x)) \right] \tag{18.17}
\]

is a pure surface term. The symbol \( \delta_E \) denotes the substantial change under Einstein
transformations at fixed argument \( x \) [see (3.133),(11.79)], i.e.,

\[
\delta_E g_{\mu\nu}(x) = \bar{D}_\mu \xi_\nu(x) + \bar{D}_\nu \xi_\mu(x). \tag{18.18}
\]

Under Einstein transformations, the metric transforms as

\[
\delta_E \sqrt{\text{−}g} = -\frac{1}{2} \sqrt{\text{−}g} g_{\mu\nu} \delta_E g^{\mu\nu} = \frac{1}{2} \sqrt{\text{−}g} g^{\mu\nu} \delta_E g_{\mu\nu}, \tag{18.19}
\]

which, upon inserting (18.18), yields

\[
\frac{1}{2} \sqrt{\text{−}g} g^{\mu\nu} \left[ \xi^\lambda \partial_\lambda g_{\mu\nu} + (\partial_\mu \xi^\lambda) g_{\lambda\nu} + (\partial_\nu \xi^\lambda) g_{\mu\lambda} \right]. \tag{18.20}
\]
Therefore
\[ \delta_E \sqrt{-g} = \xi^\lambda \partial_\lambda \sqrt{-g} + \sqrt{-g} \partial_\lambda \xi^\lambda = \partial_\lambda \left( \xi^\lambda \sqrt{-g} \right) \]  
\[ (18.21) \]
and
\[ \delta_E \int d^4 x \sqrt{-g} = \int d^4 x \sqrt{g} D_\lambda \xi^\lambda = \int d^4 x \partial_\lambda \left( \xi^\lambda \sqrt{-g} \right). \]
\[ (18.22) \]
This shows that the trivial action \( \int d^4 x \sqrt{-g} \) indeed changes by a pure surface term. There is complete invariance if we require \( \xi^\lambda (x) \) to vanish at the surface.

The same result holds for a general action if \( L \) is a scalar Lagrangian density satisfying
\[ L' (x') = L (x) \]
\[ (18.23) \]
and therefore
\[ \delta_E L (x) \equiv L' (x') - L (x) = L' (x') - L (x) = \xi^\lambda \partial_\lambda L (x). \]
\[ (18.24) \]
The variation of \( A \) is
\[ \delta_E A = \int d^4 x \left( \delta_E \left( h^\mu_\alpha \right) + \frac{d A}{d h^\mu_\alpha} \delta_E h^\mu_\alpha \right) \]
\[ = \int d^4 x \left( \sqrt{-g} \Theta^\alpha_\mu \delta_E h^\mu_\alpha - \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma \delta_E A^\mu_\beta \right). \]
\[ (18.25) \]
We can now derive the covariant conservation law associated with Einstein invariance by using the substantial variations \( \delta_E h^\mu_\alpha \) and \( \delta_E A^\mu_\beta \) and calculating \( \delta_E A \) once more as follows:
\[ \delta_E A = \int d^4 x \left( \frac{\delta A}{d h^\mu_\alpha} \delta_E h^\mu_\alpha + \frac{d A}{d A^\mu_\beta} \delta_E A^\mu_\beta \right) \]
\[ = \int d^4 x \left( \sqrt{-g} \Theta^\alpha_\mu \delta_E h^\mu_\alpha - \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma \delta_E A^\mu_\beta \right). \]
\[ (18.26) \]
The substantial variations of the vierbein fields \( h^\mu_\alpha \) and \( A^\mu_\beta \) are those of a vector with a super- or subscripts \( \mu \) [recall (11.74), (11.75)]:
\[ \delta_E h^\mu_\alpha = \xi^\lambda \partial_\lambda h^\mu_\alpha - \partial_\alpha \xi^\mu h^\kappa_\kappa, \quad \delta_E A^\mu_\beta \gamma = \xi^\lambda \partial_\lambda A^\mu_\beta \gamma + \partial_\mu \xi^\lambda A^\lambda_\beta \gamma. \]
\[ (18.27) \]
Inserting these into (18.26), we find
\[ \delta_E A = \int d^4 x \left\{ \sqrt{-g} \Theta^\alpha_\mu \left( \xi^\lambda \partial_\lambda h^\mu_\alpha - \partial_\alpha \xi^\mu h^\kappa_\kappa \right) - \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma \left( \xi^\lambda \partial_\lambda A^\mu_\beta \gamma + \partial_\mu \xi^\lambda A^\lambda_\beta \gamma \right) \right\}. \]
\[ (18.28) \]
After partial integrations and letting $\xi^\lambda$ be zero everywhere, except for a $\delta$-function singularity at some place $x$, gives

$$\partial_\kappa \left( \sqrt{-g} \Theta_\lambda^\alpha h_\alpha^\kappa \right) + \sqrt{-g} \Theta_\mu^\alpha \partial_\lambda h_\alpha^\mu + \frac{1}{2} \partial_\mu \left( \sqrt{-g} \Sigma^\beta_\gamma^\mu A_{\lambda\beta}^\gamma \right) - \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu \partial_\lambda A_{\mu\beta}^\gamma = 0. \quad (18.29)$$

The second line can be rewritten as

$$\frac{1}{2} \partial_\mu \left( -\sqrt{-g} \Sigma^\beta_\gamma^\mu \right) A_{\lambda\beta}^\gamma + \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu \left( \partial_\mu A_{\lambda\beta}^\gamma - \partial_\lambda A_{\mu\beta}^\gamma \right). \quad (18.30)$$

If we introduce the covariant curl of the $A$ field,

$$F_{\mu\lambda\beta}^\gamma \equiv \partial_\mu A_{\lambda\beta}^\gamma - \partial_\lambda A_{\mu\beta}^\gamma - \left[ A_{\mu\beta}^\gamma A_{\lambda\delta}^\gamma - (\mu \leftrightarrow \lambda) \right], \quad (18.31)$$

then (18.30) becomes

$$\frac{1}{2} \partial_\mu \left( \sqrt{-g} \Sigma^\beta_\gamma^\mu \right) A_{\lambda\beta}^\gamma + \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu \left[ A_{\mu\beta}^\gamma A_{\mu\delta}^\gamma - (\mu \leftrightarrow \lambda) \right] + \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu F_{\mu\lambda\beta}^\gamma. \quad (18.32)$$

The first three terms in (18.29) can now be collected into a covariant derivative $D^*_\mu$ defined in (15.39), i.e.,

$$\frac{1}{2} \sqrt{-g} D^*_\mu \Sigma^\beta_\gamma^\mu A_{\lambda\beta}^\gamma, \quad (18.33)$$

so that the second line in (18.29) becomes, after using the conservation law (18.13),

$$-\sqrt{-g} \Theta_\gamma^\beta A_{\lambda\beta}^\gamma + \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu F_{\mu\lambda\beta}^\gamma. \quad (18.34)$$

In the first line of (18.31) we write

$$\Theta_\mu^\alpha \partial_\lambda h_\alpha^\mu = \Theta_\mu^\alpha D^L_\lambda h_\alpha^\mu + \Theta_\mu^\alpha A_{\lambda\alpha}^\beta h_\beta^\mu \quad (18.35)$$

and (18.29) takes the form

$$\partial_\kappa \left( \sqrt{-g} \Theta_\lambda^\kappa \right) + \sqrt{-g} \Theta_\mu^\alpha D^L_\lambda h_\alpha^\mu - \frac{1}{2} \sqrt{-g} \Sigma^\beta_\gamma^\mu F_{\mu\beta\gamma} = 0. \quad (18.36)$$

This equation is covariant under local Lorentz transformations but not yet manifestly so under Einstein transformations. In order to verify the latter we observe that the derivative $D^L_\lambda$ of $h$ can be rewritten as

$$D^L_\lambda h_\alpha^\mu = \partial_\lambda h_\alpha^\mu - A_{\lambda\alpha}^\beta h_\beta^\mu$$

$$= - \Gamma_\lambda^\mu_\kappa h_\alpha^\kappa - \left( \Gamma_\lambda^\mu_\sigma - \Gamma_\lambda^\mu_\sigma \right) h_\alpha^\sigma = - \Gamma_\lambda^\mu_\sigma h_\alpha^\sigma, \quad (18.37)$$
in accordance with the identity $D_\lambda h_\alpha^\mu = 0$. Then the second term is
\[ -\sqrt{-g} \Gamma^\mu_{\lambda\sigma} \Theta_\mu^\sigma. \] (18.38)

We now rewrite the first term as
\[ \sqrt{-g} (D^\ast_\lambda \Theta_\lambda^\kappa + \Gamma^\tau_{\kappa\lambda} \Theta_\tau^\kappa), \] (18.39)
and obtain the completely covariant conservation law for the energy momentum tensor $[1, 2, 3, 4, 5]$.

\[ D^\ast_\lambda \Theta_\lambda^\kappa + 2S^\tau_{\kappa\lambda} \Theta_\tau^\kappa - \frac{1}{2} \Sigma^\beta_\beta \mu F_{\lambda\mu\beta}^\gamma = 0. \] (18.40)

### 18.3 Covariant Derivation of Conservation Laws

The conservation laws of energy, momentum and angular momentum can be derived somewhat more efficiently, if some initial effort is spent in preparing the Einstein and local Lorentz transformations (18.27), (18.3), (18.4) of $h_\alpha^\mu$ and $A^\mu_\alpha$ in a covariant form. Take $\delta E h_\alpha^\mu$. It can be rewritten as
\[ \delta_E h_\alpha^\mu = \xi^\lambda \partial_\lambda h_\alpha^\mu + \Gamma^\mu_{\lambda\sigma} h_\alpha^\lambda \xi^\sigma - D_\lambda \xi^\mu h_\alpha^\lambda. \] (18.41)

Using the identity
\[ \partial_\lambda h_\alpha^\mu = -h_{\lambda\nu}^\mu h_\alpha^\nu = A^\beta_\lambda h_\beta^\mu - \Gamma^\mu_{\lambda\nu} h_\alpha^\nu, \] (18.42)
we can rewrite (18.41) in the covariant form
\[ \delta_E h_\alpha^\mu = -D_\alpha \xi^\mu + (A^\mu_\alpha + 2S^\mu_\alpha) \xi^\beta. \] (18.43)

The reciprocal vierbein field $h_\alpha^\mu$ transforms as
\[ \delta_E h_\alpha^\mu = D_\mu \xi^\alpha - (A^\alpha_\beta - 2S^\alpha_\beta) \xi^\beta. \] (18.44)

Similarly, we find
\[ \delta_E A^\mu_\alpha = \xi^\lambda \partial_\lambda A^\mu_\alpha + D_\mu \xi^\lambda A^\lambda_\alpha - \Gamma^\lambda_{\mu\kappa} A^\beta_\lambda \xi^\kappa = D_\mu \left( \xi^\alpha A^\mu_\alpha \right) - \xi^\lambda \left( D_\mu A^\lambda_\alpha - \partial_\lambda A^\mu_\alpha \right) - \Gamma^\lambda_{\mu\kappa} A^\beta_\lambda \xi^\kappa = D_\mu \left( \xi^\alpha A^\mu_\alpha \right) - \xi^\lambda F^\lambda_{\mu\alpha}. \] (18.45)

Under local Lorentz transformations, the vierbein field has already its simplest possible form,
\[ \delta_L h_\alpha^\mu = \omega_\alpha^\beta h_\alpha^\mu. \] (18.46)
while $A_{\mu\alpha}^{\beta}$ acquires the typical additive term of a gauge field
\[ \delta_L A_{\mu\alpha}^{\beta} = D_\mu \omega_\alpha^{\beta}. \] (18.47)

Using these covariant transformation rules, the variations of the action (18.6), (18.28) become
\[ \delta_L A = \int d^4x \sqrt{-g} \left\{ \Theta_\beta^{\alpha} \omega_\alpha^{\beta} h_\beta^\mu - \frac{1}{2} \Sigma_\beta^{\alpha} \mu D_\mu \omega_\alpha^{\beta} \right\}, \] (18.48)
\[ \delta_E A = \int d^4x \sqrt{-g} \left\{ \Theta_\mu^{\alpha} \left( -D_\lambda \xi^\mu h_\lambda^\alpha + (A_\lambda \mu - 2 S_\lambda \mu) \xi^\alpha \right) - \frac{1}{2} \Sigma_\beta^{\alpha} \mu \left[ D_\mu \left( \xi^\lambda A_\lambda^{\beta} \right) - \xi^\lambda F_\mu^{\lambda \beta} \right] \right\}. \] (18.49)

A partial integration of (18.48) [using (15.34), (15.38)] then gives directly the divergence of the spin current (18.13). A partial integration of (18.49) leads to
\[ D^\lambda \Theta_\mu^\lambda + \left( A_\mu^{\beta} - 2 S_\mu^{\beta} \right) \Theta_\beta^{\alpha} + \frac{1}{2} D^\nu \Sigma_\nu^{\alpha} \beta + \frac{1}{2} \Sigma_\beta^{\alpha} \nu F_\nu^{\mu \beta} = 0, \] (18.50)
which, after inserting (18.13), reduces correctly to the covariant, conservation law for the canonical energy-momentum tensor (18.40).

## 18.4 Matter with Integer Spin

If matter fields only carried integer spin it would not be necessary to introduce the $h_\alpha^{\mu \beta}$, $A_{\mu\alpha}^{\beta}$ fields. Then the theory could be formulated only with indices $\mu$ in an Einstein-invariant way. Let us derive the conservation of angular momentum for this situation. The action may be viewed as a functional of $g_{\mu\nu}$ and $K_{\mu\nu}^\lambda$, which enters via the affine connection $\Gamma_{\mu\nu}^\lambda = \bar{\Gamma}_{\mu\nu}^\lambda + K_{\mu\nu}^\lambda$. Einstein invariance implies the vanishing of the symmetry variation
\[ \delta_E A = \int d^4x \left( \frac{\delta A}{\delta g_{\mu\nu}}|_{g_{\mu\nu}} \delta_E g_{\mu\nu} + \frac{\delta A}{\delta K_{\mu\nu}^\lambda}|_{K_{\mu\nu}^\lambda} \delta_E K_{\mu\nu}^\lambda \right) \]
\[ = -\frac{1}{2} \int d^4x \sqrt{-g} \left\{ T_{\mu\nu} \left( \xi^\lambda \partial_\lambda g_{\mu\nu} + \partial_\mu \xi^\lambda g_{\lambda\nu} + \partial_\nu \xi^\lambda g_{\mu\lambda} \right) + \Sigma_\nu^{\mu \kappa} \left( \xi^\lambda \partial_\lambda K_{\mu\nu}^\kappa + \partial_\mu \xi^\lambda K_{\lambda\nu}^\kappa + \partial_\nu \xi^\lambda K_{\mu\lambda}^\kappa - \partial_\lambda \xi^\kappa K_{\mu\nu}^\lambda \right) \right\}, \] (18.51)
if the Euler-Lagrange equations are fulfilled. Here we have used the definitions (15.14) and (15.16) of the energy momentum tensor and the current density, and inserted the infinitesimal transformation laws (11.77) of $g_{\mu\nu}$ and (11.107) of $S_{\mu\nu}^\lambda$ (which holds also for $K_{\mu\nu}^\lambda$). We have omitted the superscripts $m$ since the equations in this section apply just as well to the gravitational field action $\hat{A}$, if we use the definitions (15.15) and (15.17). The further calculations are simplified by defining the symmetrized canonical energy-momentum tensor as follows:
\[ \frac{\delta A}{\delta g_{\mu\nu}}|_{g_{\mu\nu}} = \text{const.} \equiv -\frac{1}{2} \sqrt{-g} \left( \Theta_{\mu\nu} + \Theta_{\nu\mu} \right). \] (18.52)
It is easy to see that this definition agrees with (17.135) if there are no spin $-1/2$ fields. This is done by differentiating $A$ with respect to $h_{\alpha}^{\mu}$ at fixed $A_{\mu\alpha}$ and changing the index $\alpha$ to $\nu$). It may also be verified by forming

$$\frac{\delta A}{\delta g_{\mu\nu}(x)}|_{\Gamma_{\nu\lambda}^{\lambda}=\text{const.}} = \frac{\delta A}{\delta g_{\mu\nu}(x)}|_{\Gamma_{\mu\nu}^{\lambda}=\text{const.}} + \int dy \frac{\delta A}{\delta \Gamma_{\sigma\tau}^{\lambda}(y)}|_{g_{\nu\mu}=\text{const.}} \frac{\delta \Gamma_{\sigma\tau}^{\lambda}(y)}{\delta g_{\mu\nu}(x)}|_{\Gamma_{\nu\lambda}^{\lambda}=\text{const.}},$$

so that one obtains the standard Belinfante relation (17.148) between $T_{\mu\nu}$ and $\Theta_{\mu\nu}$:

$$T_{\mu\nu} = \Theta_{\mu\nu} - \frac{1}{2} \partial_\lambda (\Sigma^{\mu\lambda\nu} - \Sigma^{\nu\lambda\mu} + \Sigma^{\lambda\mu\nu}).$$

(18.54)

For pure gravity, (18.52) is in accord with (17.139) which states that $\Theta_{\mu\nu}$ is the Einstein tensor [recall (17.139)] up to a factor $-\kappa$

$$-\kappa \Theta_{\mu\nu} = G_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R,$$

as can be seen from (15.27) and the Belinfante relation (17.148) again coincides with (15.57).

Thus we can evaluate the consequences of Einstein invariance by using $\Theta$ and $\Sigma$ and considering, instead of (18.51), the variation

$$0 = \delta E A = \int d^4 x \left\{ \frac{\delta A}{\delta g_{\mu\nu}(x)} \delta E g_{\mu\nu} + \frac{\delta A}{\delta \Gamma_{\mu\nu}^{\lambda}} \delta E \Gamma_{\mu\nu}^{\lambda} \right\}$$

$$= -\frac{1}{2} \int d^4 x \sqrt{-g} \left\{ \Theta_{\mu\nu} \xi^\mu \left( \xi^\lambda \partial_\lambda g_{\mu\nu} + \partial_\mu \xi^\lambda g_{\lambda\nu} + \partial_\nu \xi^\lambda g_{\mu\lambda} \right) - \Sigma^\nu_\mu \left( \xi^\lambda \partial_\lambda \Gamma_{\mu\nu}^{\kappa} + \partial_\mu \xi^\lambda \Gamma_{\nu\lambda}^{\kappa} + \partial_\nu \xi^\lambda \Gamma_{\lambda\mu}^{\kappa} - \partial_\lambda \xi^\lambda \Gamma_{\mu\nu}^{\kappa} + \partial_\mu \partial_\nu \xi^\kappa \right) \right\}. (18.55)$$

It is again useful to bring the variations $\delta E g_{\mu\nu}, \delta E \Gamma_{\mu\nu}^{\lambda}$ into covariant form. We rewrite the Einstein variation of the metric as

$$\delta E g_{\mu\nu} = D_\mu \xi_\nu + D_\nu \xi_\mu = D_\mu \xi_\nu + D_\nu \xi_\mu + \left[ K_{\mu\nu}^{\lambda} + (\mu \leftrightarrow \nu) \right] \xi_\lambda$$

$$= D_\mu \xi_\nu + D_\nu \xi_\mu + 2 \left[ S_{\lambda\mu\nu} + (\mu \leftrightarrow \nu) \right] \xi_\lambda,$$

(18.56)

and the variation of the connection

$$\delta E \Gamma_{\mu\nu}^{\kappa} = D_\mu D_\nu \xi^\kappa - 2 D_\mu \left( S_{\nu\lambda}^{\kappa} \xi_\lambda \right) + R_{\lambda\mu\nu}^{\kappa} \xi_\lambda.$$

(18.57)

Inserting this into (18.53) gives

$$\delta E A = \int d^4 x \sqrt{-g} \left\{ (\Theta^{\mu\nu} + \Theta^{\nu\mu}) \left( D_\nu \xi_\mu + 2 S_{\lambda\mu\nu} \right) \xi^\lambda + \Sigma^\nu_\mu \left[ D_\mu D_\nu \xi^\kappa - 2 D_\mu \left( S_{\nu\lambda}^{\kappa} \xi_\lambda \right) + R_{\lambda\mu\nu}^{\kappa} \xi_\lambda \right] \right\}.$$

H. Kleinert, GRAVITY WITH TORSION
18.5 Relations between Conservation Laws and Bianchi Identities

By partially integrating the Σ term and using the spin divergence law (18.13), we obtain immediately

$$\delta E_A = 2 \int d^4x \sqrt{-g} \left\{ -D_\mu \Theta^\mu_\lambda - 2S_\mu_\lambda \nu \Theta^\nu_\nu + \frac{1}{2} \Sigma^{\nu_\kappa_\mu} \Sigma^\mu_\kappa_\rho R_{\lambda_\mu_\rho}^\kappa \right\} \xi^\lambda, \quad (18.58)$$

leading directly to the covariant conservation law

$$D^*_\mu \Theta^\mu_\lambda + 2S_\mu_\lambda \nu \Theta^\nu_\nu - \frac{1}{2} \Sigma^{\nu_\kappa_\mu} \Sigma^\mu_\kappa_\rho R_{\lambda_\mu_\rho}^\kappa = 0. \quad (18.59)$$

This is not in manifest agreement with (18.40) since the last term is $\Sigma^{\nu_\kappa_\mu} R_{\lambda_\mu_\rho}^\kappa$, while we had $\Sigma^{\beta_\gamma_\mu} S_{\lambda_\gamma}^\beta \rho_\mu$, which is the same due to Eq. (17.75).

When expressing the energy-momentum tensor and the spin-current density in terms of the Einstein and Palatini tensors $G_{\mu_\nu} = R_{\mu_\nu} - g_{\mu_\nu} R$ and $(1/2)S^{\nu_\kappa_\mu} = S^{\nu_\kappa_\mu} + g^{\nu_\mu} S_{\kappa}^\rho - \delta^{\nu_\mu} S^\rho$, the two covariant conservations laws (18.13) and (18.59) of a pure gravitational field take the form [recall (15.54), (17.139)]

$$\frac{1}{2} D^*_\mu S^{\lambda_\kappa_\mu} = G^{[\lambda,\kappa]}, \quad (18.60)$$

$$D^*_\mu G^\mu_\lambda = 2S^\nu_\mu S^{\lambda_\kappa_\nu} + 1/2 S^{\nu_\kappa_\mu} R_{\lambda_\mu_\nu}^\kappa = 0. \quad (18.61)$$

18.5 Relations between Conservation Laws and Bianchi Identities

For the gravitational field itself, both covariant laws are automatically satisfied irrespective of the presence of matter due to the fundamental identity (12.103) and the Bianchi identity (12.115). To see this we apply (15.39) to (15.45) and obtain

$$\frac{1}{2} D^*_\lambda S^{\nu_\mu_\lambda} = D^*_\lambda (S_{\nu_\mu_\lambda} + S^{\mu_\nu} S_{\mu_\lambda} - S^{\rho_\lambda} S_{\mu_\rho}) = D^*_\lambda S^{\mu_\nu} + D^*_\nu S_{\mu_\lambda} - D^*_\mu S_{\nu_\lambda} = D^*_\lambda S^{\mu_\nu} + D^*_\nu S_{\mu_\lambda} - D^*_\mu S_{\nu_\lambda} + 2S_{\nu_\mu_\lambda} \chi. \quad (18.62)$$

Now we take (12.103) and contract the subscript \( \nu \) with the superscript \( \kappa \) to obtain

$$R_{\mu_\kappa} - R_{\kappa_\mu} = 2 \left( D^*_\kappa S_{\nu_\mu_\kappa} + D^*_\mu S_{\nu_\kappa} + D^*_\nu S_{\mu_\kappa} \right) - 4 \left( S_{\rho_\mu_\kappa} S_{\lambda_\rho} - S_{\rho_\mu_\kappa} S_{\nu_\nu} + S_{\lambda_\mu} S_{\rho_\rho} - S_{\lambda_\mu} S_{\nu_\nu} \right) = 2 \left( D^*_\kappa S_{\nu_\mu_\kappa} + D^*_\mu S_{\nu_\kappa} - D^*_\nu S_{\mu_\kappa} \right) + 4S_{\rho_\mu_\kappa} S_{\nu_\nu} = D^*_\lambda S^{\nu_\mu_\lambda} \chi. \quad (18.63)$$

Since $R_{\mu_\kappa}$ differs from $G_{\mu_\kappa}$ only by the symmetric tensor $g_{\mu_\kappa}/2$, the same equation holds for $G_{\mu_\kappa} - G_{\lambda_\kappa}$, so that we find

$$D^*_\lambda S_{\nu_\mu_\lambda} = G_{\nu_\mu} - G_{\mu_\nu} \quad (18.64)$$

in agreement with (18.60).
Similarly, using (12.115) and permuting the indices we have

\[ D_\tau R_{\sigma \nu \mu} + D_\sigma R_{\nu \tau \mu} + D_\nu R_{\tau \sigma \mu} = 2 S_{\tau \sigma} R_{\nu \lambda \mu}^\lambda + 2 S_{\nu \sigma} R_{\tau \lambda \mu}^\lambda + 2 S_{\nu \tau} R_{\sigma \lambda \mu}^\lambda. \]  

(18.65)

Contracting \( \nu \) and \( \mu \), this becomes

\[ 2D_\tau R_{\sigma}^\tau - D_\sigma R^\tau = 2D_\tau G_{\sigma}^\tau = -2S_{\tau \sigma}^\lambda + 2S_{\nu}^\mu R_{\lambda \mu}^\lambda + 2S_{\nu}^\mu R_{\lambda \mu}^\tau \]

\[ = -4S_{\tau \sigma}^\lambda R_{\lambda}^\tau + 2S_{\nu}^\mu R_{\lambda \mu}^\tau \]  

(18.66)

or

\[ D^\mu G_{\sigma}^\mu - 2S_{\mu} \left( R_{\lambda}^\mu - \frac{1}{2} \delta_{\lambda}^\mu R \right) + S_{\tau \sigma} \left( G_{\lambda}^\tau + \frac{1}{2} \delta_{\lambda}^\tau R \right) - S_{\nu}^\lambda R_{\sigma \lambda \mu}^\tau = 0, \]  

(18.67)

in agreement with (18.61).

Within the defect interpretation of curvature and torsion, we observed before that the fundamental identities are nonlinear generalizations of the conservation laws of defect densities. From what we have just learned, the same equation can be obtained as conservation laws of energy-momentum and angular momentum from an Einstein action.

The two laws follow from the invariance of the Einstein action under general coordinate transformations, which are local translations, and under local Lorentz transformations, respectively.

These transformations correspond to elastic deformations (translational and rotational) of the world crystal and the invariance of the action expresses the fact that elastic deformations do not change the defect structure.

It is important to realize that due to the intimate relationship between the conservation laws and the fundamental identities for the gravitational fields, they remain valid in the presence of any matter distribution. Then, by the field equations (17.128), (17.149), the spin density and energy-momentum tensor of the matter fields have to satisfy the same divergence laws by themselves. Indeed, it can easily be seen that this is a direct consequence of the Einstein invariance of the matter action in an arbitrary but fixed affine space, i.e., in a space whose geometry is specified from the outset rather than being determined by the matter fields via the field equations.

18.6 Particle Trajectories from Energy-Momentum Conservation

The classical equations of motion for a point particle have the consequence that its energy-momentum tensor \( T^\mu_\lambda = 0 \) in Eq. (15.23) satisfies the covariant conservation law (18.40) all by itself. Otherwise the Einstein equations (18.52) would not be satisfied. Since the symmetric energy-momentum tensors \( T^m_\lambda \mu = 0 \) and \( T^f_\lambda \mu = 0 \) of gravitational field and matter are proportional to each other, they must separately satisfy the covariant conservation law.
Consider first a particle without spin in a space without torsion. Starting point is the covariant conservation law (18.40) which reads now

$$\bar{D}_\kappa T^\kappa_\lambda(x) = 0.$$  \hfill (18.68)

Expressing the covariant derivative in terms of the Riemann connection, and this further in terms of derivatives of the metric using the identity

$$\frac{1}{\sqrt{-g}} \partial_\nu \sqrt{-g} = \frac{1}{2} g^{\lambda\kappa} \partial_\nu g_{\lambda\kappa} = \bar{\Gamma}_\nu^{\lambda\kappa},$$ \hfill (18.69)

Eq. (18.68) becomes

$$\partial_\nu [\sqrt{-g} T^{\mu\nu}(x)] + \sqrt{-g} \bar{\Gamma}_\nu^{\mu\lambda}(x) T_{\lambda\nu}(x) = 0.$$ \hfill (18.70)

This must hold for the energy-momentum tensor of the particle trajectories (15.23). Inserting this gives

$$m \int d\tau [\dot{q}^\mu(\tau) \dot{q}^\nu(\tau) \partial_\nu \delta^{(4)}(x - q(\tau)) + \bar{\Gamma}_\nu^{\mu\lambda}(q) \dot{q}^\nu(\tau) \dot{q}^\lambda(\tau) \delta^{(4)}(x - q(\tau))] = 0.$$ \hfill (18.71)

The first term in the integrand can also be written as $-\dot{q}^\mu(\tau) \partial_\nu \delta^{(4)}(x - q(\tau))$, so that a partial integration leads to

$$m \int d\tau [\ddot{q}^\mu(\tau) + \bar{\Gamma}_\nu^{\mu\lambda}(q) \dot{q}^\nu(\tau) \dot{q}^\lambda(\tau)] \delta^{(4)}(x - q(\tau)) = 0.$$ \hfill (18.72)

Integrating this over a thin tube around the trajectory $q^\mu(\tau)$, we obtain the equation (11.25) for the geodesic trajectory [6].

The same result may be derived from the following consideration. According to Eq. (18.58), the variation of the action under Einstein transformations of the coordinates $\delta_E x^\mu = -\xi^\mu$ is in Riemannian space

$$\delta_E A = -2 \int d^4 x \sqrt{-g} \bar{D}_\mu T^{\mu\lambda}(x) \xi^\lambda.$$ \hfill (18.73)

Due to Einstein’s equation (15.61) this holds separately for field and matter parts. If matter consists of point particles only, we obtain:

$$\delta_E A = -\int d\tau \frac{m}{\delta q^\mu(\tau)} \xi^\mu(q(\tau)).$$ \hfill (18.74)

This vanishes along the geodesic trajectory implying that the energy-momentum tensor is covariantly conserved.

Let us now allow for torsion in the curved spacetime, where the covariant conservation law (18.68) becomes (18.59):

$$D^\mu \Theta^\lambda_\mu + 2 S^{\mu\lambda}_\nu \Theta^\nu_\mu - \frac{1}{2} \Sigma^\lambda_\nu \kappa R_{\lambda\mu\nu}^\kappa = 0.$$ \hfill (18.75)
Recalling the Belinfante relation (18.54) and the definition of $D^\kappa_\mu$ in Eq. (15.39), we obtain for scalar particles with $\Sigma^\nu_\kappa J^\mu = 0$:

$$\bar{D}_\mu T^\mu_\lambda = 0.$$  \hspace{1cm} (18.76)

This coincides with the conservation law (18.68) in Riemannian space and leads once more the geodesic trajectories (11.25).

How can we remove the discrepancy with respect to the autoparallel trajectory found by the multivalued mapping procedure in Eq. (14.7)? Apparently, the variational procedure of the metric which has led to the Einstein equation (15.61) must be modified to account for the torsion. This question is further discussed in Subsection 25.2.1.

Notes and References

Gravitation of Spinning Matter as a Gauge Theory

The alert reader will have noticed by now that the theory of gravity of spinning matter, when formulated in terms of fields $h_\alpha^\mu$, $A_\mu^\alpha$ introduced in Eqs. (17.17) and (18.31), is really a gauge theory of local Lorentz transformations. Gauge properties have become apparent before in when we observed in Eq. (11.105) that the connection $\Gamma^\lambda_{\mu\nu}$ transforms like a nonabelian gauge field under general coordinate transformations. But at that early stage, we could not have really spoken about a gauge theory since the connection $\Gamma^\lambda_{\mu\nu}$ was not an independent field of the system. When introducing spinning particles, the metric $g_{\mu\nu}(x)$ as a fundamental field was replaced by the vierbein field $h_\alpha^\mu(x)$ which transforms like a gauge field under translations. The Dirac theory in curved space, in which the covariant derivative contains the spin connection $A_\alpha^\beta\gamma$ of Eq. (17.66) with only the objects of anholonomity and no torsion tensor $K^\mu_{\lambda\nu}0$, is a bona fide gauge theory of local Einstein and Lorentz transformations. If the space has also torsion, the spin connection $A_\alpha^\beta\gamma$ becomes a gauge field which is completely independent of the gauge field $h_\alpha^\mu$. Let us study the properties of such a theory in more detail.

19.1 Local Lorentz Transformations

Recall that under infinitesimal Lorentz transformations a vector field behaves like [see Eq. (17.102)]

$$\delta_L v_\alpha(x) = \omega_\alpha^\beta v_\beta(x), \quad \delta_L v^\alpha(x) = \omega^\alpha_\beta v^\beta(x),$$

(19.1)

where the physical coordinates $x^\mu$ remain unchanged since only the local directions are transformed. Due to the antisymmetry of the matrix $\omega$ this can also be written as

$$\delta_L v^\alpha(x) = -v^\beta(x)\omega^\alpha_\beta.$$  

(19.2)

This shows that the transformation law (I.3.108b) coincides precisely with (18.4) for the special case of the local Lorentz group:

$$\delta_L A_\mu^\gamma = \omega_\mu^\beta A_\mu^\beta\gamma + \omega^\gamma_\gamma A_\mu^\beta\gamma + \partial_\mu \omega_\beta^\gamma.$$  

(19.3)
Observe that the spacetime variables $x^\mu$ are not transformed so that the the Lorentz group acts like an internal symmetry group. There is, however, a certain similarity with external gauge symmetries discussed in Section 3.5, Part I. This is because $h^\alpha_{\mu}$ can couple Lorentz and Einstein indices, just as in (I.3.135), thus giving rise to more invariants. For instance, there is no need of forming $(F_{\mu\nu})^2$ in order to get an invariant action. There also exists an invariant expression linear in the field strength,

$$A_f = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} h^{\alpha\mu} h^\beta_\nu F_{\nu\mu\alpha}.$$  \hfill (19.4)

In fact, from (17.75) this is just the Einstein-Cartan action (15.8).

For completeness, let us see once more how the spin current and energy-momentum tensor follow from this action with independent fields $h^\alpha_{\mu}, A_{\mu\alpha\beta}$. First we calculate the spin current of the field. By definition,

$$\frac{1}{2} \sqrt{-g} \sum_{\alpha\beta} \Sigma^{\alpha\beta}_{\mu}(x) = -\frac{\delta f_A}{\delta A_{\mu\alpha\beta}(x)} (19.5)$$

$$= \frac{1}{2\kappa} \int d^4x \sqrt{-g} h^{\alpha\mu} h^\beta_\nu \left( \partial_{\mu} A_{\nu\alpha\beta} - \partial_{\nu} A_{\mu\alpha\beta} - A_{\nu\alpha} \gamma A_{\mu\gamma} + A_{\mu\alpha} \gamma A_{\nu\beta} \right)$$

$$= -\frac{1}{2\kappa} \left\{ \partial_\nu \sqrt{-g} \left[ h^\alpha_{\mu} h^\beta_\nu - (\alpha \leftrightarrow \beta) \right] + \sqrt{-g} \left[ (A_{\nu\alpha'} h^\alpha_{\mu} h^\beta_\nu + A_{\nu\beta'} h^\alpha_{\mu} h^\beta_\nu) - (\alpha \leftrightarrow \beta) \right] \right\}. $$

We may write this in terms of the partially covariant derivatives (18.10), (18.11) as

$$-\kappa \sum_{\alpha\beta} \Sigma^{\alpha\beta}_{\mu} = L^\nu \left[ h^\alpha_{\mu} h^\beta_\nu - (\alpha \leftrightarrow \beta) \right] + \Gamma^\nu_{\sigma} \left[ h^\alpha_{\mu} h^\beta_\nu - (\alpha \leftrightarrow \beta) \right].$$

Applying the chain rule of differentiation this becomes

$$-\kappa \sum_{\alpha\beta} \Sigma^{\alpha\beta}_{\mu} = \left( L^\nu \left[ h^\alpha_{\mu} h^\beta_\nu - (\alpha \leftrightarrow \beta) \right] + \Gamma^\nu_{\sigma} \left[ h^\alpha_{\mu} h^\beta_\nu - (\alpha \leftrightarrow \beta) \right] \right) - (\alpha \leftrightarrow \beta).$$ \hfill (19.6)

We now observe that, due to the identity $D^\mu h^\mu_{\alpha\nu} \equiv 0$, the connection can be rewritten as

$$\Gamma^\lambda_{\mu\nu} = h^{\alpha\lambda} L^\mu h^\alpha_{\nu} = -h^{\alpha\nu} L^\mu h^\alpha_{\lambda}. \hfill (19.7)$$

This relation is complementary to the relation $\bar{\Gamma}^\gamma_{\mu\nu} = h^\gamma_{\nu} L^\mu h^\mu_{\lambda}$ of Eq. (17.141). Using (19.7), the spin current of the field becomes

$$-\kappa \sum_{\alpha\beta} \Sigma^{\alpha\beta}_{\mu} = 2 \left( S_{\alpha\beta}^\mu + h^\mu_{\alpha} S_{\beta} - h^\mu_{\beta} S_{\alpha} \right) = S_{\alpha\beta}^\mu,$$ \hfill (19.8)
We now calculate the functional derivative of the action with respect to $h_{\alpha}^{\mu}$. It shows directly that the canonical energy-momentum tensor of the gravitational field coincides with the Einstein tensor,

$$
\sqrt{-g} \Theta_{\mu}^{\alpha} = \left( \frac{\delta A_f}{\delta h_{\alpha}^{\mu}} \right) = \sqrt{-g} \left( h^{\delta\nu} F_{\mu\nu\delta}^{\alpha} - h^{\beta}_{\mu} F_{\beta\delta}^{\delta\alpha} \right)
$$

$$
\sqrt{-g} \left( R_{\mu\nu} - g_{\mu\nu} R \right) = \sqrt{-g} G_{\mu}^{\alpha}.
$$

(19.9)

The use of the field $h_{\alpha}^{\mu}$ has made it possible to retrieve the Einstein tensor without projecting out the symmetric part of it, as in the previous formulas, (15.27) and (18.52).

### 19.2 Local Translations

In the literature one often finds the statement that the vierbein field may be considered as a *gauge field of local translations*. In fact, Einstein’s transformations

$$
x' = x - \xi(x)
$$

(19.10)

can be considered as local translations and the vierbein field does ensure that the theory is invariant under these, just as any *bona fide* gauge field is supposed to. The covariant derivative

$$
D_{\alpha} \equiv h_{\alpha}^{\mu} \partial_{\mu} + \frac{i}{2} A_{\alpha\beta} \Sigma_{\gamma}^{\beta}
$$

may be viewed as a combination of $h_{\alpha}^{\mu}$ times the translational “functional matrix” $\partial_{\mu}$ and $(i/2) A_{\alpha\beta} \Sigma_{\gamma}^{\beta}$ times the Lorentz matrix $\Sigma_{\gamma}^{\beta}$. This viewpoint becomes most transparent by considering the expression in (17.83), the commutator of two covariant derivatives with respect to the dislocation coordinates,

$$
[D_{\alpha}, D_{\beta}] \psi = \frac{i}{2} F_{\alpha\beta\gamma} \Sigma_{\gamma}^{\delta} \psi + i2S_{\alpha\beta} \gamma iD_{\gamma} \psi.
$$

(19.12)

Since the factor of $F_{\alpha\beta\gamma}$ is the curl of the gauge field of Lorentz transformations, the factor $2S_{\alpha\beta} \gamma$ of $D_{\gamma} \psi$ may be considered as the curl of the gauge field of translations. Indeed, if we write $2S_{\alpha\beta} \gamma$ in the form

$$
2S_{\alpha\beta} \gamma = -h^{\gamma}_{\nu} \left[ h_{\alpha}^{\mu} \frac{\partial}{\partial \nu} h_{\beta}^{\nu} - (\alpha \leftrightarrow \beta) \right]
$$

$$
= h_{\alpha}^{\mu} h_{\beta}^{\nu} \left[ \frac{\partial}{\partial \mu} h^{\gamma}_{\nu} - (\mu \leftrightarrow \nu) \right],
$$

(19.13)

we arrive at the standard form of a curl, and the present formulation of gravity of spinning matter can be considered as a gauge theory of both local Lorentz transformations and local translations.
If the space has no torsion, then $A_{µα}^{β}$ is completely composed of derivatives of vierbein fields [recall Eq. (17.63)]

$$A_{αβ}^{γ} = -h_{γ}^{λ}h_{α}^{μ}h_{β}^{ν} h_{ρ}^{λ} Γ_{μν}^{λ}. \tag{19.14}$$

Inserting this into (19.13) we verify that this is equivalent to a vanishing torsion.

In recent years, this aspect of gravitational theory has received increasing attention, due to the shift in emphasis from geometric principles to gauge principles.

**Notes and References**

For more details on general relativity, the reader may consult

The mathematics of metric-affine spaces is discussed in
whose notation we use.

The gauge aspects of gravity are discussed in
(http://www.physik.fu-berlin.de/~kleinert/b1/contents2.html)

The symmetric energy-momentum tensor was first constructed by
F.J. Belinfante, *Physica* 6 (1939) 887,
Linearized Einstein Gravity of Point Particles

The gravitational field equations (1.41) and (17.130) are complicated non-linear equations and hard to solve in general. For many applications, the non-linearities are irrelevant. The deviations of the metric $g_{\mu\nu}$ from the Minkowski-form $g_{ab}$ in Eq. (1.29) are so small that a linear approximation is perfectly adequate. We shall neglect torsion and find that the linearized field equations acquire great similarity to Maxwell’s equations.

20.1 Action and Equation of Motion

In the linearized theory, the deviations of the metric from the Minkowski metric can be considered as an ordinary tensor field in Minkowski space. To define this correctly, we shall from now on use the notation $\eta_{ab}$ for the Minkowski metric $g_{ab}$ in Eq. (1.29). Then the curved-space metric $g_{\mu\nu}(x)$ is written as

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x). \quad (20.1)$$

In order to emphasize the tensor transformation properties in Minkowski space, we shall relabel the field $h_{\mu\nu}(x)$ with Minkowski indices $h_{ab}(x)$. As will be seen below, this tensor field possesses gauge properties very similar to those of the vector potential $A_a(x)$ of electromagnetism. The gauge transformations which leave physical phenomena invariant are linearized remnants of the infinitesimal coordinate transformations.

If point particles are included into the linearized gravitational action, the linearized theory becomes very similar to the Maxwell-Lorentz theory of electrons in an electromagnetic field. It is worth studying this structure in detail. In order to stress the analogy we shall perform the linear approximation right at the level of the action and derive the field equations from the extremal principle. Starting point is the action (15.11) for the gravity of point particle:

$$A = A + \mathcal{A} = -mc \int_{\tau_a}^{\tau_b} d\tau \sqrt{g_{\mu\nu}(\tau)q^{\mu}(\tau)q^{\nu}(\tau)} - \frac{1}{2\kappa} \int d^4x \sqrt{-g}R. \quad (20.2)$$
Inserting the decomposition (20.1) written with Minkowski indices, \( g_{ab} = \eta_{ab} + h_{ab} \), the linearized matter action becomes
\[
\mathcal{A}^m = -mc \int_{\tau_a}^{\tau_b} d\tau \sqrt{\eta_{ab} \dot{q}^a(\tau) \dot{q}^b(\tau)} - \frac{m}{2} \int_{\tau_a}^{\tau_b} d\tau h_{ab}(q(\tau)) \dot{q}^a(\tau) \dot{q}^b(\tau).
\] (20.3)

In order to calculate the linearized field action we observe that the Christoffel symbols (11.22) have the linear approximation
\[
\bar{\Gamma}^c_{ab} \approx \frac{1}{2} \left( \partial_a h^c_b + \partial_b h^c_a - \partial^c h_{ab} \right).
\] (20.4)

With this, the curvature tensor (11.144) becomes
\[
R_{abcd} \approx \frac{1}{2} \left[ \partial_a \partial_c h_{bd} - \partial_b \partial_d h_{ac} - (a \leftrightarrow b) \right],
\] (20.5)

and the Ricci tensor (11.140):
\[
R_{ad} \approx \frac{1}{2} \left( \partial_a \partial_c h_{cd} + \partial_d \partial_c h_{ca} - \partial_a \partial_d h - \partial^2 h_{ad} \right),
\] (20.6)

where \( h \) is defined to be the trace of the tensor \( h_{ab} \), i.e., \( h \equiv h_{cc} \). The ensuing scalar curvature (11.141) reads
\[
R \approx - (\partial^2 h - \partial_a \partial_b h^{ab}),
\] (20.7)

so that the Einstein tensor (11.142) becomes
\[
G^{ad} = R^{ad} - \frac{1}{2} g^{ad} R \approx - \frac{1}{2} \left( \partial^2 h^{ad} + \partial^a \partial^d h - \partial^a \partial_c h^{cd} - \partial^d \partial_c h^{ca} \right) + \frac{1}{2} \eta^{ad} (\partial^2 h - \partial_a \partial_b h^{ab}).
\] (20.8)

Since torsion is zero by assumption, we shall admit the bars on top of all Riemannian quantities.

The right-hand side of (27.28), which we shall denote by \( G^{ad}_{(1)} \) to emphasize its linear dependence on \( h_{ab} \), can be written as a four-dimensional version of a double curl
\[
G^{ad}_{(1)} = \frac{1}{2} \epsilon^{abef}_{\ h} \partial_e \partial_f h_{cd}.
\] (20.9)

This can easily be verified using the well known identity for the product of the \( \epsilon \)-tensors:
\[
\epsilon_{abcd} \epsilon_{efgh} = \sum_P \epsilon_P \delta_{aP(c)} \delta_{bP(f)} \delta_{cP(g)} \delta_{dP(h)},
\] (20.10)

where the sum runs over all permutations \( P \) of the second indices, with a minus sign \( \epsilon_P = -1 \) for odd permutations.
The linear expressions for $R$ are not yet sufficient to calculate the Einstein action up to second order in $h_{ab}$. In fact, the linear pieces are pure surface terms and do not contribute at all to the equation of motion. To find these we shall rewrite $\sqrt{-g} R$ in such a way that it exhibits the quadratic terms in the connection $\Gamma^{\lambda}_{\mu\nu}$. The appropriate formula was given in (11.151):

$$\sqrt{-g} R = \partial_\lambda \left[ (g^{\mu\nu} \sqrt{-g}) \left( \Gamma^\lambda_{\mu\nu} - \delta_\mu^\lambda \Gamma^\kappa_{\nu\kappa} \right) \right] + \sqrt{-g} g^{\mu\nu} \left( \Gamma^\lambda_{\mu\nu} \Gamma^\kappa_{\nu\kappa} - \Gamma^\lambda_{\mu\nu} \Gamma^\kappa_{\nu\lambda} \right). \quad (20.11)$$

The first term is a pure divergence, and may be omitted in the action, keeping only the second term. After inserting (27.23), setting $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, $g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$, and going to Minkowski indices, this becomes directly the desired quadratic action in $h_{ab}$.

$$\mathcal{A} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} R \quad (20.12)$$

$$\approx \frac{1}{8\kappa} \int d^4x \left\{ \left( \partial_a h^a - \partial^e h_{ae} - \partial^f h_{af} \right) \left( \partial_a h^b + \partial^a h_b - \partial^b h_a \right) - \left( \partial_a h^{ac} + \partial_{c} h^{ac} - \partial^a h^c \right) \left( \partial_a h_{bd} + \partial_{d} h_{bd} - \partial^d h_{ad} \right) \right\}. \quad (20.13)$$

The right-hand side can be rearranged, using the symmetry of $h_{ab}$, to

$$\mathcal{A} = \frac{1}{8\kappa} \int d^4x (\partial_a h^{bc} \partial_a h_{bc} - \partial^e h_{ae} \partial_{e} h_{bc} - \partial^a h^{bc} \partial_{d} h_{bd} + 2 \partial^a h^{ab} \partial_{b} h - \partial_a h \partial^a h). \quad (20.14)$$

A partial integration brings the third term to the same form as the second, apart from irrelevant terms, and we arrive at

$$\mathcal{A} = \frac{1}{8\kappa} \int d^4x (\partial_a h^{bc} \partial^a h_{bc} - 2 \partial^c h_{ab} \partial^a h^{bc} + 2 \partial^a h^{ab} \partial_{b} h - \partial_a h \partial^a h). \quad (20.14)$$

Another useful form of the action is the following, differing (20.14) by surface terms:

$$\mathcal{A} = \frac{1}{8\kappa} \int d^4x h_{ab} \epsilon^{acde} \epsilon^{bcef} \partial_d \partial_f h_{eg}. \quad (20.15)$$

The equivalence can be verified with the help of the identity

$$\epsilon^{acde} \epsilon^{bcef} h_{eg} = \delta^{ab} \delta^{df} \delta^{eg} + \delta^{af} \delta^{eg} \delta^{db} + \delta^{ag} \delta^{df} \delta^{eb} - \delta^{ab} \delta^{df} \delta^{eg} - \delta^{af} \delta^{dg} \delta^{eg} - \delta^{ag} \delta^{df} \delta^{eb}, \quad (20.16)$$

which follows from (20.10) by one contraction. Another way of writing the field action is

$$\mathcal{A} = \frac{1}{2} \int d^4x (\pi_{cab} \partial_c h_{ab}), \quad (20.17)$$

where $\pi_{cab}$ is the following tensor antisymmetric under the exchange $c \leftrightarrow a$, and symmetric in $a \leftrightarrow b$:

$$\pi_{cab} = \frac{1}{8\kappa} \left[ (\partial_c h_{ab} - \partial_a h_{cb}) (\eta_{cb} \partial_a h - \eta_{ab} \partial_c h) - \eta_{cb} \partial^e h_{ea} + \eta_{ab} \partial^e h_{ec} \right] + (a \leftrightarrow b). \quad (20.18)$$
The action (20.17) will be a convenient basis for our further discussion of the linearized theory of gravitation.

The equations of motion can now be derived in the standard manner. For the point particles in a gravitational field we find the canonical momentum from the Lagrangian (11.12). Watching out for the correct sign discussed in the derivation of (2.16), and using the proper time $\tau$ to parametrize the orbit, we obtain from the linearized action (20.3)

$$p_{\text{can}}^a(\tau) = -\frac{\partial m}{\partial q^a(\tau)} = m \left[ \dot{q}^a(\tau) + h^a_b(q(\tau)) \dot{q}^b(\tau) \right] = p^a(\tau) + h^a_b(q(\tau)) p^b(\tau).$$

(20.19)

This expression is reminiscent of electromagnetism, where the presence of an electromagnetic field modifies the canonical momentum of an electron from $p^a$ to [recall (2.131)]

$$P^a = p^a + \frac{e}{c} A^a.$$  

(20.20)

We shall therefore denote the canonical momentum $p_{\text{can}}^a$ also here by $P^a$.

The equation of motion for the canonical momentum in a weak gravitational field is then

$$\frac{d}{d\tau} \frac{\partial L}{\partial \dot{q}^a(\tau)} = \frac{\partial L}{\partial q^a(\tau)} = -\frac{1}{2} \delta^a_b h_{bc}(q(\tau)) \dot{q}^b(\tau) p^c(\tau).$$

(20.21)

Inserting on the left-hand side the first line of (20.19), and neglecting the term $h^a_b \ddot{q}^b$ since it is of second order in $h^a_b$, we obtain

$$\ddot{q}^a + \left( \partial_a h^a_b - \frac{1}{2} \delta^a_b h_{bc} \right) \dot{q}^b \dot{q}^c \equiv 0.$$  

(20.22)

This is is the gravitational analog of the Lorentz equation (2.129). Its solutions are the geodesic curves (14.7) in a weak gravitational field with linearized connection (27.23).

Let us now derive and study the linearized equations of motion for the gravitational field itself. We rewrite the interacting term in the matter action (20.3) as

$$\mathcal{A}_{\text{int}}^m = -\frac{m}{2} \int_{\tau_a}^{\tau_b} d\tau \int d^4 x h_{ab}(x) \dot{q}^a(\tau) \dot{q}^b(\tau) \delta^{(4)}(x - q(\tau)),$$

(20.23)

and differentiate this functionally with respect to $h_{ab}(x)$ to find the symmetric energy-momentum tensor of matter

$$\mathcal{T}_{\text{int}}^{ab}(x) = -2 \frac{\delta \mathcal{A}_{\text{int}}^m}{\delta h_{ab}(x)} = m \int d\tau \dot{q}^a(\tau) \dot{q}^b(\tau) \delta^{(4)}(x - q(\tau)).$$  

(20.24)
From the action (20.2) of the gravitational field, on the other hand, we find a symmetric energy-momentum tensor

\[ f_{T}^{ab} = -2 \frac{\delta A}{\delta h_{ab}} = 2\partial \pi_{cab} \]

\[ = \frac{1}{2\kappa} \left[ \partial^{2} h_{ab} - \partial^{a} \partial_{b} h_{bc} - \partial^{b} \partial_{c} h_{ac} + \partial_{a} \partial_{b} h - \frac{1}{2} \eta_{ab} (\partial^{2} h - \partial^{c} \partial^{d} h_{cd}) \right]. \]  

(20.25)

Comparison with (27.28) shows that this is precisely \(-1/\kappa\) times the linearized Einstein tensor, which we shall denote by

\[ f_{T}^{ab} = \frac{1}{\kappa} G_{ab}^{(1)}. \]

(20.26)

According to Einstein’s equation (15.61), the sum of field and matter energy-momentum tensor must vanish, i.e., in this approximation:

\[ G_{\mu \nu}^{(1)} = \kappa m T_{\mu \nu}. \]

(20.27)

Within the present field theory in Minkowski space governed by (20.2) with the linearized field action (20.17), the same equation is found from the Euler-Lagrange equation

\[ \partial_{a} \frac{\partial L}{\partial \partial_{a} h_{bc}} = \frac{\partial L}{\partial h_{bc}}. \]

(20.28)

This reads

\[ -\partial^{c} \pi_{cab}(x) = \frac{1}{\kappa} G_{ab}(x) = T_{ab}^{m}(x) = \frac{m}{2} \int_{\tau_{a}}^{\tau_{b}} d\tau \dot{q}_{a}(\tau) \dot{q}_{b}(\tau) \delta^{(4)}(x - q(\tau)). \]

(20.29)

This may be compared with Maxwell’s equation for the electromagnetic field created by a single electron which may be written as [recall (16.20), (2.86)]

\[ -\partial_{c} e^{\pi}_{ca}(x) = \frac{1}{c} \partial_{c} F^{ca}(x) = -\frac{1}{c} j^{a}(x) = -e \int_{\tau_{a}}^{\tau_{b}} d\tau \dot{q}^{a}(\tau) \delta^{(4)}(x - q(\tau)). \]

(20.30)

where \(e^{\pi}_{ca}\) is the canonical momentum of the electromagnetic field:

\[ e^{\pi}_{ca} \equiv \frac{\partial L}{\partial \partial_{c} A_{a}} = -\frac{1}{c} F^{ca}. \]

(20.31)

There is perfect analogy between Eqs. (20.30) and (20.29).

**20.2 Gauge Invariance**

As remarked in the beginning of the last section, the similarity between the two sets of equations goes even further. The gravitational equation also share an important property of the Maxwell-Lorentz equation: the gauge invariance. Under an infinitesimal coordinate transformation

\[ x^{\mu} \rightarrow x^{\mu} - \xi^{\mu}(x), \]

(20.32)
the metric $g_{\mu\nu}(x)$ changes by $\bar{D}_\mu \xi_\nu(x) + \bar{D}_\nu \xi_\mu(x)$ [recall (11.79)]. This implies the linearized metric $g_{ab}(x) = \eta_{ab} + h_{ab}(x)$ to change by $\partial_a \xi_b(x) + \partial_b \xi_a(x)$, so that the field $h_{ab}$ undergoes a gauge transformation

$$h_{ab}(x) \rightarrow h_{ab}(x) + \partial_a \xi_b(x) + \partial_b \xi_a(x). \tag{20.33}$$

This is completely analogous to the gauge transformation

$$A_a(x) \rightarrow A_a(x) + \partial_a \Lambda(x). \tag{20.34}$$

Under these gauge transformations, the electromagnetic field tensor $F_{ab} = \partial_a A_b - \partial_b A_a$ is invariant. Under the gauge transformations (20.33), the curvature tensor and its contractions are invariant, in particular the Einstein tensor $G_{ab}$ in the field equation (20.29). This is most obvious in the double-curl representation (27.28) of the linearized Einstein tensor. The spatial components of $G^{ab}$ are a direct extension of the electromagnetic situation from the vector to the symmetric tensor potential. Whereas Maxwell’s equations (20.30) in three dimensions contain the magnetic field

$$B^i = \epsilon_{ijk} \partial_j A^k, \tag{20.35}$$

the field equation (20.29) contains $G_{ij}$ which by Eq. (27.28) is the double curl of $h_{ij}$:

$$G_{ij} = \frac{1}{2} \epsilon_{ikl} \epsilon_{jmn} \partial_k \partial_m h_{mn}. \tag{20.36}$$

In this expression, the three-dimensional gauge invariance is most transparent, and so is the invariance of the linearized field action (20.15) under the four-dimensional gauge transformations (20.33).

Let us verify explicitly that also the equation of motion (20.22) for a mass point is gauge invariant under (20.33). This is not directly obvious. We have to remember that the gauge transformation (20.33) arose by linearizing the coordinate transformation (20.32). Hence it has to be performed jointly with a change of coordinates. This gives a change in the invariant velocity of the particle

$$\dot{q}^a \rightarrow \dot{q}^a - \dot{\xi}^a = \dot{q}^a - \partial_b \xi^a \dot{q}^b, \tag{20.37}$$

and in the acceleration

$$\ddot{q}^a \rightarrow \ddot{q}^a - \partial_c \partial_b \xi^a \dot{q}^b \dot{q}^c. \tag{20.38}$$

A term $-\partial_b \xi^a \dot{q}^b$ has been dropped since it is of second-order in the small quantities $\xi$ and $h$. Inserting (20.38) into the equations of motion (20.22) and gauge-transforming also the force term via (20.33), we verify the invariance:

$$\ddot{q}^a + \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h^c_{\ cb} \right) \dot{q}^b \dot{q}^c \rightarrow$$

$$\ddot{q}^a - \partial_c \partial_b \xi^a \dot{q}^b \dot{q}^c + \left\{ \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h^c_{\ cb} \right) + \left[ \partial_c (\partial^a \xi_b + \partial_b \xi^a) - \partial^a \partial_b \xi_a \right] \right\} \dot{q}^b \dot{q}^c$$

$$= \ddot{q}^a + \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h^c_{\ cb} \right) \dot{q}^b \dot{q}^c. \tag{20.39}$$
The freedom of using different gauges can be used to simplify the expression (27.28) for the linearized Einstein tensor $G_{ab}^{(1)}$. By analogy with the Lorentz gauge $\partial_a A^a = 0$ of electrodynamics in Eq. (2.105), it is customary to choose the so-called Hilbert gauge which satisfies the condition
\[
\partial_a \left( h^{ab} - \frac{1}{2} \eta^{ab} h \right) = 0. \tag{20.40}
\]
In this gauge, one introduces the tensor field
\[
\phi^{ab} \equiv h^{ab} - \frac{1}{2} \eta^{ab} h, \tag{20.41}
\]
for which the Hilbert condition becomes simply
\[
\partial_a \phi^{ab} = 0, \tag{20.42}
\]
and the Einstein tensor (27.28) reduces to
\[
G_{ab}^{(1)} = -\frac{1}{2} \partial^2 \phi^{ab}. \tag{20.43}
\]
This brings the Einstein equation (20.27) to the simple form
\[
-\frac{1}{2} \partial^2 \phi^{ab} = \kappa \frac{m}{c} \int d\tau \dot{q}^a(\tau) \dot{q}^b(\tau). \tag{20.44}
\]
When fixing the gauge via the Hilbert condition it must be kept in mind that this does not eliminate all gauge degrees of freedom. After having enforced $\partial_a \phi^{ab} = 0$, we can still perform a further restricted set of gauge transformations
\[
\begin{align*}
  h^{ab} &\rightarrow h^{ab} + \partial^a \xi^b + \partial^b \xi^a \\
  \phi^{ab} &\rightarrow \phi^{ab} + \partial^a \xi^b + \partial^b \xi^a - \eta^{ab} \partial_c \xi^c \tag{20.45}
\end{align*}
\]
which satisfy
\[
\partial^2 \xi^a = 0. \tag{20.46}
\]
These transformations leave the field equation (20.44) and the Hilbert gauge $\partial_a \phi^{ab} = 0$ invariant. They are obviously the analogs of the gauge transformations of the second kind (2.108) in electrodynamics which leave the Maxwell equations and the Lorentz condition $\partial_a A^a = 0$ invariant.

### 20.3 Newton’s Field around a Mass Point at Rest

For a single mass point sitting at the origin, the energy-momentum tensor $\Theta_{\mu\nu}^m$ on the right-hand side of Eq. (20.44) has the simple form [recall (1.255), (1.258)]
\[
T_{\mu\nu}^m = m c \delta^{ab} \delta^{a0} \delta^{(3)}(x). \tag{20.47}
\]
The left-hand side of Eq. (20.44) becomes
\[ \frac{1}{2} \nabla^2 \phi^{ab}. \] (20.48)
Hence we find that only the field component \( \phi^{00} \) survives and is given by [recall (28.38)]
\[ \phi^{00} = -\frac{2c^2m}{4\pi r} = -\frac{4G_N m}{c^2 r}. \] (20.49)
The field \( h^{ab} = \phi^{ab} - \frac{1}{2} \eta^{ab} \phi \) has the components
\[ h^{00} = \frac{1}{2} \phi^{00}, \quad h^{ii} = \frac{1}{2} \phi^{00} \quad (i = 1, 2, 3) \] (20.50)
Hence \( h^{00} \) is related to the Newton potential \( \Phi(x) = -G_N m/r \) by
\[ h^{00}(x) = \frac{2}{c^2} \Phi(x). \] (20.51)

### 20.4 Newtonian Limit of Particle Motion

Consider the motion of a non-relativistic particle, i.e. with a velocity very small compared to the light velocity. Then
\[ |\dot{q}^i| = |v^i| \ll c, \quad d\tau \approx dt, \quad \dot{q}^0 \approx c. \] (20.52)
We shall assume particle to move in a weak gravitational field with a very slow time dependence \( \partial_i h_{ab} \approx 0 \). Then we can approximate the spatial components of the equations of motion (20.22) as
\[ \ddot{q}^i = -\frac{1}{2} c^2 \partial_i h^{00}. \] (20.53)
This is of course the geodesic equation (11.25) in the approximation
\[ \ddot{q}^i + \dot{\Gamma}^{00}_{0i} c^2 \approx 0, \] (20.54)
with the weak-field Christoffel symbol
\[ \dot{\Gamma}^{00}_{0i} = \frac{1}{2} g^{i\lambda} (2 \partial_\lambda g_{00} - \partial_\lambda g_{00}) \approx \frac{1}{2} \partial_i h^{00}. \] (20.55)
Equation (20.53) is precisely Newton’s law for the motion of a mass point in a gravitational field with the Newton potential (20.51). Thus, in the weak-field limit, the metric can be written in terms of the gravitational potential \( \Phi \) as
\[ g_{00}(x) = 1 + \frac{2}{c^2} \Phi(x), \quad g_{ii}(x) = -1 + \frac{2}{c^2} \Phi(x), \quad g_{ab} = 0 \text{ for } a \neq b. \] (20.56)
20.5 The Cosmological Term

For the purpose of studying models of the cosmos as a whole, Einstein introduced another term into the action of the gravitational field called the \textit{cosmological term}

\[ \mathcal{A} = -\frac{\Lambda}{\kappa c^2} \int d^4x \sqrt{-g}. \]  

(20.57)

The constant \( \Lambda \) has the dimension of an inverse square time. Inserting the weak-field expansion (20.1), into \( \sqrt{-g} \), we find the expansion in powers of \( h_{ab} \):

\begin{align*}
\sqrt{-g} &= \exp \left[ \frac{1}{2} \text{tr} \log (-\eta_{ab} - h_{ab}) \right] \\
&= \exp \left[ \frac{1}{2} \text{tr} \log (-\eta_{ab}) \right] \exp \left[ \frac{1}{2} h_a^a - (1/4)h_a^b h_b^a + \ldots \right] \\
&= 1 + \frac{1}{2} h_a^a - \frac{1}{4} h_a^b h_b^a + \frac{1}{8} h^2 + \ldots, \\
&= 1 + \frac{1}{2} h_a^a - \frac{1}{4} h_a^b h_b^a + \frac{1}{8} h^2 + \ldots,
\end{align*}

(20.58)

so that the cosmological action (20.57) becomes

\[ \mathcal{A} = -\frac{\Lambda}{\kappa c^2} \int d^4x \left[ 1 + \frac{1}{2} h_a^a - \frac{1}{4} h_a^b h_b^a + \frac{1}{8} h^2 + \ldots \right]. \]  

(20.59)

Cosmological data put a very stringent limit upon the size of \( \Lambda \). In fact, it appears to be almost zero. If it is nonzero, this would be the smallest non-zero physical constant known in physics. Hence we can certainly neglect all higher powers of \( h_{ab} \), keeping only the linear term. This changes the field equation (20.27) to

\[ G_{ab} - \frac{\Lambda}{c^2} \eta_{ab} = \kappa T_{ab}. \]  

(20.60)

and arrive at the linearized Einstein equation

\[ G_{ab} = \kappa T_{ab} + \Lambda \eta_{ab}. \]  

(20.61)

Inserting (20.43), this reads

\[ -\frac{1}{2} \partial^2 \phi_{ab} = \kappa T_{ab} + \Lambda \eta_{ab}, \]  

(20.62)

which becomes with (20.41):

\[ -\frac{1}{2} \partial^2 h_{ab} = \kappa \left( T_{ab} - \frac{1}{2} \eta_{ab} T \right) - \Lambda \eta_{ab}, \]  

(20.63)

For a static field, \( -\partial^2 \) reduces to \( \nabla^2 \).

In the absence of matter, the cosmological term \( \Lambda \) gives rise to a static gravitational field which satisfies the inhomogeneous Laplace equation

\[ \nabla^2 \Phi = -\Lambda. \]  

(20.64)
Comparing this with Newton’s equation
\[ \nabla^2 \Phi = 4\pi G_N \rho(x) \]  
we that the cosmological constant \( \Lambda \) has the same effect as a uniform negative mass density
\[ \rho(x) = \rho_\Lambda \equiv -\frac{\Lambda}{4\pi G_N}. \]  
If we choose \( \Phi(x) = 0 \) at the origin, a positive cosmological current makes \( \Phi(x) \) decrease radially outward like
\[ \Phi(x) = -\frac{1}{3} \Lambda r^2. \]  
This looks like an effective inverted harmonic oscillator potential superimposed upon the Newtonian force.

Note that in the presence of a cosmological term there exists no asymptotically flat space.

Experimentally, we can find bounds on the size of \( \Lambda \) by observations of the validity of Newton’s law at large distances. By looking at binary systems of galaxies (two galaxies which move in a close orbit around each other) with a size \( \approx 10^{23} \text{ cm} \) and a mass \( \approx 10^{11} M_\odot \) and finding evidence for Newton’s \( 1/r \) law, Page\(^1\) has given a first bound for \( \Lambda \):
\[ 0 < \Lambda < 10^{-33} \frac{1}{\text{sec}^2}. \]  
By studying clusters of galaxies of large dimensions, this limit was improved by three orders of magnitude since we now know more about the dynamics and mass distribution (in particular of the dark mass and black holes inside the galaxies).

The most sensitive limit comes from cosmology where \( \Lambda \) influences significantly the rate of expansion of the universe. A negative \( \Lambda \) would slow down, a positive \( \Lambda \) would accelerate the expansion with respect to the Newtonian forces. From the available data, one can estimate
\[ \Lambda \approx 10^{-35} \frac{1}{\text{sec}^2}. \]  
This corresponds to a length scale
\[ l_\Lambda = \frac{c}{\sqrt{\Lambda}} \approx 10^{28} \text{ cm} \approx 10^{10} \text{ light years} \]  
thereby reaching the order of magnitude of the size of the visible universe itself. The cosmological constant corresponds to an energy
\[ m_\Lambda c^2 \approx \sqrt{\Lambda} \hbar \approx 10^{-23} \text{ eV}. \]  

This is much smaller than the limit existing for the photon mass which was estimated via spacecraft measurement of the earth’s magnetic field to be

$$\frac{m_\gamma}{m_e} < 10^{22}. \quad (20.72)$$

In cosmological theories, the cosmological constant (20.69) is most naturally determined in terms of the dimensionless parameter

$$\Omega_{\lambda 0} = 0.68 \pm 0.10 \quad (20.73)$$
as

$$\Lambda = \Omega_{\lambda 0} \frac{3H_0^2 l_p^2}{8\pi} \approx 10^{-122} \frac{\hbar}{l_p} \quad (20.74)$$

where

$$l_p = \left( \frac{c}{G_N \hbar} \right)^{-1/2} \approx 1.615 \times 10^{-33} \text{ cm}. \quad (20.75)$$
is the Planck length formed from a combination of Newton’s gravitational constant (28.37), the light velocity $c \approx 3 \times 10^{10} \text{ cm/s}$, and Planck’s constant $\hbar \approx 1.05459 \times 10^{-27}$. It is the Compton wavelength $l_p \equiv \hbar/m_P c$ associated with the Planck mass

$$m_P = \left( \frac{c \hbar}{G_N} \right)^{1/2} \approx 2.177 \times 10^{-5} \text{ g} = 1.22 \times 10^{22} \text{ MeV}/c^2. \quad (20.76)$$

The parameter $H_0$ denotes the Hubble constant whose inverse is roughly the lifetime of the universe:

$$H_0^{-1} \approx 14 \times 10^9 \text{ years}. \quad (20.77)$$

Note that in comparison with the Planck mass (28.40), the mass $m_\Lambda$ associated with the cosmological constant in Eq. (20.71) is extremely small, the ratio being $m_\Lambda/m_P < 10^{-60}$.

## 20.6 Energy-Momentum Tensors in Linearized Gravity

Let us study the linearized gravitational field somewhat further. We may use the standard Noether rules of Section 3.6 for calculating the canonical energy-momentum tensor and apply them to the action (20.17). This gives

$$\Theta^{ab} = \frac{\partial \mathcal{L}}{\partial \ddot{h}^{cd}_{ab}} \ddot{h}^{cd} - \eta^{ab} \mathcal{L} = \pi_{bedc} \ddot{h}^{ed} - \eta^{ab} \mathcal{L}$$

$$= \frac{1}{2\kappa} (\partial_c \dot{h}_{cd} - \partial_d \dot{h}_{hc} + \eta_{bd} \dot{\partial}_c h - \eta_{bc} \partial_d h - \eta_{bd} \dot{\partial}^f h_{ec} + \eta_{bc} \partial_d \dot{h}_{ec}) \ddot{h}^{cd}$$

$$- \frac{\eta^{ab}}{8\kappa} (\partial_c h^{hc} \dot{\partial}^a h_{bc} - 2\partial_c h^{ab} \ddot{h}^{bc} + 2\partial_a h^{ab} \partial_b h - \partial_a h \dot{\partial}^a h). \quad (20.78)$$
The spin current density can be calculated as in Section 3.6, starting from the substantial derivative of the tensor field

$$\delta_s h^{ab} = \omega^{ad} h^{bd} + \omega^{bd} h^{ad}.$$  
(20.79)

Following the Noether procedure we find

$$\Sigma^{ab,c} = 2 \frac{\partial \phi}{\partial \partial_c h_{ad}} - (a \leftrightarrow b) = 2\left[\pi_{cd} h^{bd} - (a \leftrightarrow b)\right].$$  
(20.80)

Combining the two results according to Belinfante’s formula (17.148), we find the symmetric energy-momentum tensor

$$F^{ab} = \pi^{bcd} \partial_a h^{cd} - \frac{1}{2} \eta^{ab} \nabla^2,$$  
(20.81)

with the same form in terms of the fields $\phi^{ab}$:

$$F^{ab} = \frac{1}{8\kappa} \left[2 \partial_a \phi^{cd} \partial_b h_{cd} - \partial_a h \partial_b h - \eta^{ab} \left(\partial_c \phi^{de} \partial_d h_{ce} - \frac{1}{2} \partial_c h \partial^c \phi\right)\right].$$  
(20.82)

Using the field equation $\partial_a \pi^{abc} = 0$ and the gauge $\partial_a \phi^{ab} = 0$, in which $\partial_a h^{ab} = \partial^b h/2$, this takes the simple form

$$F^{ab} = \frac{1}{8\kappa} \left[2 \partial_a \phi^{cd} \partial_b \phi_{cd} - \partial_a \phi \partial_b \phi - \eta^{ab} \left(\partial_c \phi^{de} \partial_d \phi_{ce} - \frac{1}{2} \partial_c \phi \partial^c \phi\right)\right].$$  
(20.83)

### 20.7 Relation between Canonical Energy-Momentum Tensor and Einstein Tensor

The question arises as to the relation between this energy momentum tensor of the gravitational field and the Einstein tensor $G^{\mu\nu}$ which is, up to a factor $-1/\kappa$, the complete energy-momentum tensor of the gravitational field. As far as the Einstein equation is concerned,

$$G^{\mu\nu} = \kappa T^{\mu\nu}$$  
(20.84)

of this. If the field equation was expanded to all order in $k$, the Einstein equation could be rewritten in the form

$$G^{\mu\nu}_{(1)} = \kappa T^{\mu\nu} - G^{\mu\nu}_{(2)} - G^{\mu\nu}_{(3)} + \ldots,$$  
(20.85)

where $G^{\mu\nu}_{(2)}, G^{\mu\nu}_{(3)}, \ldots$ denote the terms in $G^{\mu\nu}$ of order $h^2, h^3, \ldots$.

The presence of such correction make the difference of the gauge theory of gravity with respect to the gauge theory of electromagnetism. In electromagnetism, the Maxwell equation

$$\partial_b F^{ab} = -\frac{1}{c} j^a$$  
(20.86)
is exact at the classical level. Forming the divergence on both sides gives zero for different reasons: on the left-hand side, the vanishing is trivial, due to the antisymmetry of $F^{ab}$. On the right-hand side, the vanishing divergence $\partial_a j^a = 0$ is due to the conservation of electric charges.

In the linearized theory of gravity, the analogous equation (20.27)

$$ G^{ab} = \kappa T^{m\, ab} \quad (20.87) $$

does not quite have the same consistency. At the linear level it does: using the explicit form (27.28) we can easily verify ourselves that the left-hand side has no divergence, while the right-hand side satisfies the law of energy-momentum conservation.

Consider now the second-order correction terms to the Einstein tensor:

$$ G^{ab}_{(2)} = R^{ab}_{(2)} - \frac{1}{2} R_{(2)}^{ab} - \frac{1}{2} h^{ab} R_{(1)}, \quad (20.88) $$

where

$$ R^{ab}_{(2)} = \frac{1}{2} h^{dc} [\partial^a \partial_c h_{db} - \partial_a \partial_d h^c_b + \partial_d \partial_b h^a_c - \partial^a \partial_b h_{dc}] $$

$$ - \frac{1}{4} \left[ (\partial^a h^c_b + \partial_b h^{ac} - \partial^c h^a_b)(2\partial^d h_{dc} - \partial^d h) - (\partial_d h^c_b + \partial_b h^c_d - \partial^c h_{bd})(\partial^a h^d_c + \partial^d h^a_c - \partial_a h^a_d) \right], \quad (20.89) $$

$$ R_{(2)} = \frac{1}{2} h^{dc} (\partial^2 h_{dc} + \partial_d \partial_c h - \partial^a \partial_a h_{dc} - \partial^a \partial_a h_{dc}) $$

$$ - \frac{1}{4} \left[ (2\partial_a h^{ac} - \partial^a h)(2\partial^d h^c_d - \partial^d h) - (\partial_d h^{ac} + \partial^a h^c_d - \partial^d h^a_d)(\partial_a h^d_c + \partial^d h^a_c - \partial_a h^a_d) \right]. \quad (20.90) $$

Due to its presence, the conservation law for the energy-momentum tensor of matter is only approximate:

$$ \partial_b m^{m\, ab} \approx 0. \quad (20.91) $$

The equivalence of all forms of energy forbids us to omit the energy-momentum tensor of the gravitational field itself from the conservation law. As long as we stay with the linear approximation, this omission is only a very small error. But if we move close to a large mass, the inconsistency becomes significant.

In Einstein’s theory, the lowest correction to the linearized equation amounts to adding to the energy-momentum tensor of matter on the right-hand side of (20.85), the energy-momentum tensor of the gravitational field

$$ m^{m\, ab}_{(2)} = -\frac{1}{\kappa} G^{ab}_{(2)} \quad (20.92) $$
One may conjecture that this is equivalent to adding the symmetric Belinfante
energy-momentum tensor of the gravitational field calculated in Eq. (20.82), which
is also quadratic in the fields $h_{ab}$:

$$T_{(2)}^{ab} \equiv \mathcal{T}_{(2)}^{ab} = \frac{1}{8\kappa} \left[ 2\partial^c \phi^{cd} \partial^b \phi^{cd} - \partial^a \phi \partial^b \phi - \eta^{ab} \left( \partial_c \phi^{de} \phi^c \phi_{de} - \frac{1}{2} \partial_c \phi \partial^c \phi \right) \right].$$

(20.93)

This would account for the change of the gravitational field caused by the energy-
momentum tensor of the field itself. But, using formula (20.88) we can immediately
see that $-(1/\kappa)G_{(2)}^{ab}$ is not the same tensor as $\mathcal{T}_{(2)}^{ab}$. For example, the terms $h \partial^2 h$ in $\eta^{ab} R_{(2)}$ have no counterpart in $\mathcal{T}_{(2)}^{ab}$.

However, the equivalence of the two tensor does hold, as far as the expectation
within a harmonic wave motion are concerned. Then the terms $h \partial^2 h$ in $\eta^{ab} R_{(2)}$ have no counterpart in $\mathcal{T}_{(2)}^{ab}$.

20.8 First-Order Linear Correction in Symmetric Space

The field equation (20.85)

$$G_{(1)}^{ab} \equiv \kappa (T^{ab} + \mathcal{T}_{(2)}^{ab})$$

(20.95)

with $G$ being the linearized Einstein tensor and $T^{ab}$ the quadratic Belinfante tensor
can be used to calculate the asymptotic field around a star one step beyond the
linear approximation. In the Hilbert gauge $\partial_a \phi^{ab} = 0$, the equation is

$$-\partial^2 \phi^{ab} = \kappa T^{ab} + \frac{1}{4} \left[ 2\partial^c \phi^{cd} \partial^b \phi_{cd} - \partial^a \phi \partial^b \phi - \eta^{ab} \left( \partial^d \phi^{ef} \partial_d \phi_{ef} - \frac{1}{2} \partial_d \phi \partial^d \phi \right) \right].$$

(20.96)

By taking, as the lowest approximation, the linear field (20.118), and inserting it
into the right-hand side we can find what terms in $\phi^{ab}$ are necessary to balance the
energy-momentum tensor of the field $\phi^{ab}$ itself. The result is, in the Hilbert gauge,

$$\phi^{ab} = \phi_{(1)}^{ab} + \phi_{(2)}^{ab},$$

(20.97)

where

$$\phi_{(1)}^{ab} = -4\frac{GM}{c^2 r^4} \delta^{a0} \delta^{b0},$$

$$\phi_{(2)}^{0j} = \frac{2G^2 M^2}{c^4 r^2}, \quad \phi_{(2)}^{ij} = 0, \quad \phi_{(2)}^{ij} = -\delta_{ij} \frac{4G^2 M^2}{c^4 r^2}.$$
We shall see in the next chapter that this quadratic solution contains all information about the Einstein field equation that has been subjected to an experimental test until very recently. The gravitational red shift, light deflection, and radar echo experiments which have been measured in the past are sensitive only to the leading Newtonian $1/r$-part in this field.

An important observation which gives information on the $1/r^2$ parts contained in (20.98) is the precession of the perihelion. For the planet Mercury, this has been done many years ago but due to the influence of all other planets, the extraction of the relevant data for the single planet introduces many errors. This will be discussed in Section 22.3. More recently, the precession in double pulsars has led to an extremely precise verification of Einstein’s equations in the higher-curvature regime.

In alternative proposals for the gravitational action, the post-Newtonian corrections can be parametrized with the help of three parameters $\alpha, \beta, \gamma$ as follows:

$$\phi_{(2)}^{00} = (\alpha \gamma - \alpha^2 + 2\beta) \frac{2G^2 M^2}{c^4 r^2}, \quad \phi_{(2)}^{ij} = -\frac{2G^2 M^2}{c^4 r^2} \left[3\gamma - \alpha + (\alpha - \gamma) \frac{x_i x_j}{r^2}\right].$$ (20.99)

In the famous Brans-Dicke theory [1] which contains, in addition to the metric, a scalar field whose source is the trace of the energy-momentum tensor of matter, the parameters are

$$\alpha = 1, \quad \beta = 1, \quad \gamma = \frac{1 + \omega}{2 + \omega}. \quad (20.100)$$

The calculation of post-Newtonian effects is an active field of research since theoretical results can now be compared with the abundant data from binary stars and pulsars. For references see [2, 3, 4].

### 20.9 Weak Gravitational Field of a Spinning Star.

#### Lense-Thirring Effect

For a spinning star at the origin of a Minkowski frame of reference, the energy-momentum tensor possesses not only the single element $T^{00}$ in Eq. (20.44), but also elements $T^{0i}$. Since in the rest frame the total angular momentum $J^i = (1/2)\epsilon^{ijk} J^j$ is equal to the spin $S^i$, the spin specifies the moments of the energy-momentum tensor.

$$J^i = \int d^3x [x^i T^{m\, j0}(x, t) - x^i T^{m\, i0}(x, t)] = \epsilon^{ijk} S^k. \quad (20.101)$$

If the shape of the star is invariant under rotations, then $T^{ab}(x, t)$ is independent of time and so is the gravitational field $\phi^{ab}(x, t)$. It is then easy to solve Eq. (20.44) by

$$\phi^{ab}(x) = -\frac{2\kappa}{4\pi} \int d^3x' \frac{1}{|x - x'|} T^{ab}(x'). \quad (20.102)$$
Far away from the star, we can expand, with \( r \equiv |\mathbf{x} - \mathbf{x}'| \),

\[
\frac{1}{|\mathbf{x} - \mathbf{x}'|} \approx \frac{1}{r} + \frac{x^k x'^k}{r^3} + \ldots
\]  

(20.103)

and the field is seen to have the asymptotic form

\[
\phi^{ab}(\mathbf{x}) = -\frac{2\kappa}{4\pi} \left[ \frac{1}{r} \int d^3x' T^{m00}(\mathbf{x}') + \frac{x^k}{r^3} \int d^3x' x'^k T^{mab}(\mathbf{x}') + \ldots \right].
\]  

(20.104)

Let us calculate the different components of the first term in brackets. The integral over \( T^{m00}(\mathbf{x}') \) gives the zeroth component of the total momentum and hence for a star at the origin:

\[
\int d^3x' T^{m00}(\mathbf{x}') = mc.
\]  

(20.105)

The spatial integrals over all other matrix elements \( T^{maj}(\mathbf{x}) \) vanish. For a point particle, this is obvious since \( T^{maj}(\mathbf{x}') \) is proportional to the velocities \( u^j \) which vanish in the rest frame. For a general mass distribution, we use the conservation law

\[
\partial^0 T^{maj} + \partial_j T^{m0j} = 0
\]  

(20.106)

to conclude that for a time-independent mass distribution \( \partial_j T^{m0j} = 0 \). Multiplying this by \( x^k \) and integrating over all space gives

\[
\int d^3x x^k \partial_j T^{m0j}(\mathbf{x}) = -\int d^3x T^{m0j}(\mathbf{x}) = 0.
\]  

(20.107)

Hence, the leading order in the expansion (20.104) is

\[
\phi^{ab}(\mathbf{x}) = -\frac{2\kappa}{4\pi} \frac{1}{r} Mc \delta^a_0 \delta^b_0 + \ldots
\]  

(20.108)

just as for a point particle, where

\[
T^{m00}(\mathbf{x}) = \delta^a_0 \delta^b_0 \delta^{(3)}(\mathbf{x}).
\]  

(20.109)

Consider now the second term in the expansion. Under the assumption of rotational symmetry, the integral over \( \mathbf{x}'T^{00}(\mathbf{x}') \) certainly vanishes:

\[
\int d^3x' x'^k T^{m00}(\mathbf{x}') = 0.
\]  

(20.110)

In order to calculate the \( i0 \) part of the integral

\[
\int d^3x' x'^k T^{mio}(\mathbf{x}')
\]  

(20.111)
we observe that it is antisymmetric in $k$ and $i$. This is seen by using once more the conservation law (20.106) for time independent mass distributions

$$\partial_j T^m_{ja} = 0$$  \hspace{1cm} (20.112)

and evaluating the trivial integral

$$\int d^3x' x'^k x'^i \partial_j T^m_{ja} = 0$$  \hspace{1cm} (20.113)

via integration by parts. This gives the result

$$-\int d^3x' (x'^m T^i_{ka} + x'^k T^m_{ia}) = 0,$$  \hspace{1cm} (20.114)

Thus proving the antisymmetry of (20.111) in $k$ and $i$. Hence we can write

$$\int d^3x' x'^m T^i_{m0} = \frac{1}{2} \int d^3x' (x'^k T^m_{i0} - x'^i T^m_{k0})$$

$$= \frac{1}{2} J_{ki} = \frac{1}{2} \epsilon_{kil} S_l.$$  \hspace{1cm} (20.115)

Finally, the $ij$-part of the integral $\int d^3x' x'^k T^m_{ij}(x')$ vanishes identically. This follows from the identity

$$\int d^3x' x'^k T^m_{ij} = \frac{1}{2} \int d^3x' x'^k \partial_l (x'^j T^m_{il} + x'^i T^m_{jl})$$  \hspace{1cm} (20.116)

Upon a partial integration this becomes

$$-\frac{1}{2} \int d^3x' (x'^j T^m_{ik} + x'^i T^m_{jk})$$  \hspace{1cm} (20.117)

which vanishes due to (20.114).

From (20.108) and (20.115) we find the gravitational field far away from the spinning star:

$$\phi^{00}(x) = -\frac{2\kappa Mc}{4\pi r} = -\frac{4GM}{c^2 r} + \ldots$$

$$\phi^{0i}(x) = \phi^{i0}(x) = 2\frac{G}{c^2 r^3} (x \times S)^i + \ldots$$

$$\phi^{ij}(x) = 0$$  \hspace{1cm} (20.118)

From this we obtain the deviations of the metric from the Minkowski form

$$h^{00}(x) = -2\frac{GM}{c^2 r} + \ldots,$$

$$h^{0i}(x) = h^{i0}(x) = 2\frac{G}{c^2 r^3} (x \times S)^i + \ldots,$$  \hspace{1cm} (20.119)

$$h^{ij}(x) = -\delta^{ij}\frac{2GM}{c^2 r} + \ldots,$$  \hspace{1cm} (20.120)
so that $g_{ab} \approx \eta_{ab} + h_{ab}$ becomes

$$
g_{ab} = \begin{pmatrix}
1 - 2 \frac{GM}{c^2 r} & -2 \frac{G}{c^3 r^3} x^2 S^3 & 2 \frac{G}{c^4 r^4} x^1 S^3 & 0 \\
-2 \frac{G}{c^3 r^3} x^2 S^3 & -1 - 2 \frac{2GM}{c^2 r} & 0 & 0 \\
2 \frac{G}{c^3 r^3} x^1 S^3 & 0 & -1 - 2 \frac{2GM}{c^2 r} & 0 \\
0 & 0 & 0 & -1 - 2 \frac{2GM}{c^2 r}
\end{pmatrix}
$$

(20.121)

Thus the metric of a spinning star is characterized by non-zero off-diagonal elements.

Such off-diagonal elements have a direct physical consequence. A gyroscope placed in a stationary orbit around a star processes with a rate depending on the spin of the star. For an orbit over the equator, the extra rate is $-\frac{G}{r^2} S$. The negative sign is due to the fact that the gravitational drag close to the earth is larger than away from it. For a polar orbit, the sign is positive. The derivation of this effect will be given in the next section.

## 20.10 Free Fall of a Spinning Point Particle

Consider a point particle with spin which may be realized approximately by a very small gyroscope. In Minkowski space, its equations of motion are

$$
\frac{du^a}{d\tau} = \frac{d^2 q^a}{d\tau^2} = 0, \quad \frac{dS^a}{d\tau} = 0.
$$

(20.122)

By construction, the spin vector $S^a = (1/2c)\epsilon^{abcd} J_{bc} u_d$ is orthogonal to $u^a$ [recall (1.299)], i.e.,

$$
S^a u_a = 0.
$$

(20.123)

In a space with curvature and torsion, Eq. (20.122) are simply replaced by

$$
\frac{D}{d\tau} u^\mu = 0, \quad \frac{DS^\mu}{d\tau} = 0,
$$

(20.124)

which read, more explicitly:

$$
\frac{du^\mu}{d\tau} + \Gamma^\mu_{\lambda\kappa} u^\lambda u^\kappa = 0, \quad \frac{dS^\mu}{d\tau} + \Gamma^\mu_{\lambda\kappa} u^\lambda S^\kappa = 0.
$$

(20.125)

During the fall, the length of the spin four-vector does not change since

$$
\frac{d}{d\tau} (S_a S^a) = 2 \left( \frac{D}{d\tau} S^\mu \right) S^\mu = 0
$$

(20.126)

In a comoving frame, the Christoffel symbols vanish at the position of the point particle, so that the spin does not precess.
20.11 Precession of Spinning Planet due to Tidal Forces

A planet is not a point particle and thus subjected to tidal acceleration. In a co-moving frame, the tidal acceleration \((21.42)\) between two test particles at a distance \(\delta q^i\) is

\[
\frac{d^2\delta q^i}{dt^2} = -R^i_{j00} \delta q^j c^2.
\]  

The acceleration tries to pull the second particle towards the first. Let us denote the expression on the right-hand side by \(f^i(\delta q)\). It is the force per unit mass pushing the particles apart. An extended body with a mass distribution

\[
m d^3(\delta q) n(\delta q)
\]

around the center-of-mass, which defines the comoving frame, experiences a torque per unit mass

\[
\tau^k = \int d^3(\delta q) \epsilon^{klm} \delta q^l f^m(\delta q) = -mc^2 \epsilon^{kli} R^i_{j00} \int d^3(\delta q) \delta q^l \delta q^j n(\delta q).
\]

The latter integral can be expressed in terms of the moments of inertia \(I^{ij}\) or the quadrupole moments \(Q^{ij}\) of the body, which are defined by

\[
I^{ij} = m \int d^3(\delta q) [\delta^{ij}(\delta q)^2 - (\delta q)^i (\delta q)^j] n(\delta q),
\]

\[
Q^{ij} = m \int d^3(\delta q) [3(\delta q)^i \delta q^j - \delta^{ij}(\delta q)^2] n(\delta q)
= -3I^{ij} + \delta^{ij} I^{kk}.
\]

Since \(R^i_{j00}\) is antisymmetric in \(j\) and \(i\), only the traceless part of \(I^{ij}\) contributes and we may write

\[
\tau^k = -c^2 \epsilon^{kli} R^i_{j00} \left( I^{ij} - \frac{1}{3} \delta^{ij} I^{kk} \right)
= \frac{1}{3} c^2 \epsilon^{kli} R^i_{j00} Q^{ij}.
\]

If the body has a spin \(S\) and precesses around its principal axis of inertia, the torque changes the spin vector at a rate

\[
\frac{dS^k}{dt} = \tau^k.
\]

This formula can be applied to the earth as a spinning object. Since the earth is not spherical (the equatorial diameter is 0.3% larger than the polar one), the formula leads to the equinoctial precession of the spin of the earth. It is caused by the joint solar and lunar pull acting on the equatorial bulge of the earth.

This behavior is in contrast to a completely spherical gyroscope while as we have seen in Section 11.4 remains always parallel to itself during a free fall. Thus, a set of freely falling spherical gyroscopes can be used to define a parallel vector field \(v^a\), parallel in the sense \(D_a v^a = 0\).
20.12 Precession of Gyroscope in a Satellite Orbit

In Eqs. (20.131) and (20.132) we found that in a comoving frame, the tidal forces cause a precession of a rotating body proportional to its quadrupole moment $Q^{kl}$. A rotating body without a quadrupole moment does not precess in free fall. However, if we observe such a gyroscope from another frame of reference, for example from a distance observer where the metric is asymptotically flat, the spin vector $n^\mu$ does show precession.

20.12.1 Geodetic Precession

In order to be specific, consider a gyroscope in a circular satellite orbit around the earth and let $n^i$ be the unit vector along which its spin points. Its equation of motion is most simple in a comoving frame of reference $x'\mu$ since there $\tau = t$ and $dx'\mu/d\tau = c(1, 0, 0, 0)$. In addition, the spin has only spatial components so that the convenient equation of motion $Dn^\mu/d\tau = 0$ reduces to

$$\frac{dn'^i}{d\tau} = -\Gamma^i_{0j} cn'^j.$$  

(20.133)

It is possible to use this equation and calculate from it the precession rate of $n'^i$ with respect to the distant star. For this we go into a comoving frame which maintains a fixed orientation with respect to the distant stars. This is not a falling frame, so that its Christoffel symbols do not vanish, leading to a non-zero precession rate with respect to the distant stars.

Let us focus attention upon a fixed time $t'_0$ where the orbit is parallel to the $x$-axis in coordinate frame anchored to the fixed star, and the earth lies below the gyroscope on the $z$-axis, at a distance $r$. Then there is an acceleration

$$a = \frac{GM}{r^2}$$  

(20.134)

along the $z$-axis. Let the velocity with respect to the fixed star be $v = ve_x$. It is easy to calculate, for the instant $t'_0$, the coordinate transformation to the comoving frame of reference. The motion is stopped by going to the new Lorentz frame with coordinates

$$x^0 = x'^0 + \frac{v}{c} x'^1,$$

$$x^1 = x'^1 + \frac{v}{c} x'^0,$$

$$1x^2 = x'^2,$$

$$x^3 = x'^3.$$  

(20.135)

We have assumed $v \ll c$, for simplicity. The acceleration along the negative $z$-direction is removed by an additional transformation

$$x^1 = x'^1,$$
\[ x^2 = x'^2, \]
\[ x^3 = x'^3 + \frac{a}{2c^2}(x'^0)^2. \]  
(20.136)

It is useful to define the comoving frame \( x', z' \) such that the transformed metric \( g'_{\mu\nu} \) is Minkowskian. This is achieved by the following transformation of the time:

\[ x^0 = x'^0 + \frac{v}{c}x'^1 + \frac{a}{c^2}x'^0 x'^3. \]  
(20.137)

The total transformation has the matrix form

\[
\frac{\partial x'^a}{\partial x^b} = \begin{pmatrix}
1 + ax'^3/c^2 & v/c & ax'^0/c^2 \\
v/c & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} = \alpha^a_{\phantom{a}b}.
\]  
(20.138)

When subjecting the almost Minkowskian metric \( g_{ab} = \eta_{ab} + h_{ab} \) around the earth to this transformation we obtain

\[
g'_{ab} = \frac{\partial x'^c}{\partial x^a} \frac{\partial x'^d}{\partial x^b} g_{cd} = \alpha^c_{\phantom{c}a} \alpha^d_{\phantom{d}b} h_{cd} + \alpha^c_{\phantom{c}b} \alpha^d_{\phantom{d}a} h_{cd}.
\]  
(20.139)
Neglecting quantities of the order $a^2$ and $aGM/r$, this becomes

$$g'_{ab} = \eta_{ab} + h_{ab} + \begin{pmatrix}
0 & -\frac{4vGM}{c^2r} & 0 & 0 \\
\frac{4vGM}{c^2r} & 0 & 0 & \frac{v}{c^2r}x^0 \\
0 & 0 & 0 & 0 \\
\frac{v}{c^2r}x^0 & 0 & 0 & 0
\end{pmatrix}. \quad (20.140)$$

Adding the weak gravitational field of the earth (20.50), the metric reads

$$g'_{ab} = \begin{pmatrix}
1 - \frac{2GM}{c^2r} & -\frac{4vGM}{c^2r} & 0 & 0 \\
-\frac{4vGM}{c^2r} & -1 - \frac{2GM}{c^2r} & 0 & \frac{v}{c^2r}x^0 \\
0 & 0 & -1 - \frac{2GM}{c^2r} & 0 \\
0 & \frac{v}{c^2r}x^0 & 0 & -1 - \frac{2GM}{c^2r}
\end{pmatrix}. \quad (20.141)$$

From this we calculate the Christoffel symbols

$$\bar{\Gamma}_{03}^{1'} = -\bar{\Gamma}_{01}^{3'} = -\frac{3}{2} \frac{v}{c^2r} \frac{GM}{r^2} \quad (20.142)$$

All other components vanish.

The second term is purely kinematic in origin and related to the Thomas precession. Inserting (20.134) we see that it removes $1/4$ of the first term caused by the gravitational field,

$$\bar{\Gamma}_{03}^{1'} = -\bar{\Gamma}_{01}^{3'} = -\frac{3}{2} \frac{v}{c^2r} \frac{GM}{r^2}. \quad (20.143)$$

Going back to (20.133) we find that the rate of change of the spin direction in the comoving frame is given by

$$\frac{dn'^1}{dt'} = \frac{3}{2} \frac{v}{c^2r^3} \bar{\eta}^{13},$$
$$\frac{dn'^2}{dt'} = 0,$$
$$\frac{dn'^3}{dt'} = -\frac{3}{2} \frac{v}{c^2r^3} \bar{\eta}^{1}. \quad (20.144)$$

In vector notation, this reads

$$\frac{dn'}{dt'} = 3 \frac{GM}{2c^2r^3} \times v = \Omega_g \times n'. \quad (20.145)$$

H. Kleinert, GRAVITY WITH TORSION
where the vector $\mathbf{\Omega}_g$ is the angular velocity vector of the geodetic precession. The result can also be expressed in terms of the gravitational acceleration vector as follows

$$\frac{dn'}{dt'} = \frac{3}{2c^2} (\mathbf{v} \times \dot{\mathbf{v}}) \times \mathbf{n}'.$$  \hspace{1cm} (20.146)

This result was first obtained by de Sitter in 1916.

Equation (20.146) may also be interpreted in another way. Since the spin directions can be used to define the orientation of freely falling frames of reference, Eq. (20.146) tells us how such a frame of reference is rotating with respect to the distant stars.

Just as in atomic physics, we may interpret this result as being the consequence of a gravitational spin-orbit interaction energy

$$U = \frac{2GM}{mc^2r^3} \mathbf{L} \cdot \mathbf{S} = \frac{2}{mc^2r} \frac{1}{r} \frac{\partial U}{\partial r},$$  \hspace{1cm} (20.147)

which must be corrected by the Thomas precession by replacing the number 2 by $2 - \frac{1}{2} = \frac{3}{2}$.

Note that the existence of such a spin-orbit interaction causes spinning particles to show an apparent violation of the equivalence principle: spinning electrons, protons, etc., do not move along geodesics.

For a gyroscope 500 miles above the earth in a polar orbit with a spin orthogonal to the angular momentum (i.e., in the plane of the orbit), the rate of precession is maximal and has the value

$$3 \frac{GM}{2c^2r^3} \sqrt{\frac{GM}{r}} \approx 6.9'' \text{ year}.$$  \hspace{1cm} (20.148)

If the spin is parallel to the angular momentum, there is no precession.

20.12.2 Lense-Thirring of Frame-Dragging Precession

These results are true only if we neglect the rotation of the earth around its axis. After the preparatory work in the last section it is easy to take also this effect into account. According to Eq. (20.120), the gravitational field of the rotating earth has the additional matrix elements:

$$\bar{h}_{0i} = 2 \frac{G}{r^3} (\mathbf{x} \times \mathbf{S}_e)^i = 2 \frac{G}{r^3} \epsilon^{ijk} x^j S^k_e,$$  \hspace{1cm} (20.149)

where $\mathbf{S}_e$ is the spin of the earth. It contributes to the Christoffel symbol a term

$$\Gamma_{0j}^{ij} = \frac{1}{2} (\partial_j k_{0i} - \partial_i k_{0j}) = -G \epsilon_{ijk} \epsilon_{kln} \frac{x^m}{r^3} S^{n}_{e},$$  \hspace{1cm} (20.150)

which leads to the following contribution to $dn^{ij}/dt$:

$$-\Gamma_{0j}^{ij} n^{ij} = \epsilon_{ijk} \frac{G}{r^3} \left[ S^l_{e} - \frac{3}{r^2} (\mathbf{S}_e \cdot \mathbf{x}) \right] n^{ij}.$$

$$\approx \frac{2GM}{c^2} \frac{1}{r} \frac{\partial U}{\partial r} - \frac{1}{2} \Gamma_{0j}^{ij} n^{ij}.$$

$$-\Gamma_{0j}^{ij} n^{ij} = \epsilon_{ijk} \frac{G}{r^3} \left[ S^l_{e} - \frac{3}{r^2} (\mathbf{S}_e \cdot \mathbf{x}) \right] n^{ij}.$$  \hspace{1cm} (20.151)
Writing this as $\Delta \Omega \times \mathbf{n}'$, we read off the additional precession rate

$$\Delta \Omega = -\frac{G}{r^3} \left[ S_e - \frac{3x}{r} (\mathbf{x} \cdot S_e) \right].$$  \hspace{1cm} (20.152)

This is the frame-dragging or Lense-Thirring effect.

For a gyroscope in polar orbit with the spin axis parallel to the angular momentum, this gives the only contribution to the precession rate of $0.05''$ per year. For nonpolar orbits it is negligible compared with the much larger geodetic precession.

### 20.13 Torsion in Linearized Theory

If the field action has the Einstein-Cartan form

$$\mathcal{A} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} R,$$  \hspace{1cm} (20.153)

then torsion does not enter the linearized field equation for the metric. This is seen as follows. According to Eq. (11.133), the difference between the full curvature tensor and the curvature tensor calculated from the Christoffel symbols is given, to linear order in the fields, by

$$R_{\mu\nu\lambda\kappa} - \bar{R}_{\mu\nu\lambda\kappa} \approx \partial_\mu K_{\nu\lambda\kappa} - \partial_\nu K_{\mu\lambda\kappa}. \hspace{1cm} (20.154)$$

For the Ricci tensor and the scalar curvature this implies

$$R_{\mu\kappa} - \bar{R}_{\mu\kappa} \approx \partial_\mu K^{\nu}_{\nu\kappa} - \partial_\nu K^{\mu}_{\mu\kappa},$$

$$R - \bar{R} \approx \partial_\mu K^{\nu\mu} - \partial_\nu K^{\nu\mu}. \hspace{1cm} (20.155)$$

Hence

$$G_{\mu\kappa} - \bar{G}_{\mu\kappa} \approx \partial_\mu K^{\nu}_{\nu\kappa} - \partial_\nu K^{\mu}_{\mu\kappa} - \frac{1}{2} \eta_{\mu\nu} (\partial_\mu K^{\nu}_{\nu\kappa} - \partial_\nu K^{\mu}_{\mu\kappa}). \hspace{1cm} (20.156)$$

Expressing each $K_{\mu\nu\lambda}$, except for the second, as $S_{\mu\nu\lambda} - S_{\nu\lambda\mu} + S_{\lambda\mu\nu}$, this becomes

$$G_{\mu\kappa} - \bar{G}_{\mu\kappa} = 2\partial_\mu S_{\kappa} - \partial_\nu K^{\nu}_{\kappa\mu} - 2\eta_{\mu\nu} \partial_\mu S^{\nu}_{\kappa} \hspace{1cm} (20.157)$$

This difference is precisely equal to the Belinfante correction required to make $G_{\mu\kappa}$ symmetric. Indeed, if we form the combination (18.45)

$$G_{\mu\kappa} - \frac{1}{2} \partial^\nu (S_{\mu\kappa,\nu} - S_{\kappa\nu,\mu} + S_{\nu\kappa,\mu}), \hspace{1cm} (20.158)$$

and insert the Palatini tensor from (18.40), we find

$$G_{\mu\kappa} - \partial^\nu (K_{\mu\kappa\nu} + 2\eta_{\mu\nu} S_{\kappa} - 2\eta_{\mu\kappa} S_{\nu}) = \bar{G}_{\mu\kappa}. \hspace{1cm} (20.159)$$

Thus, also in the presence of torsion, the linearized Einstein equation for $h_{ab}$ reads

$$\bar{G}_{ab} = \kappa \bar{T}^{m}_{ab}. \hspace{1cm} (20.160)$$
The equation for the torsion

$$S_{ab,c} = \kappa \Sigma_{ab,c}$$  \hspace{1cm} (20.161)

is completely independent of this and does not influence the metric. This will happen only to next order in $h_{\mu \nu}$. The only physical consequence which is changed by the presence of torsion at this linear level is the motion of spinning matter through spinning matter, as we shall see later. Such consequences will, however, remain undetectable even in the distant future.

**Notes and References**

[1] For the scalar tensor theory see


[3] W. Goldberger and I. Rothstein, hep-th/0409156. For a recent review see

21

Experimental Tests in Weak Field

Although this booklet is not meant to replace standard textbooks on Einstein gravity, let us derive, for completeness, some immediate consequence of the linearized Einstein theory. The possible consequences of torsion will be discussed Chapter 25.

21.1 Gravitational Red Shift

Given the weak-field limit (20.56) of \( g_{\mu\nu} \), it is easy to calculate the gravitational red shift of spectral lines emitted from the atmosphere of a star. The light can be thought of coming from an oscillator of frequency \( \omega \) which, in the absence of a gravitational field, would have a period oscillation \( \Delta \tau = 1/\omega_0 \). In the presence of the gravitational field, this period is maintained only in a freely falling frame, of reference. When considering the atom at rest, this period is to be identified with a proper time interval \( \Delta \tau \). Since for the action at rest \( dq^i = 0 \), this proper time interval transforms into the local coordinate time interval the surface of the star

\[
\Delta t = \frac{1}{\sqrt{g_{00}}} \Delta \tau.
\]  

(21.1)

Hence the proper frequency \( \omega_0 \) changes to the local frequency

\[
\omega = \sqrt{g_{00}} \omega_0.
\]  

(21.2)

Inserting the above formula (20.56), this implies a frequency reduction

\[
\frac{\Delta \omega}{\omega} = \frac{\Phi}{c^2} = -\frac{G M}{c^2 r}.
\]  

(21.3)

Let us estimate the size of this effect following Ref. [1]. The mass and radius of the sun are

\[
M_\odot = 1.97 \times 10^{33} \text{g},
\]

\[
R_\odot = 0.695 \times 10^6 \text{km}.
\]  

(21.4)

With the gravitational constant being

\[
G = 6.672 \times 10^{-8} \text{cm}^3\text{g}^{-1}\text{sec}^{-2},
\]  

(21.5)
we find on the surface of the sun a gravitational field

\[ \Phi = -\frac{GM_\odot}{c^2 R_\odot} \approx -2.12 \times 10^{-6}. \]  

(21.6)

Such a small red shift of light emitted from the sun’s atmosphere is very hard to detect. The reason is that a relative velocity of only 600 m/sec will produce the same frequency shift via the Doppler effect. Since the sun’s atmosphere has a temperature of about 3000°K, the thermal velocity of light element such as, C, N, O is about three times as large as that. The red shift is therefore hidden well within the line width due to thermal Doppler broadening. In addition, there are unknown Doppler shifts due to the violent convection in the sun’s atmosphere which vary from place to place. These shifts can easily move a spectral line red shifted via (21.6) completely into the blue. Only in recent years have the experimentalists claimed to have sufficient control over this problem such as to verify the theoretical number up to 5% accuracy.

Heavy small stars, called white dwarfs, are in principle more suitable for a test since they have about 1/2 the sun’s mass but are smaller by a factor 10 to 100. They pose, however, another problem, namely that mass and radius are not directly measurable. The mass is accessible only in double stars like 40 Eridani A and B. They are so far apart that one can localize the center of gravity and find, from the period of revolution, both masses. The radius can only be inferred theoretically, from astrophysical calculations relating mass and radius. In this way one finds on the surface of 40 Eridani B

\[ \Phi \sim -(5.7 \pm 1) \times 10^{-5}, \]  

(21.7)

while the red shift is

\[ \frac{\Delta \omega}{\omega} = -(7 \pm 1) \times 10^{-5}. \]  

(21.8)

The most accurate determination has been made in a laboratory experiment on the earth by Pound and Reblea. They looked at a 14.41 keV X-ray emitted by Fe^{57} and reabsorbed it 22.6 m lower in another sample of Fe^{57}. By using crystalline Fe^{57}, they could take advantage of the Mößbauer effect, according to which there is a certain probability that the X-ray is emitted without recoil guaranteeing that the spectral line has not Doppler red shift.

Across the 22.6 m, the earth’s gravitational potential changes by

\[ \frac{\Delta \Phi}{c^2} = -980 \frac{\text{cm}}{\text{sec}^2} \cdot 2260 \frac{\text{cm}}{c^2} = -2.46 \times 10^{-15}. \]  

(21.9)

At first sight it seems impossible to measure such a small red shift since the lifetime \( \Gamma^{-1} \) is only \( \approx 1 \mu \text{sec} \) such that the relative line width is \( \Gamma/\nu \sim 1.3 \times 10^{-12} \) is 1000 times larger than the red shift. In order to increase the sensitivity Pound and Reblea
mounted the source on a fast rotating wheel with peripheral velocity $v_0$ such that $\Delta \nu/\nu$ receives an additional Doppler shift and becomes

$$\frac{\Delta \nu'}{\nu} = \frac{\Delta \nu}{\nu} - \frac{v_0}{c} \cos \omega t.$$  (21.10)

Then the absorption cross section has the temporal behavior

$$\frac{\Gamma^2}{(\Delta \nu')^2 + \Gamma^2} = \frac{(\Gamma/\nu)^2}{\left(\frac{\Delta \nu}{\nu} - \frac{v_0}{c} \cos \omega t\right)^2 + (\Gamma/\nu)^2} \approx \frac{(\Gamma/\nu)^2}{\frac{v_0^2}{c^2} \cos^2 \omega t + (\Gamma/\nu)^2} \left\{ 1 + \frac{2 \Delta \nu v_0}{\nu c \cos \omega t} \right\}.$$

(21.11)

The gravitational red shift $\Delta \nu/\nu$ was extracted from the deviation of the observed behavior from the ideal behavior in the prefactor of (21.11). This gave

$$\frac{\Delta \nu}{\nu} \sim (2.575 \pm .26) \times 10^{-15}.$$  (21.12)

### 21.2 Deflection of Light Grazing the Sun

Let us now turn to the gravitational effect which historically represented the first spectacular success of Einstein theory. For this we consider the equation of motion (20.22), and multiply it by the particel mass $m$ to write it as

$$\dot{p}^a + \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h_{cb} \right) p^b \dot{q}^c \equiv 0.$$  (21.13)

Due to the de Broglie relation $p^a = \hbar k^a$, the same equation holds for the wave vector of the particles:

$$\dot{k}^a + \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h_{cb} \right) k^b \dot{q}^c \equiv 0.$$  (21.14)

In a weak gravitational field, this yields

$$\dot{k}^a = - \left( \partial_c h^a_{\ b} - \frac{1}{2} \partial^a h_{bc} \right) k^b \dot{x}^c = -\dot{h}^a_{\ b} k^b + \frac{1}{2} \partial^a h_{bc} k^b \dot{x}^c$$

(21.15)

where we have omitted terms of order $h_{ab}^2$, and $q^c$ have been approximated by Minkowski coordinates $x^c$. For a particle which starts at infinity with momentum $k^a$ and ends up at infinity with a small deviation of the momentum, the term $\int d\tau \dot{h}^a_{\ b} k^b$ gives only a second-order contribution to the deviation and can be neglected.

Equation (21.15) contains no mass term and must therefore remain valid in the limit $m \to 0$, where it described photons.
21.2 Deflection of Light Grazing the Sun

Consider now a light ray passing the sun with an impact parameter $b$ as in Fig. 24.1. Its trajectory before it gets close to the sun is

$$dx^a = dz(1, 0, 0, 1)$$

and its initial momentum vector is

$$k^a = |k|(1, 0, 0, 1).$$

When arriving close to the sun, the momentum starts deviating by

$$dk^a = \frac{1}{2} \partial^a (h_{00} + h_{30} + h_{03} + h_{33}) |k| dz. \qquad (21.18)$$

Assuming only a very small deflection angle, the right-hand side remains correct even after the light has passed the sun. Inserting the Newtonian limit (20.50) for $h_{ab}$ which has $h_{00} = h_{33} = -2GM_\odot/c^2r$, this becomes

$$dk^0 = |k| \partial^0 h_{00} dz = 0,$$

$$dk^i = |k| \partial^i h_{00} dz = 2|k| \partial_i \left( \frac{GM_\odot}{c^2 r} \right) dz. \qquad (21.20)$$

The deflection points in the 1-direction, so that the only relevant component is

$$dk^1 = 2|k| \frac{GM_\odot}{c^2} \partial_1 \frac{1}{r}. \qquad (21.21)$$
This has to be integrated along the entire z-axis with the result
\[ \int dz \partial_1 = \left. \frac{1}{r} \right|_{y,z} = - \int_{-\infty}^{\infty} dz \frac{b}{\sqrt{z^2 + b^2}} = - \frac{2}{b}. \] (21.22)

Hence the total relative change of \( k_1 \) is
\[ \frac{\Delta k_1}{|k|} = -4 \frac{GM_{\odot}}{c^2 b}. \] (21.23)

The remaining components of \( k_2, k_3 \) do not change to this order in \( G \). Since we use the linearized theory, the quantity (21.23) is must be small, in which case it gives directly the angle of deflection \( \theta \). Inserting the solar parameters
\[ R_{\odot} \approx 6.96 \times 10^{10}, \quad M_{\odot} \approx 1.99 \times 10^{33} \text{g} \] (21.24)

one obtains, for light grazing the surface of the sun, the small angle
\[ \theta \approx -1.75^\prime. \] (21.25)

The first measurement of the light deflection was performed in May 1919 during a total eclipse of the sun on the island of Sobral (off Brazil) and in Principe, in the Gulf of Ginea. Due to the sun’s corona, the closest distance a star can be followed was really \( b \approx 2R_{\odot} \), so the angle takes only half the value (21.25). The measured values, extrapolated to \( R_{\odot} \), were \( \theta = 1.98 \pm .16 \) (Sobral) and 1.61\( \pm \).40 (Principe). A more precise measurement was performed in 1973 by the University of Texas based on observations made in Mauritania, with the result \( \theta \approx 1.58 \pm .16 \) (see Table 21.1).

More precise tests of Eq. (21.23) can be obtained from observations of the deflection of radio waves in the cm-range coming from quasistellar sources 3C279. By radio interferometry between two parabolic antennas separated by a long baseline [3] one was able to measure a value of \( \theta \) between
\[ \theta \sim 1.64 \pm 1 \quad \text{and} \quad 1.82 \pm .06. \] (21.26)

The accuracy of radio wave interferometry depends on the ratio of wave length \( \lambda \) to length \( d \) of the baseline, \( \Delta \theta \approx \lambda/d \). With \( \lambda \approx 3 \text{km} \) and \( d \equiv 3000 \text{km} \) one can obtain, in principle, an accuracy of \( .002^\prime \).

A recent experiment with a baseline 10,000 km has given a convincing verification of the theory (see Table 21.2)
\[ \theta_{\exp} / \theta_{\exp} = 1.0001 \pm 0.0001 \] (21.27)

The outcome of the lost experiment is used as an evidence against the alternative theory by Brans and Dicke which was proposed in 1961 and involves an additional scalar field which has the effect of giving the gravitational constant a space dependence.
### Table 21.1 Experimental Results on the Deflection of Light [2].

<table>
<thead>
<tr>
<th>Observatory</th>
<th>Eclipse</th>
<th>Site</th>
<th>$\theta$ (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greenwich</td>
<td>May 29, 1919</td>
<td>Sobral</td>
<td>1.98 ± 0.16</td>
</tr>
<tr>
<td>Greenwich</td>
<td>&quot;</td>
<td>Principle</td>
<td>1.16 ± 0.40</td>
</tr>
<tr>
<td>Adelaide-Greenwich</td>
<td>Sept. 21, 1922</td>
<td>Australia</td>
<td>1.16 ± 0.40</td>
</tr>
<tr>
<td>Victoria</td>
<td>&quot;</td>
<td>Australia</td>
<td>1.42 to 2.16</td>
</tr>
<tr>
<td>Lick</td>
<td></td>
<td>Australia</td>
<td>1.72 ± 0.15</td>
</tr>
<tr>
<td>Lick</td>
<td></td>
<td>Australia</td>
<td>1.82 ± 0.20</td>
</tr>
<tr>
<td>Potsdam</td>
<td>May 9, 1929</td>
<td>Sumatra</td>
<td>2.24 ± 0.10</td>
</tr>
<tr>
<td>Sternberg</td>
<td>June 19, 1936</td>
<td>U.S.S.R</td>
<td>2.73 ± 0.31</td>
</tr>
<tr>
<td>Sendai</td>
<td></td>
<td>Japan</td>
<td>1.28 to 2.13</td>
</tr>
<tr>
<td>Yerkes</td>
<td>May 20, 1947</td>
<td>Brazil</td>
<td>2.01 ± 0.27</td>
</tr>
<tr>
<td>Yerkes</td>
<td>Feb. 25, 1952</td>
<td>Sudan</td>
<td>1.70 ± 0.10</td>
</tr>
<tr>
<td>U. of Texas</td>
<td>June 30, 1973</td>
<td>Mauritania</td>
<td>1.66 ± 0.19</td>
</tr>
</tbody>
</table>

### Table 21.2 Experimental Results on the deflection of radio waves [2].

<table>
<thead>
<tr>
<th>Radio telescope</th>
<th>Observer(s)</th>
<th>Baseline (km)</th>
<th>$\theta_{\text{exp}}/\theta_{\text{theor}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owens Valley</td>
<td>Seielstadt et al. (1970)</td>
<td>1.07</td>
<td>1.01 ± 0.11</td>
</tr>
<tr>
<td>Goldstone</td>
<td>Muhleman et al. (1970)</td>
<td>21.56</td>
<td>1.04 ± 0.15</td>
</tr>
<tr>
<td>National RAO</td>
<td>Sramek (1971)</td>
<td>2.7</td>
<td>0.9 ± 0.05</td>
</tr>
<tr>
<td>Mullard RAO</td>
<td>Hill (1971)</td>
<td>≈ 1</td>
<td>1.07 ± 0.05</td>
</tr>
<tr>
<td>Cambridge</td>
<td>Riley (1973)</td>
<td>4.57</td>
<td>1.04 ± 0.08</td>
</tr>
<tr>
<td>Westerbork</td>
<td>Weiler et al. (1974)</td>
<td>1.44</td>
<td>0.96 ± 0.05</td>
</tr>
<tr>
<td>Haystack and National RAO</td>
<td>Counselman et al. (1974)</td>
<td>845</td>
<td>0.99 ± 0.03</td>
</tr>
<tr>
<td>National RAO</td>
<td>Fomalont and Sramek (1975)</td>
<td>35</td>
<td>1.015 ± 0.011</td>
</tr>
<tr>
<td>National RAO</td>
<td>Sramek (1974)</td>
<td>35</td>
<td>0.097 ± 0.08</td>
</tr>
<tr>
<td>Westerbork</td>
<td>Weiler et al. (1975)</td>
<td>≈ 1</td>
<td>1.04 ± 0.03</td>
</tr>
<tr>
<td>National RAO</td>
<td>Fomalont and Sramek (1976)</td>
<td>35</td>
<td>1.007 ± 0.009</td>
</tr>
<tr>
<td>VLBI</td>
<td>Roberston and Carter (1984)</td>
<td>≈ 10,000</td>
<td>1.004 ± 0.002</td>
</tr>
<tr>
<td>VLBI</td>
<td>Robertson et al. (1991)</td>
<td>≈ 10,000</td>
<td>1.0001 ± 0.0001</td>
</tr>
</tbody>
</table>
21.3 Time Delay of Light Grazing the Sun

Using the formula (21.20) we see that the $z$-component of the light wave vector receives a shift by

$$\Delta k^3 = 2|k| \frac{G}{c^2} \int_{-\infty}^{\infty} dz \partial_3 \frac{1}{r} = 2|k| \frac{G}{c^2} M_{\odot} \frac{1}{r}. \quad (21.28)$$

while the zeroth component of $k^0$ is constant. Thus, we find that the phase velocity of the light ray in the Minkowski background metric is equal to

$$v = \frac{c|k|}{k^3} = \frac{c}{1 + 2GM_{\odot}/c^2 r}. \quad (21.29)$$

Since this is independent of $|k|$, there is no dispersion and $r$ is also the physically relevant group velocity. Thus, light signals close to the sun travel slower than the speed of light $c$. The gravitational field in the neighborhood of the sun acts like a refractive medium. In fact the light deflection calculated in the previous section may also be viewed as a consequence of this refraction.

Notice that this state of affairs is not in contradiction with the basic principle of relativity that light has always the same velocity $c$. This principle is certainly unaffected since it refers to a freely falling observer. Even close to the sun he sees the light pass by with velocity $c$. A static observer far away from the sun registers an effective decrease in velocity, due to a change of the way time passes for him and a change in the spatial metric.

In order to see this more explicitly, we may use the property that a light particle follows a path of zero proper time

$$d\tau^2 = \frac{1}{c^2} g_{\mu
u} dx^\mu dx^\nu = 0 \quad (21.30)$$

and calculate the speed of the particle in the background Minkowski frame. Inserting $g_{ab} = \eta_{ab} + h_{ab}$ with $h_{ab}$ from (20.50), this gives

$$\left(1 - 2\frac{GM_{\odot}}{c^2 r}\right) c^2 dt^2 - \left(1 + 2\frac{GM_{\odot}}{c^2 r}\right) dx^2 = 0 \quad (21.31)$$

such that the light particle moves with the speed

$$v = \sqrt{\left(\frac{dx}{d\tau}\right)^2} = c \sqrt{\frac{1 - 2GM_{\odot}/c^2 r}{1 + 2GM_{\odot}/c^2 r}} \sim \left(1 - 2GM_{\odot}/c^2 r\right) \quad (21.32)$$

which is the same as (21.29).

The decrease in the effective velocity leads to a time delay with respect to a signal through an empty space [4]

$$\Delta t = \int_{z_1}^{z_2} \frac{dz}{v(z)} - \int_{z_1}^{z_2} \frac{dz}{c} = \int_{z_1}^{z_2} \frac{dz}{c} \frac{2GM_{\odot}}{c^2 r}. \quad (21.33)$$

H. Kleinert, GRAVITY WITH TORSION
21.4 Neighbouring Trajectories and Tidal Forces

\[ \frac{2GM}{c^3} \int_{z_1}^{z_2} \frac{dz}{\sqrt{z^2 + b^2}} = \frac{2GM}{c^3} \log \left[ \frac{z_2 + \sqrt{z_2^2 + b^2}}{z_1 + \sqrt{z_1^2 + b^2}} \right] \]

\[ = \frac{2GM}{c^3} \log \left[ \frac{(z_2 + \sqrt{z_2^2 + b^2})(z_1 + \sqrt{z_1^2 + b^2})}{b^2} \right]. \]

For \( |z_{12}| \gg b \) this gives

\[ \Delta t \sim \frac{2GM}{c^3} \log \left( \frac{4|z_1|z_2}{b^2} \right). \] (21.34)

For mercury we use \( z_1 = \text{distance earth-sun} \approx 14.9 \times 10^{12} \text{cm} \), \( z_2 = \text{distance mercury-sun} \approx 5.8 \times 10^{12} \text{cm} \) and calculate for \( b = \text{radius of the sun} \): \( \Delta t(= -110 \times 10^{-6} \text{sec (out of } \approx 100 \text{ sec. total time))}. \) For an echo, the delay time is twice as large.

When measuring the time on earth one must observe that formula (20.102) gives the time delay in the frame in which the sun is at rest while the time on earth is the proper time \( \Delta t \) (of the freely orbiting earth). The relation is

\[ \Delta \tau = \Delta t \sqrt{g_{00}} = \Delta t \sqrt{1 + h_{00}} \approx \Delta t \left( 1 - \frac{GM}{c^2 r} \right) \] (21.35)

(The effect due to the earth’s gravitational field can be neglected) where \( r \) is the distance earth-sun. This gives a 10% correction to the delay time of Mercury.

The experiments performed by Shapiro for Mercury, Venus, and Mars show at most 2% deviations from the theory. Recently, radio signals to Mariner 6 and 7 gave deviations of maximally 3%.

### 21.4 Neighbouring Trajectories and Tidal Forces

Consider a swarm of free particles in a space without torsion in which particles move along geodesics. Let us focus attention upon two neighboring material points moving along the trajectories \( x^\mu(s) \) and \( x^\mu(s) + \delta x^\mu(s) \). Inserting this into Eq. (14.7), we find that the difference vector \( \delta q^\mu(s) \) satisfies the equation of motion

\[ \frac{d^2 \delta q^\mu}{d\tau^2} + \Gamma^\mu_{\lambda\kappa}(q + \delta q) \frac{d(q^\lambda + \delta q^\lambda)}{d\tau} \frac{d(q^\kappa + \delta q^\kappa)}{d\tau} = 0, \] (21.36)

which amounts to

\[ \frac{d^2 \delta q^\mu}{d\tau^2} + \partial^\rho \Gamma^\mu_{\lambda\kappa} \delta q^\rho \frac{dq^\lambda}{d\tau} \frac{dq^\kappa}{d\tau} + 2\Gamma^\mu_{\lambda\kappa} \delta q^\lambda \frac{dq^\kappa}{d\tau} = 0. \] (21.37)

Remembering that \( S^\mu_{\nu\lambda} = 0 \) we do not have to write bars on top of \( D \) and \( \Gamma^\mu_{\nu\lambda} \).
21.4.1 Simplification

This result (21.37) is physically not very transparent since the components $\delta q^\mu$ refer to different local coordinates before and after the displacement by $d\tau$. In order to find the true $\tau$-dependence it is better to calculate the covariant acceleration

$$\frac{D^2 \delta q^\mu}{d\tau^2} = \frac{D}{d\tau} \frac{d}{d\tau} \delta q^\mu.$$  \hspace{1cm} (21.38)

This requires a little more work which goes as follows:

$$\frac{D^2 \delta q^\mu}{d\tau^2} = \frac{D}{d\tau} \left( \frac{d \delta q^\mu}{d\tau} + \Gamma_\lambda^\mu_{\rho} \frac{dq^\lambda}{d\tau} \delta q^\rho \right)$$  \hspace{1cm} (21.39)

$$= \frac{d^2 \delta q^\mu}{d\tau^2} + \Gamma_\lambda^\mu_{\kappa} \frac{dq^\lambda}{d\tau} \delta q^\kappa + \frac{d}{d\tau} \left( \Gamma_\lambda^\mu_{\rho} \frac{dq^\lambda}{d\tau} \delta q^\rho \right) + \Gamma_\sigma^\mu_{\kappa} \frac{dq^\sigma}{d\tau} \Gamma_\lambda^\mu_{\rho} \frac{dq^\lambda}{d\tau} \delta q^\rho.$$  \hspace{1cm} (21.40)

Replacing the first term on the right-hand side using Eq. (21.37), and writing out the third term explicitly, gives

$$\frac{D^2 \delta q^\mu}{d\tau^2} = -\partial_\rho \Gamma_\lambda^\mu_{\kappa} \delta q^\rho \frac{dq^\lambda}{d\tau} \frac{dq^\kappa}{d\tau} - 2 \Gamma_\lambda^\mu_{\rho} \frac{d\delta q^\lambda}{d\tau} \frac{dq^\kappa}{d\tau} + \Gamma_\lambda^\mu_{\kappa} \frac{dq^\lambda}{d\tau} \frac{d\delta q^\kappa}{d\tau} + \frac{d}{d\tau} \left( \Gamma_\lambda^\mu_{\rho} \frac{dq^\lambda}{d\tau} \delta q^\rho \right) + \Gamma_\sigma^\mu_{\kappa} \Gamma_\lambda^\mu_{\rho} \frac{dq^\sigma}{d\tau} \frac{dq^\lambda}{d\tau} \delta q^\rho.$$  \hspace{1cm} (21.41)

The second, third, and sixth terms cancel each other. Inserting in the fifth term the equation of motion (14.7), and remembering the definition (11.128) of the curvature tensor, we obtain

$$\frac{D^2 \delta q^\mu}{d\tau^2} = -R_\rho^\lambda_{\kappa} \delta q^\rho \frac{dq^\lambda}{d\tau} \frac{dq^\kappa}{d\tau}.$$  \hspace{1cm} (21.41)

This shows that in a curved space the distance between neighboring particles experiences an acceleration proportional to the local curvature. The right-hand side of (21.41) is called the tidal force since it causes the tidal motion of the oceans on the earth. In a satellite around the earth’s orbit, only the center of gravity is free of gravitational forces. At that point, centrifugal and gravitational forces are exactly balanced. This is no longer true in the neighborhood of this point. Due to the tidal forces, a set of liquid droplets spilled in a satellite moves apart winding eventually up on the walls.

21.4.2 Weak-Field Limit

Going to a frame of reference which falls freely with the particle at $q(\tau)$, the proper time coincides with the true time. Moreover, in such a comoving frame, the affine
connection vanishes such that the first covariant derivative is equal to the simple time derivative. The second covariant derivative is equal to the second time derivative if we assume the relative distance vector $\delta x^i$ to have no initial speed, $d\delta q^i/dt = 0$. Thus, looking at two mass points at equal times we find from Eq. (21.41) the relative acceleration

$$\frac{d^2 \delta q^i}{dt^2} = - R_{j00}^i \delta x^j c^2. \quad (21.42)$$

In the weak-field limit, we insert the expression (27.25) for the curvature tensor in linear approximation,

$$R_{j00i} \approx \frac{1}{2} (\partial_j \partial_0 \delta h_{0i} - \partial_i \partial_j \delta h_{00} - \partial_0 \partial_i \delta h_{0j} + \partial_0 \partial_j \delta h_{i0}). \quad (21.43)$$

If the field has the Newtonian form derived in Section 20.4, this becomes simply

$$R_{j00i} \approx - \frac{1}{c^2} \partial_i \partial_j \Phi. \quad (21.44)$$

Note that the tidal force in empty space is always divergenceless. This follows directly from Eq. (21.44), according to which

$$\partial_i R_{j00i} = - \frac{1}{c^2} \partial_i \nabla^2 \Phi = 0 \quad (21.45)$$

Equation (21.42) shows that the measurement of the tidal forces determines the components $R_{j00i}$ of the curvature tensor. Since $R_{j00i}$ is antisymmetric in $j$ and $i$, these are six independent components of $\bar{R}_{\mu\nu\lambda\kappa}$. As shown in Section 12.7, the total number of independent components of $\bar{R}_{\mu\nu\lambda\kappa}$ is 20. How can we determine the remaining components?

They are found from measurements of the tidal forces in different freely falling coordinate frames. These frames have all the same acceleration but different velocities. Thus, they differ only by a local Lorentz transformation. It can be verified that by measuring $R_{j00i}$ in all such frames we can determine all components

$$R_{j00i} = \Lambda_j^\mu \Lambda_0^\nu \Lambda_0^\lambda \Lambda_i^\kappa R_{\mu\nu\lambda\kappa}. \quad (21.46)$$

Notes and References

[1] See more details in


The gravitational red shift of the sun has been claimed to be $1.5 \pm 0.05$ the theoretical value by J. Brault, Bull. Am. Phys. Soc. 8, 28 (1963)


[2] Table taken from

[3] For the high-accuracy very long baseline interferometry experiment see

[4] The time delay measurements of radar echoes are reported in
The mariner results are given in
Simple Exact Solutions of Einstein Equations

The simplest solutions of Einstein’s equations which describe the gravitational field around a celestial object are static and have spherical symmetry. In order to describe them, let first us prepare a convenient coordinate system.

22.1 Static Spherically Symmetric Coordinates

We ignore a possible time dependence of the metric $g_{\mu\nu}(x)$ and assume it to have the general rotationally invariant form

$$
g_{\mu\nu} = \begin{pmatrix} F(r) & -E(r)x^i \\ -E(r)x^i & -C(r)\delta^{ij} - D(r)x^ix^j \end{pmatrix}, \tag{22.1}$$

where $r = \sqrt{x^2}$ is the spatial radius. The invariant distance is therefore

$$
ds^2 = F(r)c^2 dt^2 - 2E(r)c x^i dt dx^i - C(r)(dx^i)^2 - D(r)(x^i dx^i)^2. \tag{22.2}$$

The spatial vectors $x^i$ can be expressed in terms of the usual spherical angles $r$, $\theta$, and $\phi$, leading to

$$
ds^2 = F(r)c^2 dt^2 - 2E(r)rc dt dr - D(r)r^2 dr^2 - C(r)[dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2]. \tag{22.3}$$

It is useful to simplify this expression by removing the off-diagonal contribution proportional to $dt dr$. This is possible with the help of Einstein transformation of the time

$$
dt \rightarrow dt - \frac{r E(r)}{c F(r)} dr. \tag{22.4}$$

This leaves us with the invariant length

$$
ds^2 = F(r)c^2 dt^2 - G(r'')(dr'')^2 - C(r'')[dr'']^2 + r''^2(d\theta)^2 + r''^2 \sin^2 \theta (d\phi)^2 \tag{22.5}$$
where

\[ G(r) \equiv r^2 \left[ D(r) + \frac{E^2(r)}{F(r)} \right] \quad (22.6) \]

It is further useful to rescaling the radius \( r \) to, say, \( r' \) in such a way that the last two terms in (22.5) becomes \( r'^2 (d\theta^2 + \sin^2 \theta d\phi^2) \) with no extra prefactor. This is possible by setting

\[ r'^2 \equiv r^2 C(r). \quad (22.7) \]

Then \( ds^2 \) is simplified further to what is called the standard form of the invariant length square

\[ ds^2 = B(r')c^2 (dt)^2 - A(r')(dr')^2 - r'^2 [(d\theta)^2 + \sin^2 \theta (d\phi)^2]. \quad (22.8) \]

It is specified by only two \( r \)-dependent functions:

\[ B(r') \equiv F(r), \quad A(r') \equiv \left[ 1 + \frac{G(r)}{C(r)} \right] \left[ 1 + \frac{r C'(r)}{2 C(r)} \right]^{-2}. \quad (22.9) \]

We could also have gone to the spatially isotropic form

\[ ds^2 = H(r'')c^2 (dt)^2 - J(r'')[(dr'')^2 + r''^2 (d\theta)^2 + r''^2 \sin^2 \theta (d\phi)^2], \quad (22.10) \]

with the help of the transformation

\[ H(r'') = F(r), \quad J(r'') = \frac{C(r) r'^2}{r''^2}, \quad dr'' = \frac{dr}{r} \left[ 1 + \frac{G(r)}{C(r)} \right]^{1/2}, \quad (22.11) \]

The coefficients of the standard form (22.8) are related to \( H(r'') \) and \( J(r'') \) by

\[ B(r') = H(r''), \quad A(r') = \left[ 1 + \frac{r'' J'(r'')}{2 J(r'')} \right]^{-2}, \quad r'^2 = r''^2 J(r''), \quad \frac{dr'}{dr''} = \frac{J(r'')}{A(r')} . \quad (22.12) \]

In the sequel, we shall mostly use the standard form (22.8). When doing so, we shall omit the primes. If the coordinates \( dx'^\mu = (dx^0, dx^1, dx^2, dx^3) \) denote the differentials of time and the spherical coordinates, i.e., \( dx'^\mu = (dt, dr, d\theta, d\phi) \), the standard metric has the diagonal matrix form

\[ g_{\mu\nu} = \begin{pmatrix}
B(r)c^2 & 0 & 0 & 0 \\
0 & -A(r) & 0 & 0 \\
0 & 0 & -r^2 & 0 \\
0 & 0 & 0 & -r^2 \sin^2 \theta
\end{pmatrix}, \quad (22.13) \]

which has the simple determinant

\[ g = -r^4 A(r) B(r)c^2 \sin^2 \theta. \quad (22.14) \]

H. Kleinert, GRAVITY WITH TORSION
The inverse is simply

\[ g^{\mu\nu} = \begin{pmatrix} B^{-1}(r)c^{-2} & 0 & 0 & 0 \\ 0 & -A^{-1}(r) & 0 & 0 \\ 0 & 0 & -r^{-2} & 0 \\ 0 & 0 & 0 & -r^{-2}\sin^{-2}\theta \end{pmatrix}. \tag{22.15} \]

Invariant volume integrals have the measure

\[ \int d^4x\sqrt{-g} = c \int dt dr r^2 d\cos\theta d\phi \sqrt{A(r)B(r)}. \tag{22.16} \]

Let us now calculate the Riemann connection. Its only non-zero components are (omitting those implied by the symmetry in the first two indices):

\[ \bar{\Gamma}^{\mu\nu}_{r} = \frac{B^{'}c^2}{2A}, \quad \bar{\Gamma}^{\mu\tau}_{t} = \frac{B^{'}}{2B}, \]
\[ \bar{\Gamma}^{\mu\nu}_{r} = \frac{A^{'}}{2A}, \quad \bar{\Gamma}^{\mu\theta}_{\theta} = \frac{1}{r}, \quad \bar{\Gamma}^{\mu\phi}_{\phi} = \frac{1}{r}, \]
\[ \bar{\Gamma}^{\mu\nu}_{r} = -\frac{r}{A}, \quad \bar{\Gamma}^{\mu\theta}_{\theta} = \cot\theta, \]
\[ \bar{\Gamma}^{\mu\nu}_{r} = -\sin\theta\cos\theta, \quad \bar{\Gamma}^{\mu\phi}_{\phi} = -\frac{r}{A}\sin^2\theta. \tag{22.17} \]

The remaining non-zero components are obtained using the symmetry \( \bar{\Gamma}_{\mu\nu}^{\lambda} = \bar{\Gamma}_{\nu\mu}^{\lambda} \).

The resulting Ricci tensor is

\[ \bar{R}_{\mu\nu} \equiv \bar{R}_{\kappa\mu\nu}^{\kappa} = \partial_{\kappa}\bar{\Gamma}_{\mu\nu}^{\kappa} - \partial_{\nu}\bar{\Gamma}_{\kappa\mu}^{\kappa} - \bar{\Gamma}_{\kappa\nu}^{\tau}\bar{\Gamma}_{\mu\tau}^{\kappa} + \bar{\Gamma}_{\mu\nu}^{\tau}\bar{\Gamma}_{\kappa\tau}^{\kappa}. \tag{22.18} \]

Remembering the definition of \( \bar{\Gamma}_{\mu\nu}^{\lambda} \) in Eq. (11.22), we can rewrite the first term as

\[ \partial_{\mu}\bar{\Gamma}_{\kappa\nu}^{\kappa} = \frac{1}{2}\partial_{\mu}(g^{\kappa\lambda}\partial_{\nu}g_{\kappa\lambda}) = \frac{1}{2}\partial_{\mu}\log(-g), \tag{22.19} \]

and find a diagonal matrix with the elements [see also Eq. (12.171)]

\[ \bar{R}_{tt} = \frac{c^2B^{'}}{2A} - \frac{c^2}{4A}B^{'\left(A^{'}/A + B^{'}/B\right)} + \frac{c^2}{r}B^{'}, \tag{22.20} \]
\[ \bar{R}_{rr} = -\frac{B^{'\left(A^{'}/A + B^{'}/B\right)} + 1}{2B} + \frac{1}{4B}, \tag{22.21} \]
\[ \bar{R}_{\theta\theta} = 1 - \frac{r}{2A}\left(-\frac{A^{'}/A + B^{'}/B\right)} - \frac{1}{A^{'}, \tag{22.22} \]
\[ \bar{R}_{\phi\phi} = \sin^2\theta\bar{R}_{\theta\theta}. \tag{22.23} \]

The diagonality with respect to the spatial components \( ij \) and the vanishing of \( R_{t\theta}, R_{t\phi} \) are a consequence of the rotational invariance. The vanishing of \( R_{rt} \) follows from time reversal invariance.
Due to the occurrence of the logarithmic derivatives \( A' / A \) and \( B' / B \) one sometimes parametrizes \( A = e^a \) and \( B = e^b \) and states the components \( \bar{R}_{\mu}^\nu \) in the form

\[
\bar{R}_i^t = \frac{1}{A} \left( \frac{b' + b''}{r} + \frac{b'^2}{2} + \frac{a'b'}{4} - \frac{a'b'}{4} \right), \tag{22.24}
\]

\[
\bar{R}_r^r = \frac{1}{A} \left( \frac{-b' + b''}{r} + \frac{b'^2}{2} + \frac{a'b'}{4} \right), \tag{22.25}
\]

\[
\bar{R}_\theta^\theta = -\frac{1}{A} \left( \frac{A - 1}{r^2} + \frac{a' - b'}{2r} \right) = \bar{R}_\phi. \tag{22.26}
\]

The scalar curvature is

\[
\bar{R} = \bar{R}_\mu^\mu = \frac{1}{A} \left[ b'' - \frac{a'b'}{2} + \frac{b'^2}{2} + \frac{2(b' - a')}{r} + \frac{2(1 - A)}{r^2} \right]. \tag{22.27}
\]

The Einstein tensor (11.142) has the components

\[
\bar{G}_i^t = \frac{1}{A} \left( \frac{a' - 1 - A}{r^2} \right), \tag{22.28}
\]

\[
\bar{G}_r^r = \frac{1}{A} \left( \frac{b' - 1 - A}{r^2} \right), \tag{22.29}
\]

\[
\bar{G}_\theta^\theta = -\frac{1}{A} \left( \frac{b''}{2} - \frac{a'b'}{4} + \frac{b'^2}{4} - \frac{a' - b'}{2r} \right) = \bar{G}_\phi, \tag{22.30}
\]

and the trace \( \bar{G} = -\bar{R} \).

For completeness, we also give the relation between the components of the Ricci tensor in the spatially isotropic coordinates (22.10) with differentials \( dx^\mu = (dt, dx^{\nu_1}, dx^{\nu_2}, dx^{\nu_3}) \) and those in the standard coordinates. The spatial components \( G_{ij} \) of the Einstein tensor may be parametrized as \( \bar{G}^{(0)}(r''')\delta_{ij} + \bar{G}^{(2)}(r'')x^i x^j / r'^2 \). Contracting this with \( dx^{\nu_1} dx^{\nu_2} \), we obtain \( \bar{G}_{ij} dx^{\nu_1} dx^{\nu_2} = [\bar{G}^{(0)}(r'') + \bar{G}^{(2)}(r'') dr^{\nu_2} + \bar{G}^{(0)}(r'') r'^2 (d\theta'^2 + \sin^2 \theta d\phi'^2) \), and identify using (22.12)

\[
\bar{G}_{rr}(r) = \frac{H(r''')}{J(r)} \left[ \bar{G}^{(0)}(r'') + \bar{G}^{(2)}(r'') \right], \quad \bar{G}_{\theta\theta}(r) = \frac{\bar{G}_{\theta\theta}(r)}{\sin^2 \theta} = r'^2 \bar{G}^{(2)}(r'') = \frac{r^2}{J(r'')} \bar{G}^{(2)}(r''), \tag{22.31}
\]

or

\[
\bar{G}_r^r(r) = \frac{1}{J(r)} \left[ \bar{G}^{(0)}(r'') + \bar{G}^{(2)}(r'') \right], \quad \bar{G}_\theta^\theta(r) = \bar{G}_\phi^\phi(r) = \frac{1}{J} \bar{G}^{(2)}(r''), \tag{22.32}
\]

where \( r = r'' \sqrt{J(r''')} \). For the component \( \bar{G}_i^t(r) \) the relation is \( \bar{G}_i^t(r) = \bar{G}_i^t(r'') \). The Einstein tensor in spatially isotropic coordinates with the metric (22.10) has the form

\[
\bar{G}_i^t = -\frac{1}{J} \left( \bar{j} + \frac{2}{r^2} \bar{j} + \frac{1}{4} r^2 \bar{j}^2 \right), \tag{22.33}
\]

H. Kleinert, GRAVITY WITH TORSION
\[ \bar{G}^{(0)} = -\frac{1}{2} \left[ \dot{h} + \dot{j} + \frac{1}{r^2} \left( \dot{h} + \dot{j} \right) + \frac{1}{2} h^2 \right], \quad (22.34) \]
\[ \bar{G}^{(2)} = \frac{1}{2} \left[ \dot{h} + \dot{j} - \frac{1}{r^2} \left( \dot{h} + \dot{j} \right) - \frac{1}{2} h^2 + \dot{h} j - \frac{1}{2} j^2 \right], \quad (22.35) \]

where
\[ H \equiv e^h, \quad J \equiv e^j, \quad (22.36) \]

and a dot denotes the derivative with respect to \( r'' \), so that \( \dot{f} = f' \sqrt{J/A} \). The trace of the Einstein tensor is
\[ \bar{G} = -\frac{1}{J(r')} \left[ \dot{h} + 2 \dot{j} + \frac{1}{2} \left( h^2 + \dot{h} j + \dot{j}^2 \right) + \frac{2}{r^2} \left( \dot{h} + 2 \dot{j} \right) \right] = \bar{R}. \quad (22.37) \]

Relation with the previous standard expressions (22.28)–(22.30) is established by inserting from (22.12):
\[ \dot{j} = \frac{2}{r^2} \left( A^{-1/2} - 1 \right), \quad (22.38) \]
so that
\[ \ddot{j} = -\frac{2}{r^2} \left( A^{-1/2} - 1 \right) - \frac{\sqrt{J}}{r^3} a' = -\frac{2J}{r^3} \left( A^{-1/2} - 1 + ra' \right), \quad (22.39) \]
and (22.33) reduces immediately to (22.28). The relations (22.32) yield
\[ G_r^r = -\frac{1}{2J} \left[ \frac{2}{r^2} \left( \dot{h} + j \right) + h j + \frac{1}{2} j^2 \right], \quad (22.40) \]
\[ G_\theta^\theta = G_{\phi}^{\phi} = \frac{1}{2} G_r^r + \frac{1}{2J} \left( \dot{h} + j + \frac{1}{2} h^2 - \frac{1}{2} h j - \frac{1}{4} j^2 \right), \quad (22.41) \]

which coincide indeed with (22.29) and (22.30).

### 22.2 The Schwarzschild Solution

We are now ready to solve Einstein’s equation with a point mass at the origin. The Einstein tensor \( G_{\mu\nu} \) has to vanish everywhere in spacetime, except at the origin. Hence also \( \bar{R}_{\mu\nu} = \bar{G}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \bar{G} \) vanishes. We shall solve the equations first in the standard metric (22.13).

#### 22.2.1 Standard Coordinate System

Setting the components of the Ricci tensor in Eqs. (22.21)–(22.23) equal to zero we have
\[ \bar{R}_t^t - \bar{R}_r^r = -\frac{1}{r A} \left( a' + b' \right) = 0 \quad (22.42) \]
and thus $a' = -b'$, implying that $A(r)B(r) = \text{const}$. Imposing the boundary condition that far away from the central mass, the coordinates $r, \theta, \phi$ should reduce to the spherical coordinates of a Minkowski space, we require

$$g_{\mu\nu} \to \begin{pmatrix} c^2 & -1 \\ -r^2 & -r^2 \sin^2 \theta \end{pmatrix}.$$  \hfill (22.43)

Hence $A(r), B(r)$ must tend to unity:

$$A(r) \to 1, \quad B(r) \to 1,$$  \hfill (22.44)

and this fixes the constant product $A(r)B(r)$ to be equal to unity, so that

$$A(r) = \frac{1}{B(r)}.$$  \hfill (22.45)

Inserting this into Eq. (22.22) and setting $\bar{R}_{\theta\theta} = 0$, we find the equation

$$\bar{R}_{\theta\theta} = 1 - B' r - B = 1 - (Br)' = 0,$$  \hfill (22.46)

which is solved by

$$B = 1 + \frac{b}{r},$$  \hfill (22.47)

Since $B(r)$ specifies the component $g_{00}(r)$ of the metric tensor and since we want that $g_{00}(r)$ has, for large $r$, the Newtonian limit $1 - 2GM/c^2 r$, we choose the constant $b = -2GM/c^2$, and see that within the present parametrization, the Newtonian limit also happens to coincide with the exact solution for $g_{00}(r)$:

$$B(r) = g_{00}(r) = 1 - \frac{2GM}{c^2 r}.$$  \hfill (22.48)

Inserting $A = -1/B$ from Eq. (22.45) into (22.21), we find that the $rr$ -component of the Ricci tensor is related to the $\theta\theta$ -component in (22.22) by

$$\bar{R}_{rr} = -\frac{B''}{2B} - \frac{B'}{rB} = \frac{\bar{R}_{\theta\theta}}{2rB}.$$  \hfill (22.49)

This vanishes automatically with $\bar{R}_{\theta\theta} = 0$. The same is true for $\bar{R}_{\phi\phi}$.

Thus we arrive at the Schwarzschild metric in the standard form:

$$ds^2 = \left(1 - \frac{2MG}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2MG}{c^2 r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin \theta d\phi^2).$$  \hfill (22.50)
Using equations (22.11)–(22.10), we find the isotropic form

\[ ds^2 = \frac{(1 - MG/2r''c^2)^2}{(1 + MG/2r''c^2)^2} c^2 dt^2 - \left( 1 + \frac{MG}{2r''c^2} \right)^4 \left[ (dr'')^2 + r''^2 (d\theta)^2 + r''^2 \sin^2 \theta d\phi^2 \right]. \]

(22.51)

where

\[ r'' = \frac{r}{2} \left[ 1 - \frac{MG}{c^2 r} + \left( \frac{1 - 4MG}{c^2 r} \right)^{1/2} \right], \]

(22.52)

or

\[ r = r'' \left( 1 + \frac{MG}{2r''c^2} \right)^2. \]

(22.53)

When discussing the corrections to Newton’s equations of motion due to general relativity it has become customary to follow Eddington and Robertson\(^1\) and expand the isotropic metric (22.51) in power of the small parameter \(MG/2r''c^2\), leading to the invariant square distance

\[ ds^2 = \left( 1 - 2\alpha \frac{MG}{r''c^2} + 2\beta \frac{M^2G^2}{r''^2c^4} + \ldots \right) c^2 dt^2 - \left( 1 + 2\gamma \frac{MG}{r''c^2} + 2\delta \frac{M^2G^2}{r''^2c^4} + \ldots \right) (dr'')^2 + r''^2 (d\theta)^2 + r''^2 \sin^2 \theta d\phi^2, \]

(22.54)

where Newton’s law implies \(\alpha = 1\), and Einstein’s field equations specify

\[ \beta = 1, \quad \gamma = 1, \quad \delta = 3/2, \quad \ldots. \]

(22.55)

At the level of the linearized theory where \(h_{ab} = g_{ab} - \eta_{ab}\) was determined by (18.32), we can only determine \(\gamma = 1\) beyond the Newton parameter \(\alpha = 1\). The parameters \(\beta, \delta, \ldots\) lie beyond the linear approximation.

It is conceivable that experimental observations require eventually a revision of Einstein’s field equations. It is therefore useful to test the higher expansion terms separately in different experiments. Alternative theories will lead to directly verifiable consequences if they differ in the lowest corrections \(\alpha, \beta, \gamma\). For example, the scalar tensor theory of Brans and Dicke [1] has \(\gamma = (\omega + 1)/(\omega + 2)\) instead of \(\gamma = 1\), where \(\omega\) is an adjustable parameter. This theory was invented to explain a three arc-second discrepancy in the precession of the perihelion of mercury to be discussed in the next section.

Going back to the standard form of the Schwarzschild metric, the parameters \(\beta, \gamma\) appear as follows

\[ r = r' \left( 1 + \frac{MG}{r''c^2} + \ldots \right) \]

(22.56)

---

\(^1\)See H.P. Robertson, in *Space Age Astronomy*, ed. by A.J. Deutsch, W.B. Klemperer (Ac. Press, New York, 1962) p. 228
\[ ds^2 = \left[ 1 - 2\gamma \frac{MG}{r c^2} + 2(\beta - \gamma) \frac{M^2 G^2}{r^2 c^4} + \ldots \right] c^2 dt^2 - \left( 1 + 2\gamma \frac{MG}{r c^2} + \ldots \right) dr^2 - r^2 \left( d\theta^2 + \sin^2 \theta d\phi^2 \right). \quad (22.57) \]

### 22.2.2 Spatially Isotropic Coordinate System

Let us find the spatially isotropic metric (22.51) directly from the vanishing of the components of the Einstein tensor in Eqs. (22.35)–(22.35), or equivalently, in (26.52), (26.53). The vanishing of the combination \( \bar{G}_r^r + \bar{G}_\theta^\theta \) implies the differential equation

\[ \ddot{h} + \ddot{j} + \frac{3}{r} (\dot{h} + \dot{j}) + \frac{1}{2} (\dot{h} + \dot{j})^2 = 0, \quad (22.58) \]

which can be rewritten as

\[ \frac{d}{dr} \left[ r^3 (\dot{h} + \dot{j}) \right] + \frac{r^3}{2} (\dot{h} + \dot{j})^2 = 0. \quad (22.59) \]

This is solved by

\[ \dot{h} + \dot{j} = \frac{1}{\sqrt{H J}} c_1^2 r^3, \quad (22.60) \]

where \( c_1^2 \) is a constant of integration. From the vanishing of \( G_t^t + 2G_\theta^\theta - 3G_r^r \) we obtain

\[ \frac{d}{dr} \left[ r^{n2} \dot{h} \right] + \frac{1}{2r^{n2}} \dot{h} (\dot{h} + \dot{j}) = 0, \quad (22.61) \]

which is solved by

\[ \dot{h} = \frac{1}{\sqrt{H J}} c_2 r^{n2}, \quad (22.62) \]

where \( c_2 \) is a second constant of integration. Inserting this into (22.60) yields

\[ \dot{j} = \frac{1}{\sqrt{H J}} \left( \frac{c_1^2}{r^{n3}} - \frac{c_2}{r^{n2}} \right), \quad (22.63) \]

Equation (22.60) implies that

\[ \frac{d}{dr} \sqrt{H J} = \frac{1}{2} c_1^2 r^{n2}, \quad (22.64) \]

from which we find

\[ \sqrt{H J} = c_3 - \frac{c_1^2}{4r^{n2}}, \quad (22.65) \]

with a third constant of integration \( c_3 \).
At infinity, the metric is supposed to be Minkowskian, so that $c_3 = 1$. Inserting (22.65) into (22.62) yields the differential equation

$$\dot{h} = \frac{c_2/r''^2}{1 - c_1^2/4r''^2} = \frac{c_2}{c_1} \left( \frac{1}{r'' - c_1/2} - \frac{1}{r'' + c_1/2} \right),$$

(22.66)

which is solved by

$$H = \left[ \frac{1 - c_1/2r''}{1 + c_1/2r''} \right]^{c_2/c_1}.$$

(22.67)

Together with (22.65) we obtain

$$J = \frac{(1 + c_1/2r'')^{2+c_2/c_1}}{(1 - c_1/2r'')^{-2+c_2/c_1}}.$$

(22.68)

The condition that the asymptotic behavior of $H$:

$$H \equiv 1 - c_2/r'' + \ldots$$

(22.69)

yields Newton’s law requires $c_2 = 2MG/c^2$ [compare (22.54)].

The two constants of integration $c_1$ and $c_2$ are not independent of one another. Inserting (22.62) and (22.63) into $\bar{G}_r^r = 0$ of Eq. (22.33), we obtain

$$\bar{G}_r^r = -\frac{C}{2JHJ} \left( 2c_2^2 - \frac{1}{2}c_1^2 \right).$$

(22.70)

Hence the field equation $\bar{G}_r^r$ fixes $c_2 = 2c_1$, so that

$$H = \frac{(1 - MG/2c^2r'')^2}{(1 + MG/2c^2r'')^2}, \quad J = \left( 1 + MG/2c^2r'' \right)^4,$$

(22.71)

as in the spatially isotropic metric (22.51).

### 22.3 Planetary Motion in Schwarzschild Metric

Consider now the motions of a point particle in the Schwarzschild metric associated with a massive particle at the origin. We have to solve the differential equation (11.25) for the geodesic trajectory:

$$\frac{d^2x^\mu}{d\tau^2} + \bar{\Gamma}_\lambda^\mu_{\nu} \frac{dx^\lambda}{d\tau} \frac{dx^\nu}{d\tau} = 0.$$

(22.72)

Using the components of the Riemann connection listed in Eq. (22.17), this leads to the following equations:

$$\frac{d^2t}{d\tau^2} + B' \frac{dt}{d\tau} \frac{dr}{d\tau} = 0,$$

(22.73)
\[
\frac{d^2 r}{d\tau^2} + \frac{A'}{2A} \left( \frac{dr}{d\tau} \right)^2 - \frac{r}{A} \left( \frac{dt}{d\tau} \right)^2 - \frac{r \sin^2 \theta}{A} \left( \frac{d\phi}{d\tau} \right)^2 + \frac{B' c^2}{2A} \left( \frac{dt}{d\tau} \right)^2 = 0, \tag{22.74}
\]

\[
\frac{d^2 \theta}{d\tau^2} + \frac{2 d\theta}{r \, d\tau} \frac{dr}{d\tau} - \sin \theta \cos \theta \left( \frac{d\phi}{d\tau} \right)^2 = 0, \tag{22.75}
\]

\[
\frac{d^2 \phi}{d\tau^2} + \frac{2 d\phi}{r \, d\tau} \frac{dr}{d\tau} + 2 \cot \theta \frac{d\phi}{d\tau} \frac{d\theta}{d\tau} = 0. \tag{22.76}
\]

Because of rotational invariance, the particle motion must take place in a plane through the origin which can be chosen to be as the equatorial plane \( \theta \equiv \pi/2 \) of the spherical coordinate system. Then we have fulfilled the equation for \( \theta \) identically. We now take the equation of motion for \( t \) and \( \phi \), divide them by \( dt/d\tau \), \( d\phi/d\tau \), respectively, and find

\[
\frac{d}{d\tau} \left( \log \frac{dt}{d\tau} + \log B \right) = 0, \tag{22.77}
\]

\[
\frac{d}{d\tau} \left( \log \frac{d\phi}{d\tau} + \log r^2 \right) = 0, \tag{22.78}
\]

which yields two constants of motion. The first is chosen such as to fix the relation between \( t \) and \( \tau \):

\[
\frac{dt}{d\tau} = \frac{1}{B(r)}. \tag{22.79}
\]

The other is used to set

\[
r^2 \frac{d\phi}{d\tau} = l = \text{const.}, \tag{22.80}
\]

the constant \( l \) being the orbital angular momentum per unit mass. The same constant of motion is introduced in the derivation of ordinary Kepler motion from Newton’s equation of motion. Using (22.79) and (22.80), the radial equation becomes

\[
\frac{d^2 r}{d\tau^2} + A' \left( \frac{dr}{d\tau} \right)^2 - \frac{l^2}{r^2 A} + \frac{B' c^2}{2AB^2} = 0. \tag{22.81}
\]

For this we can find a first integral of motion by multiplying it by \( 2A \frac{dr}{d\tau} \) and rewriting it as

\[
\frac{d}{d\tau} \left[ A \left( \frac{dr}{d\tau} \right)^2 + \frac{l^2}{r^2} - \frac{c^2}{B(r)} \right] = 0. \tag{22.82}
\]

This gives the constant of motion

\[
A \left( \frac{dr}{d\tau} \right)^2 + \frac{l^2}{r^2} - \frac{c^2}{B(r)} = \text{const} \equiv -\varepsilon c^2. \tag{22.83}
\]
This constant is determined by inserting for \( l \) the expression \( r^2 d\phi / dr \), for \( 1/B \) the expression \( dt/d\tau = B (dt/d\tau)^2 \), and observing that the left-hand side becomes \(-ds^2/d\tau^2\). This fixes \( \varepsilon = 1 \).

Note that it is also possible to study light rays in the Schwarzschild field. For a light ray \( ds = 0 \), i.e., the proper time does not change. This is achieved by choosing the constant of integration \( \varepsilon \) to vanish, in which case the parameter \( \tau \) introduced initially to describe the trajectory is no longer the proper time, as its notation suggests, but just an arbitrary parameter to describe the light trajectory. After all, the equation of motion (22.72) is homogeneous in \( \tau \) such that an overall factor cancels out. We may imagine having chosen for this parameter the proper time \( \tau \) divided by \( \sqrt{m} \) for a massive particle and then gone to the limit of zero mass, in which case \( d\tau\to 0 \), while \( d\tau/\sqrt{m} \) still serves to parametrize the orbit. In order to keep the freedom in describing massive as well as massless trajectory it is useful to leave the constant of motion \( \varepsilon \) arbitrary and simply remember that the proper time is \( ds^2 = \varepsilon c^2 dr^2 \).

For large enough orbits, \( A(r) \) is positive and we find the inequality
\[
\varepsilon + \frac{l^2}{c^2 r^2} \leq \frac{c^2}{B},
\]
where the equal sign holds for the extrema of the orbit.

In terms of the time variable \( t \), the equations of motion read
\[
\begin{align*}
B &= 1 - \frac{2MG}{rc^2} = 1 + \frac{2\Phi}{c^2}, \\
A &= 1, \\
\frac{1 - \varepsilon}{2} &= \frac{\text{energy}}{\text{mass} c^2} = \frac{E}{mc^2}.
\end{align*}
\]
Indeed, neglecting \( 2MG/rc^2 \) in \( B \), we arrive at
\[
\begin{align*}
\dot{r}^2 &\approx l, \\
\dot{r}^2 + \frac{l^2}{2r^2} + \Phi &= \frac{2E}{m}.
\end{align*}
\]

The full equations (22.85)–(22.86) are solved in a similar way as Newton’s equations. We use (22.85) to change \( \dot{r} \) into \( dr/d\phi = \dot{r}^2/lB \), and arrive at the differential equation
\[
\frac{A}{r^4} \left( \frac{dr}{d\phi} \right)^2 + \frac{1}{r^2} - \frac{1}{l^2 B(r)} = -\frac{\varepsilon}{l^2}.
\]
This is directly integrated to
\[
\phi = \pm \int \frac{dr}{r^2} \sqrt{A(r)} \left[ \frac{1}{r^2 B(r)} - \frac{\varepsilon}{l^2} - \frac{1}{r^2} \right]^{-1/2}.
\] (22.94)
The zeros of the square root are the turning points of the orbit. They are given by
\[
\frac{1}{r^2_\pm} - \frac{1}{l^2 B(r_\pm)} = -\frac{\varepsilon}{l^2}.
\] (22.95)
In terms of \(r_\pm\), we can express the constants of motion as follows
\[
\varepsilon = \frac{r_+^2 B(r_+) - r_-^2 B(r_-)}{r_+^2 - r_-^2}, \quad l^2 = \frac{1}{B(r_+)} - \frac{1}{B(r_-)} \frac{1}{r_+^2 - r_-^2}.
\] (22.96)
In the limit of a circular orbit, they become
\[
\varepsilon = \frac{1}{B} \left( 1 - \frac{r B'}{2B} \right), \quad l^2 = \frac{B' r_+^3}{2B^2}.
\] (22.97)
The azimuthal advance between the two turning points is
\[
\phi(r_+) - \phi(r_-) = \int_{r_-}^{r_+} \frac{dr}{r^2} \sqrt{A(r)} \left( \frac{1}{l^2 B(r)} - \frac{\varepsilon}{l^2} - \frac{1}{r^2} \right)^{-1/2}
\] (22.98)
\[
= \int_{r_-}^{r_+} \frac{dr}{r^2} \sqrt{A(r)} \left\{ \frac{r_-^2 [B^{-1}(r) - B^{-1}(r_-)] - r_+^2 [B^{-1}(r) - B^{-1}(r_+)]}{r_-^2 r_+^2 (B^{-1}(r) - B^{-1}(r_-))} - \frac{1}{r^2} \right\}^{-1/2}.
\]
Inserting the metric (22.57), i.e.,
\[
A(r) = 1 + 2\gamma \frac{MG}{rc^2} + \ldots, \quad B(r) = 1 - 2\frac{MG}{rc^2} + 2(\beta - \gamma) \frac{M^2 G^2}{r^2 c^4} + \ldots, \]
(22.99) (22.100)
such that
\[
B^{-1}(r) = 1 + \frac{2MG}{rc^2} + 2(2 - \beta + \gamma) \frac{M^2 G^2}{r^2 c^4} + \ldots, \]
(22.101)
we see that, to this order, the expression in the curly is a quadratic function in \(1/r\) which vanishes at \(r = r_+\) and \(r = r_-\). It may be written as a ratio \(-[r_-^2 N_-(r_-) + r_+^2 N_+(r_+)]/r_-^2 r_+^2 N_-(r_-, r_+)+\) with
\[
N_\pm(r, r_\pm) = \frac{2MG}{c^2} \left( \frac{1}{r} - \frac{1}{r_\pm} \right) + (2 \mp \beta \pm \gamma) \frac{2MG}{c^2} \left( \frac{1}{r_-^2} - \frac{1}{r_+^2} \right),
\] (22.102)
and therefore as
\[ \{ \ldots \} = C \left( \frac{1}{r_+} - \frac{1}{r} \right) \left( \frac{1}{r} - \frac{1}{r_-} \right). \tag{22.103} \]

where the constant \( C \) is determined by going in the curly brackets of Eq. (22.98) to the limit \( r \to \infty \), where \( B(r) \to 1 \). This yields
\[ C = \frac{r_+^2 [1 - B^{-1}(r_+)] - r_-^2 [1 - B^{-1}(r_-)]}{r_+ r_- [B^{-1}(r_+) - B^{-1}(r_-)]}. \tag{22.104} \]

Expanding the right-hand side with respect to the small quantities \( MG/r \pm c^2 \), we obtain
\[ C \approx 1 - (2 - \beta + \gamma) \frac{MG}{c^2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right), \tag{22.105} \]

Inserting \( C \) into (22.103), and this further into the curly brackets of (22.98), the azimuthal advance becomes
\[ \phi(r_+) - \phi(r_-) = \left[ \left( \frac{1}{r_+} + \frac{1}{r_-} \right) - \frac{1}{r} \right] \frac{1 + \gamma \frac{MG}{c^2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right)}{\sqrt{\left( \frac{1}{r_-} - \frac{1}{r} \right) \left( \frac{1}{r} - \frac{1}{r_+} \right)}} \] \( \tag{22.106} \)

The integral can be performed with the help of a variable \( \vartheta \) defined by
\[ \frac{1}{r} = 1 + \frac{1}{2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right) - \frac{1}{2} \left( \frac{1}{r_-} - \frac{1}{r_+} \right) \sin \vartheta \] \( \tag{22.107} \)

where \( \vartheta \in (-\pi/2, \pi/2) \) describes the orbit from \( r_- \) to \( r_+ \). This gives the azimuthal advance
\[ \phi(r_+) - \phi(r_-) = \left[ 1 + \frac{1}{2} (2 - \beta + 2\gamma) \frac{MG}{c^2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right) \right] \vartheta \bigg|_{\pi/2}^{\pi/2} \]
\[ - \frac{1}{2} \gamma \frac{MG}{c^2} \left( \frac{1}{r_+} - \frac{1}{r_-} \right) \cos^4 \vartheta \bigg|_{-\pi/2}^{\pi/2}. \tag{22.108} \]

In Newton’s theory, the right-hand side is equal to \( \pi \), so that the elliptical orbit does not advance in accordance with Kepler’s law. The relativistic correction produces an \textit{advance per total revolution}
\[ \Delta \phi = 2 \left( 1 - \frac{\beta}{2} + \gamma \right) \frac{MG}{c^2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right) \pi. \tag{22.109} \]
Using the length of the semimajor axis $a$ and the excentricity parameter $e$, the radii $r_{\pm}$ are equal to $(1 + e)a$, and we can write

$$\Delta \phi = 6\pi MG c^2 \left( \frac{2 - \beta + 2\gamma}{3} \right) \frac{1}{(1 - e^2)a}.$$  \hspace{1cm} (22.110)

In Einstein’s theory where the parameters $\alpha, \beta, \gamma$ are unity [see Eq. (22.55)], the expression in parentheses is equal to unity:

$$\frac{2 - \beta + 2\gamma}{3} = 1.$$  \hspace{1cm} (22.111)

In the Brans-Dicke modification of Einstein’s theory which contains a scalar field in addition to the metric tensor [1]:

$$\frac{2 - \beta + 2\gamma}{3} = \frac{\omega + \frac{1}{3}}{\omega + 2}.$$  \hspace{1cm} (22.112)

Inserting the parameters of the planet Mercury

$a = 5.791 \times 10^{12}$ cm, $e = .2056$, $(1 - e^2)a = 5.53 \times 10^{12}$ cm,

this gives

$$\Delta \phi = 0.1038''.$$  \hspace{1cm} (22.113)

per revolution. In a century, Mercury makes 415 revolutions such that $\Delta \phi$ piles up to a total advance of the perihelion of

$$\Delta \phi = 43.03'' \text{ per century.}$$  \hspace{1cm} (22.114)

This agrees with the most recent analysis of astronomical data by Clemence in 1947.

A similar comparison for the planets Venus, Earth, and the asteroid Icarus yields the parameters shown in Table 22.1. The precession of the perihelion of mercury is the most significant test of Einstein’s theory since it involves the correction parameter $\beta$ of the non-linearized part of the field equation.

Notice that the table provides some evidence against the possibility that all the advance might be caused by an unexpected extraordinarily large quadrupole moment of the sun (due to some anomalous internal mass distribution). Since a quadrupole moment changes the gravitational potential by additional terms of the order $1/r^3$, the precession of different planets would have to decrease roughly like $1/r^2$; whereas the table rather indicates a $1/r$-decrease, this being in agreement with the relativity prediction [4].

<table>
<thead>
<tr>
<th>Celestial Bodies</th>
<th>a(10^{12} \text{cm})</th>
<th>e</th>
<th>Revolutions</th>
<th>$\Delta \phi$ in arc sec</th>
<th>astron. observ.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercury</td>
<td>5.791</td>
<td>0.2056</td>
<td>415</td>
<td>43.03</td>
<td>43.11± 0.45</td>
</tr>
<tr>
<td>Venus</td>
<td>10.821</td>
<td>0.0068</td>
<td>149</td>
<td>8.6</td>
<td>8.4± 4.8</td>
</tr>
<tr>
<td>Earth</td>
<td>14.96</td>
<td>0.0168</td>
<td>100</td>
<td>3.8</td>
<td>5.0± 1.2</td>
</tr>
<tr>
<td>Icarus</td>
<td>16.1</td>
<td>0.1827</td>
<td>85</td>
<td>10.3</td>
<td>9.8± 0.8</td>
</tr>
</tbody>
</table>
22.4 Orbit of Light Ray

Let us use the Schwarzschild metric to study the exact orbit of a light ray in the neighborhood of a central mass. For a small mass, the result was derived before in Section 21.2 within linearized gravity. For large masses, we have to solve directly the equation of motion (22.86) for $\varepsilon = 0$. We consider a light ray coming in from infinity, approaching the central mass up to a radius $r_0$, and leaving again to infinity. At $r_0$, the derivative $dr/d\phi$ vanishes so that we can identify

$$\frac{l^2}{r_0^2} - \frac{c^2}{B(r_0)} = 0,$$

and the equation (22.94) becomes

$$\phi = \int \frac{dr}{r} \sqrt{A(r)} \left[ \left( \frac{r}{r_0} \right)^2 \frac{B(r_0)}{B(r)} - 1 \right]^{-1/2}.$$

Inserting the expansion (22.100), (22.101) for $A, B^{-1}$, the integral becomes

$$\int \frac{dr}{r} \left[ \left( \frac{r}{r_0} \right)^2 - 1 \right]^{-1/2} \left[ 1 + \gamma \frac{MG}{c^2 r} + \frac{MGr}{c^2 r_0(r_0 + r)} + \ldots \right].$$

This can be integrated to

$$\phi(r) - \phi_0 = -\arcsin\left( \frac{r_0}{r} \right) + \frac{MG}{c^2 r_0} \left[ 1 + \gamma - \gamma \sqrt{1 - \left( \frac{r_0}{r} \right)^2} - \sqrt{\frac{r - r_0}{r + r_0}} \right] + \ldots,$$

leading to a total deflection angle

$$\Delta \phi = \frac{4MG}{c^2 r_0} \frac{1 + \gamma}{2}.$$

Within this lowest approximation, the distance $r_0$ can be replaced by the impact parameter $b$. Note that in contrast to the perihelion precession of a planet, this depends only on the parameter of the linearized theory $\gamma$. Inserting $\gamma = 1$ and setting $r_0$ approximately equal to the impact parameter $b$ leads back to our earlier result (21.23).
22.5 Schwarzschild Singularity

It should be realized that the Schwarzschild metric is singular when the parameter $MG/c^2r$ becomes equal to unity. Physically this is the point where the velocity of a mass point in a Newtonian orbit will have a velocity of the order of the speed of light. The length

$$r_S = \frac{2GM}{c^2}$$

is called Schwarzschild radius. For $r < r_S$, $g_{00}$ and $g_{rr}$ change sign which causes difficulties of physical interpretation. The equation of motion in the Schwarzschild metric makes sense only for $r > r_S$. For most stars, this is no practical problem since their radius is larger than $r_S$. The sun with $M_\odot = 2 \times 10^{33}$ g, for example, has

$$r_S = 2.952 \text{ km},$$

which lies well within the solar radius $R_\odot = 0.696 \times 10^6$ km. Since the Schwarzschild solution was derived in empty space, it is correct only outside the star’s surface and the Schwarzschild radius is never reached. In fact, if a star is to be smaller than its Schwarzschild radius, its mass density has to be larger than

$$M \frac{4}{3} \pi r_S^3 = \frac{3c^6}{32\pi G^3 M^2}.$$

For the sun, this is equal to $\approx 10^{16}$ g/cm$^3$, which is several orders of magnitude larger than the density of nuclear matter ($\approx 1.67 \times 10^{-24}$ g/10$^{-36}$ cm$^3 \sim 10^{12}$ g/cm$^3$). Only if all atoms are stripped of electrons under extreme pressures can such a density be achieved.

22.6 Black Holes

The Schwarzschild radius has interesting physical properties. A clock placed at radius $r_0$ has a proper time

$$d\tau = \sqrt{1 - \frac{r_S}{r_0}} dt, \quad r_S = \frac{2GM}{c^2},$$

which approaches zero for $r_0 \rightarrow r_S$. Thus the time at $r_0 = r_S$ runs infinitely slow as compared with a clock placed at infinity. As a consequence, a light signal emitted near the Schwarzschild radius arrives at a distant observer with an extremely large red shift.

Actually, this calculation is academic. It is impossible for any material particle to have a fixed radius $r_0 \geq r_S$. Only a light ray can, since its proper time does not change ($d\tau = 0$). Any material body will fall through the singularity towards the center. Remarkably, for a small observer sitting on such a body, nothing extraordinary will happen. He will experience small tidal forces which grow to infinity only...
when approaching the origin $r = 0$. In the neighborhood of $r \sim r_S$ everything is perfectly smooth.

In order to see this, consider the component $R_{1001}$ responsible for the tidal forces [see (21.41)]. Using $\Gamma_{\mu\nu}^\lambda$ from Eqs. (22.17), we calculate

$$R_{1010} = \frac{r_S}{r} \frac{1}{1 - r_S/r}$$

(22.124)

which has a singularity. This, however, is only present in the Schwarzschild coordinates, and these have been adapted to suit an observer at infinity. For an observer moving in the neighborhood of $r \sim r_S$, it is preferable to use geodesic coordinates defined by a freely falling particle at $r_0$. They are obtained by the coordinate transformation [recall Subsection 12.6.1]

$$x'^0 = c(t - t_0) \sqrt{1 - \frac{r_S}{r_0}} + \ldots, \quad x'^1 = (r - r_0) \sqrt{1 - \frac{r_S}{r_0}} + \ldots$$

$$x'^2 = r_0 \left( \theta - \frac{\pi}{2} \right) + \ldots, \quad x'^3 = r_0 \left( \phi - \phi_0 \right) + \ldots$$

This transformation makes the metric at the point $x'_0$ Minkowskian, $g_{\mu\nu}(x_0) = \eta_{\mu\nu}$.

The only non-vanishing matrix elements of the transformation matrix are at $r_0$

$$\frac{\partial x'^0}{\partial x^0} = \sqrt{1 - \frac{r_S}{r_0}}, \quad \frac{\partial x'^1}{\partial x^1} = \sqrt{1 - \frac{r_S}{r_0}}, \quad \frac{\partial x'^2}{\partial x^2} = \frac{1}{r_0}, \quad \frac{\partial x'^3}{\partial x^3} = \frac{1}{r_0}$$

(22.125)

and we find for the transformed Riemann tensor $x$

$$R'_{1010} = \frac{\partial x'^0}{\partial x'^0} \frac{\partial x'^\alpha}{\partial x'^\beta} \frac{\partial x'^\mu}{\partial x'^\nu} R_{\alpha\mu\nu} = \frac{\partial x^0}{\partial x^0} \frac{\partial x^1}{\partial x^1} \frac{\partial x^0}{\partial x^3} \frac{\partial x^1}{\partial x^3} R_{1010} = \frac{r_S}{r_0^3}.$$  

(22.126)

The singularity has disappeared. A freely falling observer is unaware of the Schwarzschild singularity. Only at $r = 0$ will he run into a true physical singularity.

To an observer at infinity, however, the singularity is a reality. It separates the space into two regions, an outside one, which can communicate with the observer at infinity via light signals, and an inside one which cannot. The surface $r = r_S$ is called an event horizon. This is a global property of the solution which no local observer can see.

In order to understand the motion better it is useful to reparametrize the Schwarzschild metric with the dimensionless Kruskal coordinates [5]. They make the coordinate singularity disappear but display clearly the event horizon. One chooses outside the Schwarzschild radius

$$u = \sqrt{\frac{r}{r_S} - 1} e^{r/2r_S} \cosh(t/2r_S), \quad v = \sqrt{\frac{r}{r_S} - 1} e^{r/2r_S} \sinh(t/2r_S), \quad r > r_S,$$

(22.127)

and inside

$$u = \sqrt{1 - \frac{r}{r_S}} e^{r/2r_S} \sinh(t/2r_S), \quad v = \sqrt{1 - \frac{r}{r_S}} e^{r/2r_S} \cosh(t/2r_S), \quad r < r_S.$$  

(22.128)
The inverse of this transformation is

\[
\left( \frac{r}{r_S} - 1 \right) e^{r/r_S} = \begin{cases} u^2 - v^2 & \text{for } r < r_S \\ v^2 - u^2 & \text{for } r > r_S \end{cases},
\]

(22.129)

and

\[
t = 2r_S \tanh \left\{ \frac{u}{v} \right\} \text{ for } r < r_S \quad \text{and} \quad t = 2r_S \tanh \left\{ \frac{v}{u} \right\} \text{ for } r > r_S.
\]

(22.130)

The invariant length is given by

\[
ds^2 = \frac{4r_S^3}{r} e^{-r/r_S} [(dv)^2 - (du)^2] - r^2 [(d\theta)^2 + \sin^2 \theta (d\phi)^2]
\]

(22.131)

which shows that \(u\) is spacelike and \(v\) is timelike. Apart from \(r = 0\), the metric is completely regular.

In order to understand this parametrization, let us ignore the angles \(\theta\) and \(\phi\) and consider only the temporal and radial parts. We see that the curves \(r = \text{constant}\) are parabolas in the \(uv\)-plane, and the curves \(t = \text{const}\) are straight lines through the origin. Taking \(r_S\) as the length unit, the Schwarzschild singularity lies on the diagonal (see Fig. 22.2). It separates the inside of the Schwarzschild singularity from the outside. If we drop a particle from a point outside this radius (region I), it will fall along a smooth trajectory towards the center as indicated by the dashed line on the right-hand side. An outside observer can never follow this orbit optically, since his clock will proceed to infinity when the particle crosses the Schwarzschild radius.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{schwarzschild_diagram.png}
\caption{Maximal Schwarzschild geometry in Kruskal coordinates [6]. Coordinate \(r\) is given in units of \(r_S\).}
\end{figure}
Black Holes

22.6 Black Holes

\[ r = r_S \text{ where } u = v. \] A particle dropped from any point inside the radius \( r = r_S \) (region II) will always remain inside. No signal can ever get out from there. This is why region II is called a black hole.

Note that no material particle starting from \( r > r_S \) or \( r < r_S \) can ever reach the regions III, IV in Fig. 22.2. This part of the figure is a world entirely of its own. In fact, the space has two world sheets, \( u > v \) and \( u < v \). For each radius \( r \) and time \( t \) there are two points in \( u, v \) space belonging to the two sheets. Each of the sheets has a singularity at \( r = 0 \) corresponding to \( v^2 - u^2 = 1 \) and \( v^2 - u^2 = -1 \), respectively, and an asymptotically flat region for \( r \to \infty \) and all \( t \) (corresponding to \( u \to \infty \) and \( u \to -\infty \), respectively).

What is interesting to see is that signals starting out from region IV can travel into regions III and I. Thus, if there is any light source within the region IV, it can be seen from the outside world. This is why region IV is called a white hole. It must be noted, however, that this interpretation might be false. The paths leaving the white hole correspond to the time reversed paths of those going into the black hole. Such paths require a quantum-field-theoretic interpretation similar to the negative energy states in the classical Klein-Gordon equation. The classical description presented here is therefore not expected to be the physically correct one.

Let us look at the space \( r > r_S \) at a fixed time coordinate \( v = 0 \). Focussing attention upon the surface defined by \( \theta = \pi/2 \) we write down the spatial line element on it

\[ dt^2 = \frac{dr^2}{1 - \frac{r_S}{r}} + r^2 d\phi^2. \tag{22.132} \]

It is useful to picture this surface by embedding it mentally in a three-dimensional “hyper-space” as a surface of revolution parametrized by the fictitious coordinate \( z = z(r) \). On such a surface, a line element has the general form

\[ dt^2 = dr^2 \left[ 1 + \left( \frac{dz}{dr} \right)^2 + r^2 d\phi^2 \right]. \tag{22.133} \]

Comparing the two expressions gives

\[ \left( \frac{dz}{dr} \right)^2 = \frac{1}{1 - \frac{r_S}{r}} - 1 = \frac{r_S}{r - r_S} \tag{22.134} \]

or

\[ z(r) = 2[r_S(r - r_S)]^{1/2} + \text{const.} \tag{22.135} \]

This is the surface of revolution shown in Fig. 22.2. The structure is called a wormhole, an Einstein-Rosen bridge, or a Schwarzschild throat. For a time \( \nu \) different from zero, the surface is no longer one at fixed time \( t \). If \( \nu \) becomes larger than unity, the wormhole disappear and the two sheets of space become completely disjoint. A schematic picture of this temporary bridge between two spaces is given in Fig. 22.2.
Figure 22.3 Spatial surface $\theta = \pi/2$ at $t = 0$ for the Schwarzschild solution: Schwarzschild throat [7, 8].

Figure 22.4 Evolution of the Schwarzschild geometry in $\nu$-time (schematic [7, 8]).
22.7 Black Holes as a Result of Gravitational Collapse

The question arises as to whether there can be circumstances under which the mass of a star is squeezed to such high densities that the boundary comes to lie within the Schwarzschild radius \( r < r_S = \frac{2GM}{c^2} \). It is presently believed that this will be the case with very large stars at the end of their stellar evolution. In order to understand that we have to realize that due to the long range of gravitational forces, a larger and larger accumulation of matter at a fixed density will lead to an arbitrarily large pressure at the center. In a star, this pressure is counterbalanced by the outward pressure caused by the thermonuclear reactions. These reactions can go on only as long as the nuclear fuel is available, i.e., as long as there are light elements to be fused and heavy elements to be fissioned into stable medium-heavy ones, the most stable being \( \text{Fe}^{56} \). When these reactions end, the gravitational pressure will compress the star, within some unknown time. The compression can be stopped if there is a way of building up sufficient counterpressure.

One such mechanism is operative at densities \( \rho > 10^5 \text{g/m}^2 \). At such densities, the atoms are squeezed to distances of the order of their electronic orbits. There it becomes energetically more favorable for the electrons to detach themselves from the nuclei and move through the system almost freely. The gas of nuclei and electrons is called a plasma. A white dwarf is consists of such a plasma.

An electron gas satisfies Pauli’s exclusion principle, according to which a cell in phase space of size \( \hbar^3 \) cannot be occupied by more than one electron. If free electrons are stacked as densely as possible in momentum space, up to the Fermi momentum \( p_F \), their total number per volume is

\[
\frac{N}{V} = 2 \int_{|p| \leq p_F} \frac{d^3p}{(2\pi\hbar)^3} = \frac{1}{3 \pi^2 \hbar^3} p_F^3
\]

from which we determine the Fermi momentum as

\[
p_F = \left(3 \pi^2 \hbar^3 N/V\right)^{1/3}.
\]

The associated electronic energy is \( p_F^2/2m_e \), so that we can roughly estimate the total energy of the electron gas due to the exclusion principle

\[
E_{\text{Pauli}}^{\text{non-rel}} \approx \frac{N p_F^2}{2m_e} = N \left(3 \pi^2 \hbar^3 N/V\right)^{2/3}.
\]

For relativistic electrons, Eq. (22.136) is still valid, but for them the Fermi energy is of the order of \( c \) times the Fermi momentum. This leads to a total energy due to the exclusion principle:

\[
E_{\text{Pauli}}^{\text{rel}} \approx Ncp_F = Nc(3 \pi^2 \hbar^3 N/V)^{1/3}.
\]

A rough criterion as to which of the two formulas should be applied is given by

\[
p_F < c m_e \quad \text{\{non-rel.\}}
\]

\[
\text{rel.}\}
\]

(22.140)
Due to charge neutrality, the number of electrons is roughly equal to the number of protons in a white dwarf. Furthermore, each proton is roughly accompanied by a neutron. Hence $N \approx M/2mp$.

The above energy has to be compared with the gravitational binding energy, which is of the order of

$$E_{\text{grav}} \approx -2N \frac{GMm_p}{R}. \quad (22.141)$$

Since $V = 4\pi R^3/3$ we see that for a given mass, a non-relativistic electron gas has an energy of the form $a/R^2 - b/R$, whereas the relativistic gas has $a/R - b/R$. The first has always a stable minimum at a finite $R$. A white dwarf with relativistic electrons, on the other hand, will collapse to $R \to 0$ if $a < b$. Under this condition, the Fermi repulsion between relativistic electrons is unable to resist the gravitational pressure. The limiting mass is easily estimated from this equation to be

$$M_{\text{crit}} \approx \frac{3\sqrt{\pi}}{8m_p^2} \left( \frac{\hbar c}{G} \right)^{3/2}. \quad (22.142)$$

Expressing $G$ in terms of the Planck mass $m_{\text{pl}} = (\hbar c/G)^{1/2} \approx 2.177 \times 10^{-5} \text{g} = 1.2 \times 10^{10} \text{GeV}/c^2$, we find

$$M_{\text{crit}} \approx \frac{3\sqrt{\pi}}{8m_p^2} \left( \frac{m_{\text{pl}}}{m_p} \right)^3 m_p \approx 2.45 \times 10^{33} \text{g} \approx 1.23 M_\odot. \quad (22.143)$$

so that a white dwarf will collapse if it is only slightly more massive than our sun whose mass is $M_\odot = 1.989 \times 10^{33} \text{g}$.

A more precise formulation of the above argument was given by Chandrasekhar who found the so-called Chandrasekhar limit

$$M_{\text{crit}} = 1.44M_\odot. \quad (22.144)$$

A white dwarf of mass $M > 1.44M_\odot$ collapses if the Fermi momentum $p_F$ of its electrons becomes relativistic, i.e., larger than $m_e c$.

The Chandrasekhar limit is only roughly obeyed. At very high density the electrons can be absorbed by the protons forming neutrons in a reaction which is the inverse reaction of the weak $\beta$ decay of the neutrons. This becomes possible as soon as the electronic energy becomes larger than the neutron-proton mass difference $(m_n - m_p \approx 1.2935 \text{MeV})$. For increasing pressures, this process leads to a depletion of the electronic density, transforming gradually all protons into neutrons. At densities above $10^{13} \text{g/cm}^3$, a star consist practically only of neutrons. It is called a neutron star. Since the neutrons come to lie very close to each other, the whole star acts like a gigantic nucleus (nuclear matter) stabilized by the gravitational attraction.

In a neutron star, the neutrons form a degenerate Fermi system. Due to the short-range repulsion between the nucleons, they forma liquid. Since the neutrons
are roughly 2000 times heavier than the electrons, they start out as a non-relativistic Fermi liquid which is again stable for the same reason as above. The density can increase until the Fermi momentum of the neutrons reaches $c m_p \approx 2000 \, c m_e$.

At the core of a supernova we have to imagine a highly turbulent state of nuclear matter. The pulsations of this state will dampen out rapidly due to gravitational radiation.

The result is a fast rotating flattened neutron star with a strong magnetic field ($\approx 10^{12}$ gauss). When the rotating magnetic field sweeps through the surrounding plasma, these radiate electromagnetic waves into space, in the radio as well as in the visible spectrum. Such pulsating waves do indeed arrive at the earth with frequencies of the order of a few seconds, the light waves a little earlier than the radio waves since these are slowed down by scattering on the interstellar plasma. The source of such waves is called a pulsar. The first pulsar was discovered in 1967 by Jocelyn Bell and Anthony Hewish. Pulsars are rapidly spinning neutron stars whose lighthouse-like beams of radio waves sweep the Earth and produce highly regular radio pulses. The steadiness of the pulses makes pulsars very accurate clocks, rivaling the best atomic clocks on Earth. At present, more than 1500 radio pulsars are known in our Galaxy, and a few have been found in nearby galaxies such as the two Magellanic Clouds [9].

The neutron star becomes unstable when the Fermi momentum $p_F$ becomes larger than $m_n c$. This happens in the neighborhood of $\rho \approx 10^{16}$ g/cm$^3$. The stability limit is

$$M_{\text{crit}}(\text{neutron star}) \approx 0.3 \text{ to } 1.4 M_\odot.$$  \hspace{1cm} (22.145)

What happens, if this limit is exceeded? It is the present belief that above this limit, matter will be squeezed without limit. When in this process the radius of a star becomes smaller than the Schwarzschild radius, a black hole is formed.

One may wonder whether the unknown physics of elementary particles of extremely short distances may not always provide the star with a new mechanism stopping the collapse. It has been argued by Rhoades and Ruffini that whatever might happen, a higher modulus of compression always goes along with the speed of light $c$, there is an ultimate limit of material stiffness. From this they calculate the ultimate mass $3.2 M_\odot$, above which gravitational collapse can no longer be prevented, independent of future physical discoveries [10].

As a function of time, the stellar development is roughly the following: When nuclear fuel is exhausted, the star becomes a white dwarf. This radiates off its thermal energy. If the mass is smaller than the Chandrasekhar limit, the becomes just a cold piece of matter. If it exceeds this limit, gravitational collapse sets in at the core as soon as the thermal pressure becomes too small to counter balance the gravitational pressure. The electrons are absorbed by the protons, forming neutrons. This goes on until nuclear densities are reached. At that point the Fermi pressure of the neutrons stops the collapse. This happens so rapidly, that a shock wave develops which throws the outer layers of the star far off into the universe. This process is called supernova. The Crab nebula is supposedly the debris of such an explosion.
and the core is supposed to consist of a neutron star (of a star of $\approx 4M_\odot$). If the neutron star is larger than $3.4M_\odot$, it will undergo a further collapse and eventually form a black hole.

Apart from the evolution of an individual star, also the interaction between stars may increase the mass of a star beyond its critical mass. This mechanism is believed to be active in Centaurus $X-3$, which supposedly consists of a neutron star which pulls away matter from another star circling around it. Eventually, the neutron star will become too massive to it gravitational collapse.

It must be realized that to the outside world, the collapse is not at all a catastrophic event since as matter approaches the Schwarzschild radius, it moves slower and slower. The light emitted from this matter has longer and longer wavelengths.

When trying to compare these unusual phenomena with astronomical observation it must be kept in mind that the results have been derived purely from classical Einstein’s theory. There can be two possibilities why such a theory might fail. First, in physical theories it is usually the case that when gradients of field quantities become very large, higher gradient energies become necessary to describe their short wavelength phenomena. Einstein’s action is linear in the curvature and therefore involves only two derivatives. If the curvature becomes large it could well be that higher gradient terms must be included.

The other source is quantum physics. Before the Schrödinger equation the electron in an atom would undergo an electromagnetic collapse falling into the center while continuously radiating off electromagnetic waves. Quantum mechanics stopped this collapse via the Heisenberg uncertainty principle of indeterminacy. The analogous thing could happen with gravitational collapse once the quantum properties of the gravitational field will be better understood.

### 22.8 Collapsing Thin Shell

Consider a thin collapsing spherical shell of very light material. We ignore the angular degrees of freedom and focus attention only on the two-dimensional place of radial and time coordinates. In four spacetime dimensions, the centrifugal barrier prevents particle from crossing through $r = 0$. This is accounted for in 1+1 dimensions by postulating that particles move only in the $r > 0$ regime and are are reflected at $r = 0$ if they reach this point.

Inside the shell, the spacetime is Minkowskian and has the metric

$$ds^2 = c^2d\tau^2 - dr^2. \quad (22.146)$$

The collapsing shell has a time-dependent radius $R(\tau)$. If the collapse start at $\tau = 0$ with a radius $R_0 > r_S$, the shell moves inwards with approximately light velocity, so that the time dependence of $R(\tau)$ is:

$$R(\tau) = R_0 - \Theta(\tau) c\tau, \quad (22.147)$$
where $\Theta(\tau)$ is the Heaviside function defined by
\[
\Theta(\tau) \equiv \begin{cases} 
1 & \tau \geq 0, \\
0 & \tau < 0.
\end{cases}
\tag{22.148}
\]

Outside the shell, the metric has the Schwarzschild form (22.50):
\[
ds^2 = \left(1 - \frac{r_S}{r}\right)c^2 dt^2 - \left(1 - \frac{r_S}{r}\right)^{-1} dr^2.
\tag{22.149}
\]

The relation between the Schwarzschild time $t$ and the Minkowski time $\tau$ is found by equating the two equations for $(ds/cd\tau)^2$ on the shell, i.e. for $r = R(\tau)$. This yields
\[
\frac{dt}{d\tau} = \frac{1}{R(\tau) - r_S} \left\{ R(\tau) \left[ R(\tau) - r_S - \frac{R^2 r_S}{c^2} \right] \right\}^{1/2}.
\tag{22.150}
\]

Before the collapse, this is simply
\[
\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - r_S/R_0}}.
\tag{22.151}
\]
during the collapse it becomes
\[
\frac{dt}{d\tau} = \frac{R_0 - c\tau}{R_0 - c\tau - r_S}.
\tag{22.152}
\]

Let us also define the Regge-Wheeler *turtoise coordinate* [11, 8]
\[
r^* \equiv r + r_S \log(r/r_S - 1),
\tag{22.153}
\]
and introduce the *Eddington-Finkelstein coordinates* [12]
\[
u \equiv ct - r^*, \quad v \equiv ct + r^*,
\tag{22.154}
\]
to express the geodesics in the Schwarzschild metric outside the shell as $U =$const for outgoing and $V =$const for incoming particles. We also introduce light cone coordinates in the Minkowski spacetime inside the shell:
\[
U \equiv c\tau - r, \quad V \equiv c\tau + r.
\tag{22.155}
\]
In terms of these coordinates, the metric is
\[
ds^2 = \begin{cases} 
\frac{dU dV}{1 - r_S/r} & \text{dudv}, \\
\end{cases}
\tag{22.156}
\]

The collapse starts at $u = -R_0$ and $U = -R_0$. It will be useful to shift this point to $u = U = 0$, by using the translated coordinates
\[
u \equiv ct - r^* + R_0^*, \quad v \equiv ct + r^* - R_0^*,
\tag{22.157}
\]
\[
U \equiv c\tau - r + R_0, \quad V \equiv c\tau + r - R_0.
\tag{22.158}
\]
Notes and References

[1] For the scalar tensor theory see

[2] For the astronomical determination of the precession of the perihelion of Mercury see
G.M. Clemence, Rev. Mod. Phys. 19 361 (1947);


[4] The most accurate photocell scanning measurement of the solar asphericity was performed by
A. Hill, P.D. Clayton, D.Z. Patz, A.W. Healy, R.T. Stebbins, J.R. Oleson,

[5] The Kruskal coordinates we proposed by


[9] See the internet page http://www.sciencemag.org/cgi/content/full/sci;303/5661/1143.

[10] For the speed of sound argument of the critical mass see


Problems

22.1 Using the experimental difference between equatorial and polar arc width of the sun \( \Delta \odot = .016'' \) (seconds of arc) and the total width \( \delta_{2R_\odot} = 1920'' \), calculate the quadrupole moment of the sun, assuming a uniform mass distribution.
Solution

The quadrapole moment for a uniform mass density \( \rho = M_\odot/(4\pi R_\odot^3/3) \) can be written as

\[
Q_{33} = \int d^3x (z^2 - r^2/3) \rho(x) = \frac{2}{3} \left[ \frac{2\pi}{4\pi} \int_0^1 dr \int_0^{r(\theta)} r^4 P_2(\cos \theta) \right] \frac{M_\odot}{3 R_\odot^3}
\]

where \( r(\theta) = R_\odot[1 - \epsilon P_2(\cos \theta)] \) and \( \epsilon = (2/3)(\Delta_\odot/\delta_{2R_\odot}) \). The latter identification follows directly from \( P_2(\cos \theta) = (1/2)(3 \cos^2 \theta - 1) \) being 1 and \(-1/2\) at poles and equator, respectively. Hence

\[
Q_{33} = M_\odot R_\odot \frac{1}{3} \int_0^1 d\cos \theta P_2(\cos \theta) [1 - \epsilon P_2(\cos \theta)]^3 = -M_\odot R_\odot \frac{2}{3} \frac{2}{3},
\]

from which we find \( Q_{33} = -(2/3)M_\odot \cdot 2.7 \times 10^{-5} \).
Dirac Electron in Schwarzschild Metric

Let us study the behavior of electrons in a gravitational field of a star described by a Schwarzschild metric \(22.50\). In order to find the spin connection, we choose vierbein fields

\[
\begin{align*}
 h^0_0 &= B^{1/2}(r), \quad h^1_1 = B^{-1/2}(r), \quad h^2_2 = r, \quad h^3_3 = r \sin \theta, \\
 h^0_0 &= B^{-1/2}(r), \quad h^1_1 = B^{1/2}(r), \quad h^2_2 = \frac{1}{r}, \quad h^3_3 = \frac{1}{r \sin \theta}.
\end{align*}
\]

(23.1) (23.2)

It is obvious that these satisfy the basic relation \(17.28\).

In order to write down the Dirac equation in the Schwarzschild geometry we have to calculate the spin connection. This is done from Eqs. \(17.91\) and \(17.62\).

We take care of rotational symmetry of the system by splitting the spinor wave functions into radial and angular parts

\[
\psi(x, t) = \frac{B^{-1/4}(r)}{r} \begin{pmatrix} F(r, t) y^l_{j,m}(\theta, \phi) \\ iG(r, t) y^l_{j,m}(\theta, \phi) \end{pmatrix},
\]

(23.3)

where \(y^l_{j,m}(\theta, \phi)\) denote the spinor spherical harmonics (see Appendix 23C). They are composed from the ordinary spherical harmonics \(Y_{lm}(\theta, \phi)\) (see Appendix 23A) and the two-component Pauli spinors \(\chi(s_3)\), the eigenvectors of \(\sigma_3/2\), via Clebsch-Gordan coefficients (see Appendix 23B):

\[
y^l_{j,m}(\theta, \phi) = \langle j, m | l, m'; \frac{1}{2}, s_3 \rangle Y_{lm}(\theta, \phi) \chi(s_3).
\]

(23.4)

The explicit form of the spinor spherical harmonics \(23.4\) is for \(l = l_\pm\):

\[
y^{l_+}_{j,m}(\theta, \phi) = \frac{1}{\sqrt{2l_+ + 1}} \begin{pmatrix} \sqrt{l_+ - m + \frac{1}{2}} Y_{l_+ m - \frac{1}{2}}(\theta, \phi) \\ -\sqrt{l_+ + m + \frac{1}{2}} Y_{l_+ m + \frac{1}{2}}(\theta, \phi) \end{pmatrix},
\]

(23.5)

\[
y^{l_-}_{j,m}(\theta, \phi) = \frac{1}{\sqrt{2l_- + 1}} \begin{pmatrix} \sqrt{l_- + m + \frac{1}{2}} Y_{l_- m - \frac{1}{2}}(\theta, \phi) \\ \sqrt{l_- - m + \frac{1}{2}} Y_{l_- m + \frac{1}{2}}(\theta, \phi) \end{pmatrix}.
\]

(23.6)
On these eigenfunctions, the operator \( \mathbf{L} \cdot \mathbf{\sigma} \) has the eigenvalues

\[
\mathbf{L} \cdot \mathbf{\sigma} y_{j,m}^{l \pm} (\theta, \phi) = -\left(1 + \kappa_{\pm}\right)y_{j,m}^{l \pm} (\theta, \phi),
\]

with

\[
\kappa_{\pm} = \mp \left(j + \frac{1}{2}\right), \quad j = l \pm \frac{1}{2}.
\]

Searching for stationary states with a time dependence \( e^{-iEt/\hbar} \), we find the differential equations for the radial functions \( F(r) \) and \( G(r) \)

\[
\begin{pmatrix}
\partial_r F(r) \\
\partial_r G(r)
\end{pmatrix} = B^{-1/2}(r) \begin{pmatrix}
\kappa/r & m - B^{-1/2}(r)[E-V(r)] \\
m + B^{-1/2}(r)[E-V(r)] & -\kappa/r
\end{pmatrix} \begin{pmatrix}
F(r) \\
G(r)
\end{pmatrix},
\]

where

\[
B(r) = 1 - \frac{r_S}{r},
\]

with Schwarzschild radius \( r_S \) of Eq. (22.120).

These equations have been solved for a dense star [1] consisting of an incompressible fluid of density \( \rho_0 \) and total mass

\[
M = \frac{4\pi}{3} \rho_0 r_0^3,
\]

whose outside has the Schwarzschild metric (22.50) and whose inside has the metric (22.13) with

\[
B^{1/2}(r) = \frac{3}{2} \sqrt{1 - \frac{r_S}{r_0}} - \frac{1}{2} \sqrt{1 - \frac{r^2 r_S}{r_0^3}}, \quad A^{-1}(r) = 1 - \frac{r^2 r_S}{r_0^3}.
\]

Figure 23.1 Energy levels of Dirac particle in gravotational field of an incompressible star of radius \( r_0 \).
where \( r_S \) is the Schwarzschild radius (22.120). For meaningful wave functions, the radius of the star must satisfy

\[
r_0 > \frac{8}{9} r_S,
\]

(23.13)

where the pressure becomes infinite. Numerical integration of the Eqs. (23.9) yields the eigenvalues of the energy shown in Fig. 23.1.

**Appendix 23A  Spherical Harmonics**

The spherical harmonics are defined as

\[
Y_{lm}(\theta, \varphi) \equiv (-1)^m \left[ \frac{2l + 1}{4\pi} \frac{(l - m)!}{(l + m)!} \right]^{1/2} P_l^m(\cos \theta)e^{im\varphi},
\]

(23A.1)

where \( P_l^m(z) \) are the associated Legendre polynomials

\[
P_l^m(z) = \frac{1}{2^l l!} (1 - z^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (z^2 - 1)^l.
\]

(23A.2)

The spherical harmonics are orthonormal with respect to the rotation-invariant scalar product

\[
\int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\varphi \ Y_{lm}^*(\theta, \varphi)Y_{l'm'}(\theta, \varphi) = \delta_{ll'} \delta_{mm'}.
\]

(23A.3)

Explicitly, they read for the lowest few angular momenta:

\[
\begin{align*}
Y_{0,0}(\theta, \phi) & = \frac{1}{2\sqrt{\pi}}, \\
Y_{11}(\theta, \phi) & = -\sqrt{\frac{3}{4\pi}} \sin \theta \frac{1}{\sqrt{2}} e^{i\phi}, \\
Y_{10}(\theta, \phi) & = \sqrt{\frac{3}{4\pi}} \cos \theta, \\
Y_{22}(\theta, \phi) & = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi}, \\
Y_{21}(\theta, \phi) & = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}, \\
Y_{20}(\theta, \phi) & = \sqrt{\frac{5}{4\pi}} \left( \frac{3}{2} \cos^2 \theta - \frac{1}{2} \right), \\
Y_{33}(\theta, \phi) & = -\frac{1}{4} \sqrt{\frac{35}{4\pi}} \sin^3 \theta e^{3i\phi},
\end{align*}
\]
\[ Y_{32}(\theta, \phi) = \frac{1}{4} \sqrt{\frac{105}{2\pi}} \sin^2 \theta \cos \theta e^{2i\phi}, \]
\[ Y_{31}(\theta, \phi) = -\frac{1}{4} \sqrt{\frac{21}{4\pi}} \sin \theta \left(5\cos^2 \theta - 1\right) e^{i\phi}, \]
\[ Y_{30}(\theta, \phi) = \sqrt{\frac{7}{4\pi}} \left(\frac{5}{2}\cos^3 \theta - \frac{3}{2} \cos \theta\right). \] (23A.4)

The spherical harmonics with negative magnetic quantum number \(m\) are obtained via the relation
\[ Y_{lm}(\theta, \phi) = (-1)^m Y_{l,-m}^*(\theta, -\phi). \] (23A.5)

For \(m = 0\), the spherical harmonics reduce to
\[ Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos \theta), \] (23A.6)
where
\[ P_l(z) \equiv P_0^l(z) = \frac{1}{2^l l!} \frac{d^l}{dz^l} (z^2 - 1)^l \] (23A.7)
are the Legendre polynomials.

### Appendix 23B  Clebsch-Gordan Coefficients

A direct product of irreducible representation states \(|s_1,m_1\rangle\) and \(|s_2,m_2\rangle\) can be decomposed into a sum of irreducible representation states \(|s,m\rangle\) with total angular momentum \(s = |s_1 - s_2|, \ldots, (s_1 + s_2)\). This is done with the help of Clebsch-Gordan coefficients. For this we multiply any product state with the completeness relation of all irreducible representation states
\[ \sum_s \sum_{m=-s}^s |s,m\rangle \langle s,m| = 1, \] (23B.1)
and obtain
\[ |s_1,m_1; s_2,m_2\rangle = \sum_{s=|s_1-s_2|}^{s_1+s_2} \sum_{m=-s}^s |s,m\rangle \langle s,m|s_1,m_1; s_2,m_2\rangle. \] (23B.2)

The expansion coefficients on the right-hand side are the desired Clebsch-Gordan coefficients.

The expansion (23B.2) can be inverted by means of a similar completeness relation in the product space
\[ \sum_{m_1=-s_1}^{s_1} \sum_{m_2=-s_2}^{s_2} |s_1,m_1; s_2,m_2\rangle \langle s_1,m_1; s_2,m_2| = 1, \] (23B.3)
yielding the expansion

\[ |s, m\rangle = \sum_{m_1 = -s_1}^{s_1} \sum_{m_2 = -s_2}^{s_2} |s_1, m_1; s_2, m_2\rangle \langle s_1, m_1; s_2, m_2|s, m\rangle. \tag{23B.4} \]

Under a rotation, these transform as follows:

\[ D^s_{m,m'} \langle s, m'|s_1, m_1'; s_2, m_2\rangle (D^s)^{-1}_{m_1,m_1'} (D^s)^{-1}_{m_2,m_2} = \langle s, m|s_1, m_1; s_2, m_2\rangle, \tag{23B.5} \]

or, because of unitarity of the representation matrices,

\[ D^s_{m,m'} (D^s_{m_1,m_1'})^* (D^s_{m_2,m_2'})^* \langle s, m'|s_1, m_1'; s_2, m_2'\rangle = \langle s, m|s_1, m_1; s_2, m_2\rangle, \tag{23B.6} \]

and since the Clebsch-Gordan coefficients are real following the Condon-Shortley convention we also have

\[ (D^s_{m,m'})^* D^s_{m_1,m_1'} D^s_{m_2,m_2'} \langle s, m|s_1, m_1'; s_2, m_2'\rangle = \langle s, m|s_1, m_1; s_2, m_2\rangle. \tag{23B.7} \]

The Clebsch-Gordan coefficients are related in a simple way to the more symmetric Wigner 3j-symbols by defined as follows:

\[ \langle s_3, -m_3|s_1, m_1; s_2, m_2\rangle = (-1)^{s_1-s_2-m_3}(2s_3+1)^{1/2} \begin{pmatrix} s_1 & s_2 & s_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{23B.8} \]

This has the more symmetric invariance property

\[ D^s_{m_1,m_1'} D^s_{m_2,m_2'} D^s_{m_3,m_3'} \begin{pmatrix} s_1 & s_2 & s_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} s_1 & s_2 & s_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{23B.9} \]

The Levi-Civita symbol \( \epsilon_{ijk} \) is a cartesian version of the Wigner 3j-symbol for \( s_1 = s_2 = s_3 = 1 \). It exhibits the invariance (23B.9) with respect to the 3 \times 3 defining representation matrices of the rotation group:

\[ R_{i_1i'} R_{i_2i''} R_{i_3i'''} \epsilon_{i_1i''i'''} = \epsilon_{i_1i_2i_3}. \tag{23B.10} \]

Under even permutations of columns, the 3j-symbols are invariant while under odd permutations they pick up a phase factor \((-1)^{s_1+s_2+s_3}\). Also:

\[ \begin{pmatrix} s_1 & s_2 & s_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{s_1+s_2+s_3} \begin{pmatrix} s_1 & s_2 & s_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix}. \tag{23B.11} \]

In order to calculate the Clebsch-Gordan coefficients we observe that the state of maximal quantum numbers \(|s_1, s_1; s_2, s_2\rangle\) is a state \(|s, m\rangle\) of the irreducible representation with the maximal angular momentum \( s = m = s_1 + s_2 \). By repeatedly
Table 23.1 The lowest Clebsch-Gordan coefficients \( \langle s_1, m_1; s_2, m_2 | s, m \rangle \). The table entries \( \pm CG \) are to be read as \( \pm \sqrt{CG} \). The coefficients are all real, i.e. equal to \( \langle s_1, m_1; s_2, m_2 | s, m \rangle \). For more symmetry properties see Eqs. (23B.22).

\[
\langle s_1, m_1; s_2, m_2 | s, m \rangle = (-1)^{s-s_1-s_2} \langle s_2, -m_2; s_1, -m_1 | s, -m \rangle = (-1)^{s-s_1-s_2} \langle s_2, m_2; s_1, m_1 | s, m \rangle
\]
applying the lowering operator of angular momentum to it according to the well-known rules

\[ \hat{L}_3 |s, m\rangle = m |s, m\rangle, \]  
\[ \hat{L}_+ |s, m\rangle = \sqrt{(s - m)(s + m + 1)} |s, m + 1\rangle, \]  
\[ \hat{L}_- |s, m\rangle = \sqrt{(s + m)(s - m + 1)} |s, m - 1\rangle, \]

we obtain the matrix elements (compare Fig. 23.2)

\[ L_- |s, s\rangle = \sqrt{(2s + 1)} |s, s - 1\rangle, \]
\[ L_- |s, s - 1\rangle = \sqrt{(2s - 1)} |2s, s - 1\rangle, \]
\[ L_- |s, -s + 2\rangle = \sqrt{(2s - 1)} |2s, -s + 1\rangle, \]
\[ L_- |s, -s + 1\rangle = \sqrt{(2s)} |s, -s\rangle. \]  

In the direct-product space, an application of the lowering operator \( L_- \times 1 + 1 \times L_- \) to the state \(|s_1, s_1; s_2, s_2\rangle\) yields, with the same rules as in (23B.15),

\[ (L_- \times 1 + 1 \times L_-)|s_1, s_1; s_2, s_2\rangle = \sqrt{(2s_1)} |s_1, s_1 - 1; s_2, s_2\rangle + \sqrt{(2s_2)} |s_1, s_1; s_2, s_2 - 1\rangle. \]  

Continuing this with the help of the general relation

\[ (L_- \times 1 + 1 \times L_-)|s_1, m_1; s_2, m_2\rangle = \sqrt{(s_1 + m_1)(s_1 - m_1 + 1)} |s_1, m_1 - 1; s_2, m_2\rangle + \sqrt{(s_2 + m_2)(s_2 - m_2 + 1)} |s_1, m_1; s_2, m_2 - 1\rangle, \]

we find all other states \(|s, m\rangle\) of the irreducible representation with \( s = s_1 + s_2\).
The state of the lower total angular momentum \( s_1 + s_2 - 1 \) with maximal magnetic quantum number \( m = s \) is obtained from the orthogonal combination of (23B.17):

\[
|s_1 + s_2 - 1, s_1 + s_2 - 1 \rangle = \sqrt{(2s_1 + 1)}|s_1, s_1 - 1; s_2, s_2 \rangle - \sqrt{(2s_2 + 1)}|s_1 s_1; s_2 s_2 - 1 \rangle.
\]  
(23B.18)

This can be verified by applying to it the raising operator \((L_+ \times 1 + 1 \times L_+)\); generalizing (23B.14) to the direct-product space:

\[
(L_+ \times 1 + 1 \times L_+) |s_1, m_1; s_2, m_2 \rangle = \sqrt{(s_1 - m_1)(s_1 + m_1 + 1)}|s_1, m_1 + 1; s_2, m_2 \rangle + \sqrt{(s_2 - m_2)(s_2 + m_2 + 1)}|s_1, m_1; s_2, m_2 + 1 \rangle,
\]  
(23B.19)

and finding that it is annihilated.

By applying the lowering operator to the state (23B.18), we generate all states of the irreducible representation \(|s_1 + s_2 - 1, m \rangle\) with \( m = -s_1 - s_2 + 1, \ldots, s_1 + s_2 - 1 \).

Multiplying (23B.17) from the left by \( \langle s, m \rangle \) and using the hermitian adjoint of relation (23B.14), we obtain the recursion relation

\[
\sqrt{(s + m)(s - m + 1)}|s_1, m_1; s_2, m_2 \rangle = \sqrt{(s_1 - m_1 + 1)(s_1 + m_1)}|s_1, m_1 - 1; s_2, m_2 \rangle|s, m - 1 \rangle + \sqrt{(s_2 - m_2 + 1)(s_2 + m_2)}|s_1, m_1; s_2, m_2 - 1 \rangle|s, m - 1 \rangle.
\]  
(23B.20)

Similarly we can multiply the raising operator relation (23B.19) in the direct-product space multiply it by \( \langle s, m | \) from the left and the adjoint of (23B.13) to find

\[
\sqrt{(s - m)(s + m + 1)}|s_1, m_1; s_2, m_2 \rangle = \sqrt{(s_1 + m_1 + 1)(s_1 - m_1)}|s_1, m_1 + 1; s_2, m_2 \rangle|s, m + 1 \rangle + \sqrt{(s_2 + m_2 + 1)(s_2 - m_2)}|s_1, m_1; s_2, m_2 + 1 \rangle|s, m + 1 \rangle.
\]  
(23B.21)

The Clebsch-Gordan coefficients have the following important symmetry properties:

\[
\langle s, m | s_1, m_1; s_2, m_2 \rangle = (-1)^{j_1 - s_1 - s_2} \langle s, m | s_2, m_2; s_1, m_1 \rangle = \langle s, m | s_2, m_2; s_1, m_1 \rangle
\]

\[
= (-1)^{j_1 - s_1 - s_2} \langle s, m | s_1, m_1; s_2, m_2 \rangle
\]

\[
= (-1)^{j_1 - s_1 - s_2} \langle s, -m | s_1, -m_1; s_2, -m_2 \rangle
\]  
(23B.22)

\[
= (-1)^{s_1 - m_1} \sqrt{\frac{2s + 1}{2s_2 + 1}} \langle s_2, -m_2 | s_1, m_1; s, -m \rangle
\]

\[
= (-1)^{s_2 + m_2} \sqrt{\frac{2s + 1}{2s_1 + 1}} \langle s_1, -m_1 | s, -m; s_2, -m_2 \rangle.
\]

Some frequently-needed values are listed in Table 23.1.
Appendix 23C  Spinor Spherical Harmonics

Equation (23.4) defines spinor spherical harmonics. In these, an orbital wave function of angular momentum $l_{\pm}$ is coupled with spin $1/2$ to a total angular momentum $j = l_{\pm} \pm 1/2$. For the configurations $j = l_{\pm} + 1/2$ with $m_{2} = -1/2$ the recursion relation (23B.20) for the Clebsch-Gordan coefficients $\langle s_{1}m_{1}; s_{2}m_{2}| sm \rangle$, simplifies by having no second term. Inserting $s_{1} = l_{-}$, $s_{2} = 1/2$ and $s = j = l_{-} + 1/2$, we find

$$\langle l_{-}, m + \frac{1}{2}; \frac{1}{2}; -\frac{1}{2}| l_{-} + \frac{1}{2}, m \rangle = \sqrt{\frac{l_{-} - m + 1/2}{l_{-} - m + 3/2}} \langle l_{-}, m - \frac{1}{2}; \frac{1}{2}; -\frac{1}{2}| l_{-} + \frac{1}{2}, m - 1 \rangle.$$ (23C.1)

This has to be iterated with the initial condition

$$\langle l_{-}, -l_{-}; \frac{1}{2}, -\frac{1}{2}| l_{-} + \frac{1}{2}, -l_{-} - \frac{1}{2} \rangle = 1,$$ (23C.2)

which follows from the fact that the state $\langle l_{-}, -l_{-}; \frac{1}{2}, -\frac{1}{2} \rangle$ carries a unique magnetic quantum number $m = -l_{-} - 1/2$ of the irreducible representation of total angular momentum $s = j = l_{-} + 1/2$. The result of the iteration is

$$\langle l_{+}, m - \frac{1}{2}; \frac{1}{2}; \frac{1}{2}| l_{+} - \frac{1}{2}, m \rangle = \sqrt{\frac{l_{+} - m + 1/2}{2l_{+} + 1}}.$$ (23C.3)

Similarly we may simplify the recursion relation (23B.21) for the configurations $j = l_{+} - 1/2$ with $m_{2} = 1/2$ to

$$\langle l_{-}, m - \frac{1}{2}; \frac{1}{2}; \frac{1}{2}| l_{-} + \frac{1}{2}, m \rangle = \sqrt{\frac{l_{-} + m + 1/2}{l_{-} + m + 3/2}} \langle l_{-}, m + \frac{1}{2}; \frac{1}{2}; \frac{1}{2}| l_{-} + \frac{1}{2}, m + 1 \rangle,$$ (23C.4)

and iterating this with the initial condition

$$\langle l_{-}, l_{-}; \frac{1}{2}m_{2}| l_{-} + \frac{1}{2}, l_{-} + \frac{1}{2} \rangle = 1,$$ (23C.5)

which expresses the fact that the state $\langle l_{-}, l_{-}; \frac{1}{2}m_{2} \rangle$ is the state of maximal magnetic quantum number $m = l_{-} + 1/2$ of the irreducible representation of total angular momentum $s = j = l_{-} + 1/2$. The result of the iteration is

$$\langle l_{-}, m - \frac{1}{2}; \frac{1}{2}; \frac{1}{2}| l_{+} + \frac{1}{2}, m \rangle = \sqrt{\frac{l_{+} + m + 1/2}{2l_{-} + 1}}.$$ (23C.6)

Inserting (23C.3) and (23C.6) the expression (23.4) for the spinor spherical harmonic of total angular momentum $j = l_{-} + 1/2$, which now reads

$$y_{j,m}^{l_{-}}(\theta, \phi) = \langle l_{-}, m - \frac{1}{2}; \frac{1}{2}; \frac{1}{2}| l_{-} + \frac{1}{2}, m \rangle \ Y_{l_{-}m-1/2}(\theta, \phi) \chi(\frac{1}{2})$$
$$+ \langle l_{-} m + \frac{1}{2}; \frac{1}{2}; -\frac{1}{2}| l_{-} + \frac{1}{2}, m \rangle \ Y_{l_{-}m+1/2}(\theta, \phi) \chi(-\frac{1}{2}),$$ (23C.7)
and separating the spin-up and spin-down components, we obtain precisely (23.6).

In order to find corresponding result for $j = l_+ - 1/2$, we use the orthogonality relation for states with the same $l$ but different $j = l \pm 1/2$:

$$\langle l + \frac{1}{2}, m | l - \frac{1}{2}, m \rangle = 0.$$  \hfill (23C.8)

Inserting a complete set of states in the direct product space yields

$$\langle l + \frac{1}{2}, m | l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \rangle \langle m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | l - \frac{1}{2}, m \rangle + \langle l + \frac{1}{2}, m | l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle \langle m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | l - \frac{1}{2}, m \rangle = 0.$$  \hfill (23C.9)

Together with (23C.3) and (23C.6) we find

$$\langle l_+ m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | l_+ - \frac{1}{2}, m \rangle = \sqrt{\frac{l_+ + m + 1/2}{2l_+ + 1}},$$

$$\langle l_+ m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | l_+ - \frac{1}{2}, m \rangle = -\sqrt{\frac{l_+ - m + 1/2}{2l_+ + 1}}.$$  \hfill (23C.10)

With this, the expression (23.4) for the spinor spherical harmonics written as

$$y_{l_+}^{m,\frac{1}{2}}(\theta, \phi) = \langle l_+ m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | l_+ - \frac{1}{2}, m \rangle Y_{l_+ m - 1/2}(\theta, \phi) \chi(\frac{1}{2})$$

$$+ \langle l_+ m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | l_+ - \frac{1}{2}, m \rangle Y_{l_+ m + 1/2}(\theta, \phi) \chi(\frac{-1}{2})$$  \hfill (23C.11)

has the components given in (23.5).

**Notes and References**

Gravitational Waves

24.1 Introductory Remarks

Changes in the mass distribution lead to changes of the gravitational field. Since the adjustment to a new field configuration can propagate at most with the speed of light, the universe must be filled with gravitational waves. The collapse of stars, supernovae explosion, birth of neutron stars, and similar dramatic events in the universe must all be accompanied by burst of such waves whose general properties will now be studied.

24.2 Reminder of Electromagnetic Waves

Before entering the discussions it is useful to recall the analogous situation for electromagnetic waves. They are carried by the four-vector potential which satisfies the field equations:

\[-\partial^2 A^a + \partial^a \partial_b A^b = -\frac{1}{c} j^a.\] (24.1)

In order to solve these we have to choose a specific gauge, say the Lorentz gauge

\[\partial_a A^a(x) = 0,\] (24.2)

in which case the equation reduces to

\[\partial^2 A^a = \frac{1}{c} j^a.\] (24.3)

The solution can directly be written down with the help of the retarded Coulomb Green function:

\[A^a(x) = \frac{1}{4\pi c} \int d^3 x' \frac{1}{|x - x'|} j^a(x')|_{t' = t - |x' - x|/c}.\] (24.4)

The retarded time argument in the source accounts for the difference between the time of emission and the time of arrival at the point \(x\) due to the finite light velocity.
A charge point moving along the trajectory \( \bar{x}(\tau) \) gives rise to a current

\[
j^\mu(x) = ec \int ds \dot{x}^\mu(s) \delta^{(4)}(x - \bar{x}(s)) = ec \frac{\dot{x}^\mu(\tau)}{d\bar{x}^0/ds} \delta^{(3)}(x - \bar{x}(s)). \tag{24.5}\]

Here \( s \) is the invariant length parameter, such that

\[
d\bar{x}^0/s = \sqrt{1 - \bar{v}^2(t)/c^2} = \gamma(t), \tag{24.6}\]

where \( \bar{v}(t) \) is the velocity along the trajectory \( \bar{x}(t) \):

\[
\bar{v}(t) \equiv \dot{\bar{x}}(t). \tag{24.7}\]

The current components are

\[
j^0(x) = ec\delta^{(3)}(x - \bar{x}(t)), \quad j(x) = e\bar{v}(t)\delta^{(3)}(x - \bar{x}(t)). \tag{24.8}\]

To find the electromagnetic field emerging from this current we solve the field equation (24.3) by

\[
A^\mu(x) = \frac{i}{c} \int d^4y G_R(x - y) j^\mu(y), \tag{24.9}\]

where \( G_R(x - y) \) is the retarded Green function

\[
G_R(x - x') = -i\Theta(x^0 - x'^0) \frac{1}{4\pi R} \left[ \delta(x^0 - x'^0 - R) - \delta(-x^0 + x'^0 + R) \right]. \tag{24.10}\]

The result is the Liénard-Wiechert potential:

\[
A^0(x) = ec \int dt' \Theta(t - t') \frac{1}{4\pi R} \bar{v}(t') \delta(t - t' - R(t')), \tag{24.11}\]

\[
A(x) = e \int dt' \Theta(t - t') \frac{1}{4\pi R} \bar{v}(t') \delta(t - t' - R(t')). \tag{24.12}\]

We now simplify the \( \delta \)-functions as follows:

\[
\delta(t - t' - R(t')) = \frac{1}{|d[t' + R(t')] / dt'|_{t'=t_R}} \delta(t' - t_R) = \frac{1}{1 - n(t_R) \cdot \bar{v}(t_R)} \delta(t' - t_R), \tag{24.13}\]

where \( n(t) \) is the direction of the distance vector \( \mathbf{R}(t) \equiv x - \bar{x}(t) \) rom the charge:

\[
n(t) \equiv \frac{\mathbf{R}(t)}{|\mathbf{R}(t)|}. \tag{24.14}\]

Inserting this into (24.11) and (24.12), we find the vector potential

\[
\bar{v}(t) \equiv \dot{\bar{x}}(t). \tag{24.15}\]
\[
A^0(x) = \frac{1}{4\pi} \left[ \frac{e}{(1 - \mathbf{n} \cdot \mathbf{\hat{v}} / c) R} \right]_{\text{ret}}, \quad A(x) = \frac{1}{4\pi} \left[ \frac{e\mathbf{\hat{v}} / c}{(1 - \mathbf{n} \cdot \mathbf{\hat{v}} / c) R} \right]_{\text{ret}}.
\] (24.16)

The brackets with the subscript "ret" indicate that the time argument is retarded with respect to the time on the left-hand side of the equation:
\[
t \to t_R \equiv t - R(t_R)/c.
\] (24.17)

By forming the combinations
\[
E(x) = -\frac{1}{c} \dot{A}(x) - \nabla A^0(x), \quad B(x) = \nabla \times A(x),
\] (24.18) (24.19)

from the Liénard-Wiechert potential (24.16), we find the electromagnetic fields
\[
E(x, t) = e \frac{1}{4\pi} \left[ \frac{(\mathbf{n} - \mathbf{\hat{v}} / c)}{(1 - \mathbf{n} \cdot \mathbf{\hat{v}} / c)^2} \frac{1}{R^2} + \frac{\mathbf{n} \times [(\mathbf{n} - \mathbf{\hat{v}} / c) \times \dot{\mathbf{\hat{v}}}]}{(1 - \mathbf{n} \cdot \mathbf{\hat{v}} / c)^3} \right] R_{\text{ret}} \quad \text{(24.20)}
\]
\[
B(x, t) = [\mathbf{n} \times E]_{\text{ret}}.
\] (24.21)

The two terms in \(E(x, t)\) have different falloff behaviors in \(1/R\) with increasing distance from the source. The first is a velocity field which falls off like \(1/R^2\). It is essentially the moving static field around the particle. The second term is an acceleration field, which has a slower fall-off proportional to \(1/R\). For this reason it can carry off radiation energy to infinity.

Indeed, the energy flux through a solid angle \(d\Omega\) is given by the scalar product of the Poynting vector \(E \times B\) with the area element \(dS\):
\[
\dot{E} = dS \cdot (E \times B) = d\Omega c E^2.
\] (24.22)

There should be no confusion between the radiated energy \(E\) and the electric field \(E\).

For a radiating electron at small velocities near the coordinate origin, the acceleration field simplifies to
\[
E(x, t) = \frac{e}{4\pi r c^2} [\dot{x} \times (\dot{x} \times \ddot{x})], \quad B(x, t) = \frac{e}{4\pi r c^2} (\dot{x} \times \ddot{x}).
\] (24.23)

with \(r = |x|\). The radiated power per solid angle is then
\[
\frac{d\dot{E}}{d\Omega} = \frac{e^2}{(4\pi)^2 c^3} (\dot{x} \times \ddot{x})^2 = \frac{e^2}{(4\pi)^2 c^3} \dot{x}^2 \sin^2 \beta,
\] (24.24)

where \(\beta\) is the angle between the oscillating dipole end the direction of emission. By integrating over all solid angles, we obtain the total radiated power
\[
\dot{E} = \frac{e^2}{(4\pi)^2 c^3} \int d\Omega (\dot{x} \times \ddot{x})^2 = \frac{e^2}{4\pi c^3} \dot{x}^2.
\] (24.25)
24.2 Reminder of Electromagnetic Waves

This is the famous Larmor formula of classical electrodynamics.

For a harmonically oscillating charge at position

\[ x(t) = x_0 e^{-i\omega t} + x^*_0 e^{i\omega t} = 2|x_0| \cos(\omega t + \delta), \]  

(24.26)
equation (24.24) yields the temporal average power

\[ \frac{d\dot{E}}{d\Omega} = \frac{e^2 \omega^4}{8\pi^2} \frac{\omega^4}{c^4} |\hat{x} \times x_0|^2 = \frac{e^2 \omega^4}{8\pi^2} |x_0|^2 \sin^2 \beta, \]  

(24.27)
and the total radiated power is

\[ \dot{E} = \frac{e^2}{4\pi} \frac{4 \omega^4}{3 c^4} |x_0|^2. \]  

(24.28)

Note that with respect to standard classical electrodynamics textbooks where the electromagnetic Lagrangian carries a prefactor $1/4\pi$, the square of the charge has an extra factor $1/4\pi$, and $e^2$ is related to the fine-structure constant $\alpha$ by $e^2 = 4\pi \alpha \bar{\hbar} c$.

A more general radiation formula valid for any charge configuration is obtained by rewriting the vector potential (24.4) as

\[ A^a(x) = \frac{1}{4\pi c} \int d^3 x' \frac{1}{|x - x'|} \int \frac{d\omega}{2\pi} j^a(\mathbf{x}', \omega) e^{-i\omega t + i\omega |x - x'|/c}, \]  

(24.29)
where

\[ j^a(\mathbf{x}', \omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} j^a(\mathbf{x}', t) \]  

(24.30)
are the Fourier components of the energy-momentum tensor. Approximating for large $r = |\mathbf{x}|$:

\[ \frac{e^{i\omega |\mathbf{x} - \mathbf{x}'|/c}}{|\mathbf{x} - \mathbf{x}'|} \approx \frac{e^{i\omega r/c}}{r} e^{-i\omega \mathbf{n} \cdot \mathbf{x}'/c} \]  

(24.31)
leaving an $\mathbf{x}'$-dependence only in the sensitive phase factor [recall the exact splitting (24.20) into a velocity and an acceleration field which carries off energy].

At a point $\mathbf{x}$ far away from the source, the spherically radiated field (24.31) looks like a passing plane wave with $e^{i\omega r/c} \approx e^{i\mathbf{k} \cdot \mathbf{x}}$. We therefore set $\omega \mathbf{n} \cdot \mathbf{x}' = \mathbf{k} \cdot \mathbf{x}'$, and find

\[ A^a(x, t) = \frac{1}{4\pi c} \frac{e^{i\omega r/c}}{r} \int \frac{d\omega}{2\pi} e^{-i\omega t} j^a(\mathbf{k}, \omega), \]  

(24.32)
where $j^a(\mathbf{k}, \omega)$ is the Fourier transform in spacetime:

\[ j^a(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} dt \int d^3 \mathbf{x} e^{i\omega t - i\mathbf{k} \cdot \mathbf{x}} j^a(\mathbf{x}, t). \]  

(24.33)
We can also write

\[ A^a(x, t) = \frac{1}{4\pi c} \frac{e^{i\omega r/c}}{r} j^a(\mathbf{k}, t). \]  

(24.34)
We now calculate the energy flux from formula (24.22). In terms of the vector potential, this reads

$$\frac{d\dot{E}}{d\Omega} = c\left[\frac{1}{c}\dot{\mathbf{A}} + \nabla A^0\right]^2. \quad (24.35)$$

In momentum space, the Lorentz gauge (24.2) implies

$$A^0(k,t) = \hat{k} \cdot \mathbf{A}(k,t), \quad (24.36)$$

such that we can rewrite

$$\frac{1}{c}\dot{A}(k,t) + i k A^0(k,t) = \frac{1}{c}\partial_t \left[\mathbf{A}(k,t) - \hat{k} A^0(k,t)\right] = \frac{1}{c}\Lambda(\hat{k})\partial_t \mathbf{A}(k,t), \quad (24.37)$$

where $\Lambda(\hat{k})$ is a matrix with the elements

$$\Lambda_{ij}(\hat{k}) = \delta_{ij} - \hat{k}_i \hat{k}_j. \quad (24.38)$$

It is a projection matrix:

$$\Lambda_{ij}(\hat{k})\Lambda_{kl}(\hat{k}) = \Lambda_{il}(\hat{k}). \quad (24.39)$$

At a point $\mathbf{x}$ far away from the source, the spherically radiated field (24.205) looks like a passing plane wave with $e^{i\omega r/c} \approx e^{ikx}$. Formula (24.22) yields therefore the radiated energy per unit angle and time

$$\frac{dE}{dt d\Omega} = \frac{1}{16\pi^2 c^3} \Lambda_{ij}(\hat{k})[\partial_t j^i(k,t)]^* [\partial_t j^j(k,t)]. \quad (24.40)$$

Using current conservation in momentum space:

$$j^i(k,t)(\delta_{lm} - \hat{k}_i \hat{k}_m) j^m(k,t) = j^2(k,t) - j^0(k,t) j_a(k,t), \quad (24.41)$$

this can also be written in a fully covariant form as

$$\frac{dE}{dt d\Omega} = \frac{1}{16\pi^2 c^3} |\partial_t j^a(k,t)|^2. \quad (24.42)$$

For long wavelengths, the spatial components of current density have a negligible dependence on $\mathbf{k}$: $j^i(k,t) \approx j^i(\mathbf{k} = 0, t) = \int d^3x j^i(\mathbf{x}, t)$. Then we can perform the integral over all angles in (24.40) using the angular average

$$\langle \hat{k}_i \hat{k}_j \rangle = \frac{1}{3}\delta_{ij}, \quad \langle \Lambda_{ij}(\hat{k}) \rangle = \frac{2}{3}\delta_{ij}, \quad (24.43)$$

yielding

$$\frac{dE}{dt} = \frac{1}{4\pi c^3} \frac{2}{3} \left[\partial_t \int d^3x j(\mathbf{x}, t)\right]^2. \quad (24.44)$$
For long wavelengths, we may use a partial integration and current conservation to rewrite the dipole approximation

\[
\int d^3x \, j^i(x, t) = - \int d^3x \, x^i \partial_\mu j^\mu(x, t) = - \int d^3x \, x^i \partial_0 j^0(x, t) = - \dot{d}(t),
\]  

(24.45)

where

\[
\dot{d}(t) \equiv \int d^3x \, x^i j^0(x, t)
\]

(24.46)
is the dipole moment of the charge distribution. Then (24.44) becomes

\[
\frac{dE}{dt} = \frac{2}{3} \frac{1}{c^3} \frac{1}{4\pi} |\dot{d}(t)|^2.
\]  

(24.47)

For a single nonrelativistic point particle moving along the orbit \(\bar{x}(t)\), the spatial current density is

\[
j(x, t) = e \dot{\bar{x}}(t) \delta^{(3)}(x - \bar{x}(t)).
\]

(24.48)

Then Eq. (24.44) becomes

\[
\frac{dE}{dt} = \frac{2}{3} \frac{e^2}{c^3} \frac{1}{4\pi} |\ddot{\bar{x}}(t)|^2,
\]

(24.49)
in agreement with (24.25).

The projection matrix (24.38) in formula (24.40) is a consequence of the polarization properties of the radiated field.

Far away from the source, the field equation (24.3) becomes homogenous:

\[
\partial^2 A^a(x) = 0,
\]

(24.50)

so that the field can be expanded into free waves

\[
e^{-ikx} = e^{-i(k_0 x_0 - kx)} \quad \text{with} \quad k_0 = |k|
\]

(24.51)
as follows

\[
A^a(x) = \sum_{k,h} \left[ a(k, h) \epsilon^a(k, h) e^{-ikx} + \text{c.c.} \right]
\]

(24.52)

where \(\epsilon^a(k, h)\) are called polarization vector. For a standard direction of the momentum, say \(k\) along the z-direction,

\[
k^\mu = (|k|, k),
\]

(24.53)

we may choose

\[
\epsilon^a(k, 0) = (1, 0, 0, 0),
\]

\[
\epsilon^a(k, 1) = (0, 1, 0, 0),
\]

\[
\epsilon^a(k, 2) = (0, 0, 1, 0),
\]

\[
\epsilon^a(k, 3) = (0, 0, 0, 1),
\]

(24.54)
whose scalar products satisfy
\[ \epsilon^a(k, h)\epsilon_a(k, h') = g_{hh'} \]  \hspace{1cm} (24.55)

Not all of those vectors are admissible, however. First, there is the Lorentz condition. In momentum space, it reads
\[ \sum_h a(k, h)k_a \epsilon^a(kh) = 0. \]  \hspace{1cm} (24.56)

Since \( k \) points in the \( z \)-direction, this implies that \( a(k, 1) \) and \( a(k, 2) \) are unconstrained while the amplitudes \( a(k, 0) \) and \( a(k, 3) \) satisfy
\[ a(k, 0) = a(k, 3). \]  \hspace{1cm} (24.57)

Alternatively, we may introduce from the outset the polarization vectors \( \epsilon^a(k, 1), \epsilon^a(k, 2), \epsilon^a(k, l) \),
\[ \epsilon^a(k, l) = \frac{1}{\sqrt{2}}(1, 0, 0, 1), \quad \epsilon(k, l') = \frac{1}{\sqrt{2}}(1, 0, 0, 1). \]  \hspace{1cm} (24.58)

The first three automatically satisfy the Lorentz condition. The fourth polarization condition vector is forbidden by the Lorentz condition. The label \( l \) stands for “longitudinal”, indicating that this vector is parallel to the momentum vector \( (24.53) \). For a wave propagating along any direction, the polarization vectors \( \epsilon^a(k, 1), \epsilon^a(k, 2), \epsilon^a(k, l) \) must be rotated into this direction by a rotation matrix \( \Lambda(R)^a_b \epsilon^b \). The results are denoted by
\[ \epsilon^a(k, p) = \Lambda(R)^a_b \epsilon^b(k, p), \quad p = 1, 2, l, l'. \]  \hspace{1cm} (24.59)

We can now expand the general solution of the wave equation \( \partial^2 A^a = 0 \) in the Lorentz gauge \( \partial_a A^a = 0 \) as follows
\[ A^a(x) = \sum_{k, h=1, 2, l} \left[ a(k, h)\epsilon^a(k, h)e^{-ikx} + c.c. \right]. \]  \hspace{1cm} (24.60)

The polarization state \( l' \) is eliminated by the Lorentz condition. This however, is not all. The longitudinal component \( l \) is also unphysical. It contributes neither to the energy nor to the momentum of the field, nor does it couple to external currents. The reason is that \( \epsilon^a(k, l) \) is proportional to \( k^a \). Such a polarization does not contribute to the gauge-invariant field strength \( F^{ab} = \partial^a A^b - \partial^b A^a \) where it appears, in momentum space, in the combination \( k^a k^b - k^b k^a = 0 \). It does not couple to currents \( j^a(x) \), since these are conserved, satisfying the momentum space relation \( k^a j_a = 0 \). Hence the longitudinal component can be dropped from the expansion leaving only two physical degrees of polarization \( \epsilon^a(k, 1) \) and \( \epsilon^a(k, 2) \). These correspond to the \( A^i(x) \)-field oscillating along the two directions transverse to the spatial momentum vector. Such waves are called \textit{linearly polarized}.

H. Kleinert, \textit{GRAVITY WITH TORSION}
Alternatively, one may define the complex polarization vectors

\[ \epsilon^a(k, \pm) = \mp \frac{1}{\sqrt{2}} [e^a(k, 1) \pm ie^a(k, 2)], \tag{24.61} \]

called \textit{helicity} vectors. The projection matrix (24.38) arises from the completeness relation of the physical polarization states

\[ \Lambda_{ij}(k) = \epsilon^i_{\pm}(k, \pm) \epsilon^j_{\pm}(k, \pm)^* \delta_{ij} - \hat{k}_i \hat{k}_j \tag{24.62} \]

Hence we can rewrite (24.40) also as

\[ \frac{dE}{dtd\Omega} = \frac{1}{16\pi^2 c^3} \sum_{h=\pm} \left| \partial_t \epsilon(\hat{k}, h)j(k, t) \right|^2. \tag{24.63} \]

For checking the dimensions of the above equations the following list of dimensions is useful:

\[ [A] = \hbar, \quad [j^2] = \frac{\hbar}{\text{sec}^3 \text{cm}}, \quad [E] = \frac{\hbar}{\text{sec}}, \quad [p] = \frac{\hbar}{\text{cm}}, \quad [A^2_a] = \frac{\hbar}{\text{sec} \text{cm}} \tag{24.64} \]

### 24.3 Polarization States of Gravitational Waves

After this resumé we are ready to analyze the possible polarization states in the gravitational case. We have seen in Eq. (20.44) that the field equations are simplest written down in terms of the field \( \phi^{ab} = h^{ab} - \frac{1}{2} \eta^{ab} h \) in the Hilbert gauge

\[ \partial_a \phi^{ab} = 0 \tag{24.65} \]

In this gauge, the gravitational field equation reads, in close analogy with the electromagnetic equation (24.3),

\[ -\frac{1}{2} \partial^2 \phi^{ab}(x) = \kappa \frac{m}{T^{ab}(x)} \tag{24.66} \]

The solution of the equation involves again the retarded Coulomb Green function:

\[ \phi^{ab}(x) = \frac{2\kappa}{4\pi} \int d^3x' \frac{1}{|x - x'|} T^{ab}(x') \left| t' = -(x' - x)/c \right. \tag{24.67} \]

At a large distance from the source, the field equations become homogeneous:

\[ \partial^2 \phi^{ab} = 0. \tag{24.68} \]

The solutions can be expanded into plane waves \( e^{-ikx} \) with \( k_0 = |k| \) as in the electromagnetic case (24.52):

\[ \phi^{ab}(x) = \sum_{k, \lambda} \left[ a(k, \lambda) \epsilon^{ab}(k, \lambda) e^{-ikx} + \text{c.c.} \right] \tag{24.69} \]
where \( \epsilon^{ab}(k, \lambda) \) are polarization tensors and \( \lambda \) runs through the different polarization degrees of freedom. Initially, a symmetric tensor \( \epsilon^{ab}(k, \lambda) \) allows for 10 independent polarization states. The gauge condition (24.65) reads for these tensors

\[
k_a \epsilon^{ab}(k, \lambda) = 0,
\]

and eliminates four states. The remaining six components can be spanned with the help of symmetrized tensor products of the polarization vectors of the electromagnetic field. Dropping the polarization state \( l' \) eliminated by the Hilbert gauge (24.65), we only have to take the three polarizations 1, 2, \( l \) into account. From these we form \( 3 \cdot 4/2 = 6 \) independent symmetric tensor products:

\[
\begin{align*}
\epsilon^{ab}(k, 1, 1) &= \epsilon^a(k, 1) \epsilon^b(k, 1), \\
\epsilon^{ab}(k, 1, 2) &= \frac{1}{\sqrt{2}} \left[ \epsilon^a(k, 1) \epsilon^b(k, 2) + (a \leftrightarrow b) \right], \\
\epsilon^{ab}(k, 2, 2) &= \epsilon^a(k, 2) \epsilon^b(k, 2), \\
\epsilon^{ab}(k, l, l) &= \epsilon^a(k, l) \epsilon^b(k, l), \\
\epsilon^{ab}(k, 1, l) &= \frac{1}{\sqrt{2}} \left[ \epsilon^a(k, 1) \epsilon^b(k, l) + (a \leftrightarrow b) \right], \\
\epsilon^{ab}(k, 2, l) &= \frac{1}{\sqrt{2}} \left[ \epsilon^a(k, 2) \epsilon^b(k, l) + (a \leftrightarrow b) \right].
\end{align*}
\]

All six polarization tensors (24.71)–(24.76) satisfy the Hilbert condition in momentum space (24.70).

Just as in the electromagnetic case, we expect several of the six components (24.71)–(24.76) to be unphysical. Indeed, the gauge transformations on the field \( h_{ab} \) are

\[
h_{ab} \to h_{ab} + \partial_a \xi_b + \partial_b \xi_a.
\]

For the field \( \phi^{ab} \), these become

\[
\phi^{ab} \to \phi^{ab} + \partial^a \xi^b + \partial^b \xi^a - \eta^{ab} \partial_c \xi^c.
\]

Thus, in momentum space, any combination

\[
k_a \epsilon^b + k^b \epsilon^a - \eta^{ab} k_c \xi^c
\]

does not contribute to observable quantities. Since \( \epsilon^a(k, l) \) is proportional to \( k^a \), the choice \( \xi^a = \epsilon^a(k, 1), \epsilon^a(k, 2), \epsilon^a(k, l) \) eliminates directly three components involving the longitudinal polarization vector \( \epsilon^a(k, l) \).

A forth unphysical component is obtained by choosing \( \xi^a \) to be the fourth polarization vector \( \epsilon^a(k, l') \) in (24.59) which was dropped in the expansion of the electromagnetic waves, since it did not satisfy the Lorentz condition, i.e., \( k_a \epsilon^a(k, l') \neq 0 \).
The vectors $\epsilon^a(\mathbf{k},1), \epsilon^a(\mathbf{k},2), \epsilon^a(\mathbf{k},l), \epsilon^a(\mathbf{k},l')$ form a complete basis. They satisfy the orthogonality relation
\[ \epsilon^a(\mathbf{k},p)\epsilon(\mathbf{k},p') = \eta_{pp'}, \quad (24.80) \]
with the metric
\[ \eta_{pp'} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (24.81) \]
and the completeness relation
\[ \epsilon^a(\mathbf{k},1)\epsilon^b(\mathbf{k},1) + \epsilon^a(\mathbf{k},2)\epsilon^b(\mathbf{k},2) + \left[ \epsilon^a(\mathbf{k},l)\epsilon^b(\mathbf{k},l') + (a \leftrightarrow b) \right] = \eta^{ab}. \quad (24.82) \]
The completeness relation shows that the polarization tensors (24.71)–(24.76) satisfy the identity
\[ \epsilon^{ab}(\mathbf{k},1,1) + \epsilon^{ab}(\mathbf{k},2,2) = \eta^{ab} - \left[ \epsilon^a(\mathbf{k},l)\epsilon^b(\mathbf{k},l') + (a \leftrightarrow b) \right]. \quad (24.83) \]
The right-hand side is exactly the pure gauge configuration (24.79) with $\xi^a = \epsilon^a(\mathbf{k},l')/|\mathbf{k}|$. The sum of the polarization tensors on the left-hand side is therefore unphysical. Thus we remain with their orthogonal combination
\[ \epsilon^a_{+}(\mathbf{k}) \equiv \frac{1}{\sqrt{2}} \left[ \epsilon^{ab}(\mathbf{k},1,1) - \epsilon^{ab}(\mathbf{k},2,2) \right] \quad (24.84) \]
as a physical polarization state. A second physical polarization tensor is (24.72):
\[ \epsilon^a_{\times}(\mathbf{k}) \equiv \epsilon^{ab}(\mathbf{k},1,2) \quad (24.85) \]
Both tensors are traceless in the spacetime indices and the spatial indices alone, and have therefore no 00-component:
\[ \epsilon^a_{+\ a}(\mathbf{k}) = \epsilon^a_{\times\ a}(\mathbf{k}) = 0, \quad \epsilon^i_{+\ i}(\mathbf{k}) = \epsilon^i_{\times\ i}(\mathbf{k}) = 0, \quad \epsilon^0_{+}(\mathbf{k}) = \epsilon^0_{\times}(\mathbf{k}) = 0. \quad (24.86) \]
The physical meaning of the subscripts + and × will be understood in the next section.

We now observe that these tensor be re-expressed in terms of products of two complex circularly polarization vectors
\[ \epsilon^{ab}(\mathbf{k},\pm 2) \equiv \epsilon^a(\mathbf{k},\pm 1)\epsilon^b(\mathbf{k},\pm 1) \]
\[ = \frac{1}{2} \left[ \epsilon^a(\mathbf{k},1)\epsilon^b(\mathbf{k},1) - \epsilon^a(\mathbf{k},2)\epsilon^b(\mathbf{k},2) \right] \pm \frac{i}{2} \left[ \epsilon^a(\mathbf{k},1)\epsilon^b(\mathbf{k},2) + (a \leftrightarrow b) \right] \]
\[ = \frac{1}{\sqrt{2}} \left[ \epsilon^{ab}_{+}(\mathbf{k}) \pm i\epsilon^{ab}_{\times}(\mathbf{k}) \right] \quad (24.87) \]
They are called helicity tensors. We shall see later that the tensors $\epsilon(\mathbf{k},\pm 2)$ correspond to circularly polarized gravitational waves. The inverse relation is
\[ \epsilon^a_{+}(\mathbf{k}) = \frac{1}{\sqrt{2}} \left[ \epsilon^{ab}(\mathbf{k},+2) + \epsilon^{ab}(\mathbf{k},-2) \right], \quad (24.88) \]
\[ \epsilon^a_{\times}(\mathbf{k}) = \frac{1}{\sqrt{2}i} \left[ \epsilon^{ab}(\mathbf{k},2) - \epsilon^{ab}(\mathbf{k},-2) \right]. \quad (24.89) \]
The two polarization tensors satisfy the completeness relation

$$\epsilon^a_{+}(k)\epsilon^{cd+}(k) + \epsilon^a_{\times}(k)\epsilon^{cd\times}(k) = \Lambda^{abcd}(\hat{k}), \quad (24.90)$$

where the spatial components of $\Lambda^{ab,cd}(\hat{k})$ form the projection matrix

$$\Lambda_{ij,kl}(\hat{k}) \equiv \frac{1}{2} \left[ \Lambda_{ik}(\hat{k})\Lambda_{jl}(\hat{k}) + \Lambda_{il}(\hat{k})\Lambda_{jk}(\hat{k}) - \Lambda_{ij}(\hat{k})\Lambda_{kl}(\hat{k}) \right], \quad (24.91)$$

When contracted with symmetric tensors, we can work with the simpler and more explicit expression

$$\Lambda_{ij,kl}(\hat{k}) = \delta^{ik}\delta^{jl} - \frac{1}{2}\delta^{ij}\delta^{kl} - \frac{1}{2}\hat{\epsilon}^{ik}\hat{\epsilon}^{jl} + \frac{1}{2} \left( \delta^{ij}\hat{k}^k\hat{k}^l + \delta^{jl}\hat{k}^i\hat{k}^k \right) + \frac{1}{2} \hat{k}^i\hat{k}^j\hat{k}^k\hat{k}^l. \quad (24.93)$$

### 24.4 Nonexistence of Gravitational Waves in $D=3$ and $D=2$ Spacetime Dimensions

The above counting procedure has immediate consequence upon the existence of gravitational waves in a hypothetical lower-dimensional world. In three dimensions, the symmetric tensor $\phi^{ab}$ has six independent components, three of which are eliminated by the Hilbert condition $\partial_a \phi^{ab} = 0$. This leaves only three components. These, however, are just as many as there are gauge degrees of freedom

$$\phi^{ab} \rightarrow \phi^{ab} + \partial^a \xi^b + \partial^b \xi^a - \eta^{ab}\partial_c \xi^c. \quad (24.94)$$

In fact, all field degrees of freedom are gauge degrees of freedom. We see this by choosing $k$ in the $z$-direction, where polarization vectors can be written as

$$\epsilon^a(l) = \frac{1}{\sqrt{2}} \frac{1}{|k|} k^a = \frac{1}{\sqrt{2}} (1,0,1),$$
$$\epsilon^a(l') = \frac{1}{\sqrt{2}} (1,0,-1),$$
$$\epsilon^a(1) = (0,1,0). \quad (24.95)$$

They allow us to form the three symmetric polarization tensor which satisfy the Hilbert condition

$$\epsilon^a(l)\epsilon^b(l)$$
$$\epsilon^a(l)\epsilon^b(l) + (a \leftrightarrow b)$$
$$\epsilon^a(1)\epsilon^b(1). \quad (24.96)$$
The first is proportional to $k^a k^b$, i.e., to a pure gauge of the form with $\xi^a = k^a$. The second is a pure gauge with $\xi^a = \epsilon^a(1)$. The third, finally, is equal to

$$\eta^{ab} - [\epsilon^a(l)\epsilon^b(l') + \epsilon^a(l')\epsilon^b(l)],$$

(24.97)

which has the pure gauge form (24.94) with $\xi^a = \epsilon^a(l')/\sqrt{2}|k|$.

This implies that the three-dimensional gravitational field has no dynamical degrees of freedom. There exist no freely propagating gravitational waves in three spacetime dimensions.

A three-dimensional Einstein theory would have a further disease. It would not even possess a Newtonian weak-field limit. To see this, suppose we had found a field $\phi^{ab}$ from the equation

$$\partial^2 \phi^{ab} = -2\kappa T^{ab}.$$  

(24.98)

For the solution $\phi^{ab}$, we calculate in $D$ dimensions

$$h^{ab} = \phi^{ab} - \frac{1}{D-2} \eta^{ab} \phi.$$  

(24.99)

In order for the weak-field limit to satisfy Newton’s equation of motion it is necessary that for a massive point particle of the origin, $h^{00}$ satisfies the Poisson differential equation

$$\partial^2 h^{00} = GM\delta^{(3)}(x).$$  

(24.100)

Moreover, $T^{ab}$ has only a $T^{00}$-component appearing in (24.100). Hence $\phi^{ab}$ has only a single nonvanishing component $\phi^{00}$. In three dimensions this implies that

$$h^{00} = \phi^{00} - \frac{1}{D-2} \eta^{00} \phi = \frac{D-3}{D-2} \phi^{00}.$$  

(24.101)

But $\phi^{00}$ must satisfy (24.98)! Hence the Newton potential vanishes identically.

Another equivalent place where this disease manifests itself is in the coupling of the gravitational field $\phi^{ab}$ to the energy-momentum tensor of a massive particle. Since this coupling is

$$H^{ab} T_{ab}$$  

(24.102)

we see that

$$h^{ab} T_{ab} = \left( \phi^{ab} - \frac{\eta^{ab}}{D-2} \phi \right) T_{ab} = \phi^{ab} \left( T_{ab} - \frac{\eta_{ab}}{D-2} T^c_c \right).$$  

(24.103)

With the particle being at rest and $T_{ab}$ having only a 00-component, the interaction becomes

$$\phi^{00} \left( T_{00} - \frac{\eta_{00}}{D-2} T_{00} \right),$$  

(24.104)
which vanishes for \( D = 3 \).

Only by a technical trick can we obtain a non-zero limit. The theory has to be defined for continuous dimensions \( D \) in the neighborhood of 3 with a coupling constant which diverges at \( D - 3 \):

\[
\kappa = \kappa_0 / (D - 3).
\]

The above-described diseases are not a consequence of the linear approximation to gravity. We have mentioned before that in three dimensions, the full curvature tensor \( R_{\mu\nu\lambda\kappa} \) is completely determined in terms of the Ricci tensor

\[
R_{\mu\nu} = G_{\mu\nu} - \frac{1}{D - 2} g_{\mu\nu} G,
\]

and thus in terms of the Einstein tensor \( G_{\mu\nu} = \kappa T_{\mu\nu} \). Due to Einstein’s equation \( G_{\mu\nu} = \kappa T_{\mu\nu} \), \( R_{\mu\nu} \) and thus \( R_{\mu\nu\lambda\kappa} \) vanishes identically everywhere, except right at the mass point. This implies, in particular, that the empty-space field equation

\[
G_{\mu\nu} = 0,
\]

which we have used in four dimensions to find the Schwarzschild metric, can have only the trivial solution \( g_{\mu\nu} = \eta_{\mu\nu} \), up to a trivial reparametrization of space.\(^1\)

It is curious to note that this disease makes it possible to develop a quantum theory of gravity, in contrast to four dimensions where such a theory does not yet exist. It is possible to define a wave function for the universe. In the absence of matter, the entire Hilbert space consists only of one state, the vacuum \( |0\rangle \).\(^2\)

Let us take a look at gravitational waves in two dimensions. There the only polarization tensor satisfying Hilbert’s constraint is

\[
\epsilon^{ab}(k) = \frac{1}{|k|^2} k^a k^b,
\]

and this is obviously a pure gauge. As far as the equation (24.99) is concerned, the situation is even worse than in three dimensions. It is impossible to recover from \( \phi^{ab} \) the metric \( g_{ab} \) since the equation

\[
h_{ab} = \phi_{ab} - \frac{\eta_{ab}}{D - 2} \phi
\]

is meaningless for \( D = 2 \). In order to see the origin of the problem let us choose another than the Hilbert gauge, and use the gauge freedom

\[
\begin{align*}
h_{00} &\rightarrow h_{00} + 2 \partial_0 \xi_0 \\
h_{01} &\rightarrow h_{01} + \partial_0 \xi_1 + \partial_1 \xi_0
\end{align*}
\]
to make $h_{00}$ and $h_{01}$ vanish. For the remaining component $h_{11} \equiv h$ we find from (20.24) the Lagrangian density:

$$L = \frac{1}{4} \left[ (\partial h)^2 - 2(\partial_1 h)^2 + 2(\partial_1 h)^2 - (\partial h)^2 \right] = 0$$

(24.110)

which vanishes identically. This property could, in fact, have been anticipated. It is well known in differential geometry that in two dimensions the Einstein action

$$A = \frac{1}{2\kappa} \int d^2 \xi \sqrt{g} R$$

(24.111)

is a pure surface term. By the Gauss-Bonnet theorem, it is entirely determined by the global topological properties of the space. For closed surfaces, it is equal to $(4\pi/2\kappa)(1 - h)$, where $h$ is the number of handles of the surface. This makes it impossible to derive equations of motion from such an action.

For completeness, let us compare the situation with the electromagnetism case. In three dimensions, the vector potential has three components minus one, due to the Lorentz condition. This leaves two degrees of freedom. One of them is a pure gauge mode, the other is physical. Hence there exists a freely propagating photon in three dimensions.

In two dimensions, the Lorentz gauge allows $A^a$ to be written in the form of a two-dimensional curl:

$$A^a = \epsilon^{ab} \partial_b \phi.$$  

(24.112)

In empty space, it satisfies the free field equation

$$\partial^2 A^a = 0.$$  

(24.113)

In terms of the field $\phi$, the field strength is

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = -\partial^2 \phi.$$  

(24.114)

Note the relation with the theory of complex functions. For zero field, $\phi$ is a harmonic function. It can therefore be considered as the real part of an analytic function

$$\phi(x) = \text{Re}\, f(z),$$  

(24.115)

where

$$z = x + iy.$$  

(24.116)

The vector $A^a$ is the real part of the gradient of this function.

$$A^a = \epsilon^{ab} \partial_b \text{Re}\, f(z).$$  

(24.117)
Now, the real and imaginary part of $f$ satisfy the Cauchy-Riemann differential equations,

$$\begin{align*}
\partial_1 \text{Re} f &= \partial_2 \text{Im} f \\
\partial_1 \text{Im} f &= -\partial_2 \text{Re} f.
\end{align*}$$

(24.118)

Hence we can also write

$$A^a = \partial^a \text{Im} f.$$  

(24.119)

This shows that $A^a$ is a pure gauge which, therefore, carries no electromagnetic field, in agreement with the assumption we started out from.

It is important to point out that the nonexistence of propagating electromagnetic waves does not rule out Coulomb forces. They do not need the exchange of physical electromagnetic waves. The exchange of virtual electromagnetic fields with zero momentum is sufficient. The inhomogeneous equation

$$\partial^2 A^a = -\frac{1}{c} j^a$$

(24.120)

has, outside the charge distribution, only the solution $\partial^2 A^a = 0$, and we have seen before that the only field satisfying this is a pure gauge field. Still, the Maxwell equation

$$\partial_1 F^{01} = \frac{1}{2} j^0$$

(24.121)

allows for a nonvanishing constant field

$$F^{01} = \text{const.}$$

(24.122)

It is carried by the $k = 0$-components of $A_a$ to which the previous polarization discussions do not apply. It is this component which gives rise to the Coulomb force in two space-time dimensions. Another way to see this is by considering the Maxwell action

$$A = \frac{1}{2c} \int d^2 x F_{ab} F^{ab} + \frac{1}{c^2} \int d^2 x j^a A_a,$$

(24.123)

and going to the Lorentz gauge,

$$A^a = \epsilon^{ab} \partial_b \phi$$

(24.124)

in which it becomes (up to surface terms)

$$A = \frac{1}{2c} \int d^2 x (\partial^2 \phi)^2 + \frac{1}{c^2} \int d^2 x \epsilon^{ab} (\partial_a j_b) \phi.$$  

(24.125)

Extremizing this in the field $\phi$ gives the field equation

$$ (\partial^2)^2 \phi = \frac{1}{c} \epsilon^{ab} \partial_a j_b.$$  

(24.126)
Reinserting this into the action, the extremum is found to be

\[ A_{\text{extr}} = \frac{1}{c^2} \int d^2 x \, \epsilon^{ab} \partial_a j_b \frac{1}{(\partial_j^2)^2} \epsilon^{b'} \partial_{a'} j_{b'}. \]  

(24.127)

Using

\[ \epsilon^{ab} \epsilon^{a'b'} = \eta^{aa'} \eta^{bb'} - \eta^{ab'} \eta^{ba'}, \]  

(24.128)

performing a partial integration, and taking advantage of current conservation \( \partial_a j^a = 0 \), this becomes

\[ A_{\text{extr}} = \frac{1}{c^2} \int d^2 x \, j_a \frac{1}{\partial^2} j^a \]  

(24.129)

which is precisely the Biot-Savard interaction law between currents. Using once more current conservation in the form

\[ \int d^2 x (j^0_0^2 - j^1_1^2) = \int d^2 x \left[ j^0_0 - \partial_1 j_1 \frac{1}{(\partial_1)^2} \partial_1 j_1 \right] \]

\[ = \int d^2 x \left[ \partial_1 j^0_0 - \partial_0 j^0_0 \frac{1}{(\partial_1)^2} \partial_1 j_0 \right] \]

\[ = \int d^2 x \, j^0_0 \frac{\partial^2}{(-\partial_1)^2} j^0_0 \]

we can rewrite (24.129) as

\[ A_{\text{extr}} = \frac{1}{2c^2} \int d^2 x \, j^0_0 \frac{1}{(-\partial_1)^2} j^0_0. \]  

(24.131)

This is an instantaneous linear potential between the charges carried by \( j_0 \) reflecting the constant electric field \( F^a \) allowed in empty space.

In this respect, the situation is quite different from the gravitational case in three dimensions. There the absence of \( R_{\mu\nu} \) outside a mass distribution implies a vanishing of the gauge invariant curvature tensor \( R_{\mu\nu\lambda\kappa} \) and hence the vanishing of all tidal forces.

### 24.5 Detection of Gravitational Waves

In order to detect gravitational waves we consider a test particle whose equation of motion reads

\[ \frac{du^a}{d\tau} = - \left( \partial_0 h^a_{bc} - \frac{1}{2} \partial^b h_{bc} \right) u^b u^c. \]  

(24.132)

If it is initially at rest or moving very slowly, the right-hand side reduces to

\[ - \left( \partial_0 h^a_{00} - \frac{1}{2} \partial^a h_{00} \right) c^2. \]  

(24.133)
Since the two physical polarization tensors have neither a 00 nor 0i components, this vanishes. Thus two particles retain their relative positions as seen from the background Minkowski frame.

This does not mean that their physical distance remains unchanged. This distance is evaluated not via the Minkowski metric $\eta_{ab}$, but via the proper slightly distorted metric $g_{ab} = \eta_{ab} + h_{ab}$. If a wave with a linear polarization $\epsilon_+(k)$ runs along the $z$-axis

$$h_{ab} = \frac{1}{\sqrt{2}} [\epsilon_a(1)\epsilon_b(1) - \epsilon_a(2)\epsilon_b(2)] a_+ e^{i(kz-\omega t)} + c.c.,$$  \hspace{1cm} (24.134)$$

and hits two particles at $(0, d_0/2, 0, 0)$ and $(0, -d_0/2, 0, 0)$, which initially do not move due to the vanishing of $du^a/\tau$, their spatial distance changes as follows

$$d^2 = -(0, d_0, 0, 0)^a (\eta_{ab} + h_{ab}) \left( \begin{array}{c} 0 \\ d_0 \\ 0 \\ 0 \end{array} \right)^b$$

$$= d_0^2 \left[ 1 - \frac{a}{\sqrt{2}} \cos(\omega t - \delta) \right]$$  \hspace{1cm} (24.135)$$

where $\delta$ is the phase of the wave amplitude $a_+$. For particles which are separated by $d_0$ along the $y$ axis, the sign in front of $a_+$ is the opposite. In general, a pair of mass points at positions

$$\pm \left( 0, \frac{d_0}{2} \cos \phi, \frac{d_0}{2} \sin \phi, 0 \right)$$  \hspace{1cm} (24.136)$$

has a distance

$$d^2 = d_0^2 \left[ 1 - \frac{a}{\sqrt{2}} \cos(2\phi) \cos(\omega t - \delta) \right].$$  \hspace{1cm} (24.137)$$

We can picture the change in the distance by imagining a circular necklace of mass points placed in the gravitational beam. If the momentum points orthogonal to the paper plane, the circle distorts into vertical and horizontal ellipses, as shown in Fig. 24.1. This is why we have chosen the notation $\Sigma_t$ for the polarization tensor of this wave. The distortions are of quadrupole character, and the area within the necklace remains invariant, due to the tracelessness of $\epsilon^a(1)\epsilon^b(2) + (a \leftrightarrow b)$ has the determinant $\sqrt{-g} = 1 + O(a^2)$ implying the volume element $d^3x' = \sqrt{g}d^3x$ is invariant, to lowest order in $a$). For a wave with polarization tensor $\epsilon_+(k_{zab}) = [\epsilon_a(1)\epsilon_b(2) + (a \leftrightarrow b)]/\sqrt{2}$, the distance changes with time as follows

$$d^2(\tau) = d_0^2 \left[ 1 - \frac{a}{\sqrt{2}} \sin(2\phi) \cos(\omega t - \delta) \right].$$  \hspace{1cm} (24.138)$$

H. Kleinert, GRAVITY WITH TORSION
Figure 24.1 Distortions of a circular array of mass points by the passage of a gravitational quadrupole wave. The left is caused by a polarization tensor $\epsilon^{ab}_+$, the second from $\epsilon^{ab}_\times$.

The situation is the same as before, except for a rotation $\phi \rightarrow \phi - \pi/4$. Thus the necklace undergoes the same quadruple distortions, except that the principal axes of the ellipses lie along the diagonals as indicated by the subscript $\times$ of the polarization tensor. The rotation by $45^0$ can also be displayed more directly by writing the polarization tensor $\epsilon^{ab}(k_z)$ of Eqs. (24.89), (24.72) as

$$\epsilon^{ab}_\times(k_z) = \frac{1}{\sqrt{2}} [\epsilon^a(1)\epsilon^b(2) + (a \leftrightarrow b)] = \frac{1}{\sqrt{2}} \left[ \epsilon^a(\nearrow)\epsilon^b(\swarrow) - \epsilon^a(\swarrow)\epsilon^b(\nearrow) \right]. \quad (24.139)$$

where

$$\epsilon^a(\nearrow) = \frac{1}{\sqrt{2}}[\epsilon^a(1) + \epsilon^a(2)], \quad \epsilon^a(\swarrow) = \frac{1}{\sqrt{2}}[\epsilon^a(1) - \epsilon^a(2)] \quad (24.140)$$

are the diagonal polarization vectors. Similarly we rewrite $\epsilon^{ab}_+(k_z)$ of Eqs. (24.84), (24.71), (24.73) as

$$\epsilon^{ab}_+(k_z) = \frac{1}{\sqrt{2}}[\epsilon_a(1)\epsilon_b(1) - \epsilon_a(2)\epsilon_b(2)] = \frac{1}{\sqrt{2}} \left[ \epsilon^a(\nearrow)\epsilon^b(\swarrow) + \epsilon^a(\swarrow)\epsilon^b(\nearrow) \right]. \quad (24.141)$$

The right-hand sides of (24.139) and (24.141) have the same forms as $\epsilon_+(k_z)_{ab}$ and $\epsilon_\times(k_z)_{ab}$ expressed in terms of respectively, but with $\epsilon_a(1)$ and $\epsilon_b(2)$ exchanged by the $45^0$-rotated diagonal polarization vectors $\epsilon^a(\nearrow)$ and $\epsilon^a(\swarrow)$.

The acceleration

$$\frac{d^2}{dt^2} = \frac{1}{\omega^2} \left\{ \frac{a}{\sqrt{2}} \cos 2\phi \right\} \left\{ \frac{\sin 2\phi}{\omega^2} \right\} \cos(\omega t - \delta) \quad (24.142)$$

implies the presence of tidal forces acting upon the necklace. Their field lines are shown in Fig. 24.2 If gravitational waves of helicity $\pm 2$ hit the necklace, it is deformed into an ellipse with a fixed shape which rotates clockwise or counter-clockwise.
Figure 24.2 Field lines of tidal forces of a gravitational wave in the $z$ direction with polarization tensor $\epsilon^a_{\perp}$ and $\epsilon^a_{\times}$, respectively. The field lines change direction with $\cos \omega t$.

around the direction of the wave. The wave is circularly polarized. This is seen by taking again two particles at positions

$$\pm \left(0, \frac{d}{2} \cos \phi, \frac{d}{2} \sin \phi, 0\right),$$  \hspace{1cm} (24.143)

and measuring their distances in the metric

$$g_{ab} = \eta_{ab} + \left[\epsilon_a (+) \epsilon_b (+) e^{i(kz - \omega t)} + c.c.\right]$$  \hspace{1cm} (24.144)

which gives

$$d^2(t) = d_0^2 \left\{1 - [(\cos^2 \phi - \sin^2 \phi) + i2 \cos \phi \sin \phi] a e^{-i\omega t} + c.c.\right\}$$

$$= d_0^2 [1 - |a| \cos(\omega t - 2\phi - \delta)],$$  \hspace{1cm} (24.145)

implying the above described rotations of the necklace, the azimuthal angle $\phi$ acting merely as a phase shift.

The time-dependent of the length measured by the metric $g_{ab} = \eta_{ab} + h_{ab}$ has a direct experimental equivalence. If we take a piezoelectric crystal, then it shows a pulsating voltage due to the distance changes between the atoms, even though its atoms remain at rest in the Minkowski coordinates. The distance is given by the minima of the interatomic potentials. Since the atomic interactions are due to electromagnetism which spreads through a space with the metric $g_{ab}$, the distances of these minima change according to changes in $g_{ab}$, and this gives rise to the piezoelectric voltage.

How large are the distortions caused by a gravitational wave? If we assume a typical astrophysical source (for the emission mechanism see the next section) with an energy flux of $\approx 10^{10} \text{erg}/(\text{cm}^2\text{sec})$ at $\omega \sim 10^4/\text{sec}$, we calculate the distortion of the metric to be of the order

$$h_{ab} \approx 10^{-7}.$$  \hspace{1cm} (24.146)
This will make the distance earth-moon, which is $3.8 \times 10^{10}$ cm, oscillate by $\approx 10^{-7}$ cm = 10 Å. This distance is of the same order as the distance between the atoms in matter which is of the same order.

At present, a laser pulse ranging to the moon can at most detect length changes of the order of 10 cm.

Other possible observable effects are quadrupole vibrations which can be excited by incoming gravitational waves within the earth or the moon itself. Their natural frequencies are 54 minutes and 15 minutes, respectively. By studying seismometer data of the earth’s vibrations, Weber found in 1967 an upper limit for the flux of gravitational waves

$$\frac{\text{energy flux}}{\text{frequency}} < 3 \times 10^7 \frac{\text{erg}}{\text{cm}^2 \text{sec Hertz}}$$

at a frequency $3.1 \times 10^{-4}$ Hertz.

In 1972, Weber built a gravitational detector consisting of a cylindrical aluminum block of 1.53 m length and 0.66 m diameter (weight $\approx 1.41 \times 10^6$ g). The block has an eigenfrequency of 1.66 Hz. By a piezoelectric strain transducer, Weber measured length changes in the material of the block. Setting up on block at the University of Maryland and another one at Argonne National Laboratory and looking at coincidences between the two, he eliminated random vibrations caused by the daily activities in the neighborhood of each detector.

In 1972, he observed two sudden simultaneous excitations which he interpreted as a possible response to a gravitational wave passing through.

Unfortunately, his observations have, until now, not found any recurrence in spite of collective efforts of several laboratories. A better idea may be necessary to become sensitive to such small effects.

### 24.6 Attractive Gravity versus Repulsive Electric Forces between Like Charges

The energy of gravitational waves gives a simple insight why the fields lead to an attraction between masses (which are always positive) while electromagnetism is repulsive between like charges. The physical components of gravitational waves are the purely spatial ones $\phi^{ij}$. Their energy has to be positive and hence the field energy carries plus sign when expressed in momentum space

$$E_{\text{grav}} \propto k^2 (\phi^{ij})^2.$$  \hspace{1cm} (24.148)

In electromagnetism, the same is true for the spatial part of the electromagnetic field $A^i$

$$E_{\text{elm}} \propto k^2 (A^i)^2.$$ \hspace{1cm} (24.149)
As a simple consequence of Lorentz invariance, the components $\phi^0$ and $A^0$ have to appear with opposite signs in $E$:

$$E_{\text{grav}} \propto k^2 \phi^{02}$$
$$E_{\text{em}} \propto -k^2 A^{02}.$$ (24.150)

But these components are the relevant ones coupling to the mass density $T^{00}$ or the charge density $j^0$, respectively, thereby giving rise to Newton’s or Coulomb’s law. The opposite signs cause the forces in the opposite direction.

### 24.7 Nonlinear Gravitational Waves

Plane gravitational waves are such simple phenomena that they can easily be found as solutions to the full nonlinear Einstein equations in the vacuum

$$G_{\mu\nu} = 0.$$ (24.151)

If the wave runs in $z$-direction we can make the 2 dimensional Ansatz for the invariant distance

$$(ds)^2 = (dt)^2 - L^2[e^{2\beta}(dx)^2 + e^{-2\beta}(dy)^2] - (dz)^2$$ (24.152)

where $L$ and $\beta$ are functions depend only on $z$ and $t$. It is useful to go to so-called light cone coordinates

$$u = t - z,$$
$$v = t + z,$$ (24.153)

which sit on wave crests moving along the positive and negative $z$-directions. It can easily be verified that the only non-zero component of the Ricci tensor is

$$R_{uu} = -2L^{-1}(L'' + \beta'^2 L)$$ (24.154)

so that the nonlinear wave equation in empty space reads

$$L'' + \beta'^2 L = 0.$$ (24.155)

Our previous linear waves correspond to the equation

$$L'' \approx 0,$$ (24.156)

which can be solved by

$$L \approx 1.$$ (24.157)

In this limit,

$$ds^2 \approx -(1 + 2\beta)dx^2 - (1 - 2\beta)dy^2 - dz^2$$ (24.158)
corresponding to

\[
h_{ab} = \begin{pmatrix}
0 & 0 & 0 \\
0 & -2\beta & 0 \\
2\beta & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}
\]

(24.159)
i.e., a wave with a polarization tensor \(\epsilon_+(k_z)_{ab}\).

Going back to the nonlinear case, we assume the space to be flat before the wave arrives, say at

\[t - z = -T\]

(24.160)
i.e., we take

\[\beta = 0, \quad L = 1 \quad \text{for} \quad u < -T.\]

(24.161)
We shall assume that the wave comes as a pulse of width \(2T\) and assume that, after the pulse has passed, space is left again flat. While the pulse passes, i.e., for \(T > u > -T\), we allow for an arbitrary \(\beta(u)\) but assume, for simplicity, that the pulse is not sharp:

\[|\beta'(u)| \ll \frac{1}{T}.\]

(24.162)
Then we do not have to solve the full differential equation. While the pulse is passing, we have

\[
\beta(u) = \text{arbitrary, with} \quad |\beta'| \ll \frac{1}{T}
\]

(24.163)
\[
L(u) = 1 - \int_{-T}^{u} du' \int_{-T}^{u'} du''[\beta'(u'')]^2 + O((\beta'T)^4) \quad \text{for} \quad u \in (-T, T).
\]

After it has passed, \(\beta(u)\) is again zero by assumption, and

\[
L(u) = 1 - \frac{u}{a}\quad a \equiv \frac{1}{\int_{-T}^{T} \beta'^2 du} + O((\beta'T)^2) \quad \text{for} \quad u > T.
\]

(24.164)
The change in \(L\) from 1 to \(-u/a\) is physically not observable. The space is flat after the pulse has passed. This is seen by going to new coordinates \(X, Y, U, V\) defined by

\[
x = \frac{X}{1 - U/a}, \quad y = \frac{Y}{1 - U/a}, \quad u = U, \quad v = V + \frac{1}{a} \frac{X^2 + Y^2}{1 - U/a}
\]

(24.165)
which brings the invariant distance to the Minkowski form

\[ds^2 = dUdV - d^2 - (dY)^2.\]

(24.166)
24.8 Emission of Gravitational Waves

The emission of gravitational waves is governed by equation (24.67) which we shall write more explicitly as

\[ \phi^{ab}(x, t) = -\frac{2\kappa}{4\pi} \int d^3 x' \frac{1}{|x - x'|} T^{ab}(x', t - |x - x'|/c). \]  

(24.167)

Far away from the source, in the radiation zone, we may approximate \( 1/|x - x'| \approx 1/|x| \equiv 1/r \), and find the leading term [recall once more the exact splitting (24.20) into a velocity and an acceleration field which carries off anergy]:

\[ \phi^{ab}(x, t) \approx -\frac{2\kappa}{4\pi r} \int d^3 x' m^T_{ab}(x', t - |x - x'|/c). \]  

(24.168)

In linear approximation, the energy-momentum term of the source is conserved by itself (i.e., without the additional energy-momentum tensor of the gravitational field)

\[ \partial_b m^T_{ab} = 0. \]  

(24.169)

The neglected terms are of order \( O(\kappa) \). Hence

\[ \partial_0 m^T_{i0}(x, t) = -\partial_j m^T_{ij}(x, t), \]  

(24.170)

\[ \partial_0 m^T_{00}(x, t) = -\partial_j m^T_{0j}(x, t). \]  

(24.171)

From these equations we find that

\[ \int d^3 x' m^T_{ij}(x', t) = \int d^3 x' x'_{i} \partial_k m^T_{0k}(x', t) \]  

(24.172)

and further

\[ \int d^3 x' m^T_{ij}(x', t) = -\int d^3 x' x'_{j} \partial_k m^T_{ik}(x', t) = \int d^3 x' x'_{i} \partial_k m^T_{0k}(x', t) \]  

(24.173)

and

\[ \int d^3 x' [x'_{i} m^T_{j0}(x', t) + x'_{j} m^T_{i0}(x', t)] = -\int d^3 x' x'_{i} x'_{j} \partial_k m^T_{k0}(x', t) \]  

(24.174)

Hence

\[ \int d^3 x' m^T_{ij}(x', t') = \frac{1}{2} \partial_0 \int d^3 x' x'_{i} x'_{j} m^T_{00}(x', t'). \]  

(24.175)
Using this, it is straightforward to find the components $\phi^{00}(x, t)$ and $\phi^{ij}(x, t)$ from the time dependent energy density $T^{m\,00}(x', t')$:

$$
\phi^{00}(x, t) = -\frac{2\kappa}{4\pi r} \int d^3 x' \frac{m}{T^{00}(x', t') | t' = t - r/c} \tag{24.176}
$$

$$
\phi^{ij}(x, t) = -\frac{2\kappa}{4\pi r} \partial_0 \int d^3 x' \int d^3 x' x'^i x'^j \frac{m}{T^{00}(x', t') | t' = t - r/c} \tag{24.177}
$$

$$
\phi^{ij}(x, t) = -\frac{2\kappa}{4\pi r} \frac{1}{2} \partial_0^2 \int d^3 x' x'^i x'^j \frac{m}{T^{00}(x', t') | t' = t - r/c}. \tag{24.178}
$$

We now introduce the quadrupole moment of the mass distribution

$$
Q^{ij}(t) \equiv \frac{1}{c} \int d^3 x' \left(x'^i x'^j - \frac{1}{3} \delta^{ij} \right) \frac{m}{T^{00}(x', t)}, \tag{24.179}
$$

defined as the traceless second moment of the mass distribution, and the trace part of the second moment

$$
S(t) \equiv \frac{1}{c} \int d^3 x' r'^2 \frac{m}{T^{00}}, \tag{24.180}
$$

we can write

$$
\phi^{ij}(x, t) = -\frac{\kappa}{4\pi r} c \partial_0^2 \left[ Q^{ij}(t') + \frac{1}{3} \delta^{ij} S(t') \right] | t' = t - r/c. \tag{24.181}
$$

Note that this relation is true only if the forces acting upon the matter points are entirely due to gravity, in which case the neglected field part in the energy-momentum conservation law $\partial_\alpha \left( T^{m\,ab} + f^{m\,ab} \right)$ are of second order in $\kappa$ and thus negligible. If there are other forces, for instance electromagnetic ones, this adds another term to the conservation law. In this case we must use the original source in (24.168).

Far away from the source, $\phi^{ab}(x, t)$ behaves locally like an outgoing plane wave. Thus it must be purely transverse (some linear combination of the polarization tensors $\epsilon_+$ and $\epsilon_\times$ in Eqs. (24.88) and (24.89)]. Thus only the spatially traceless quadrupole part will contribute

$$
\phi^{ij}_{\text{phys}}(x, t) = -\frac{\kappa}{4\pi r c} Q^{ij}(t) | t' = t - r/c. \tag{24.182}
$$

Indeed, by the arguments leading to (24.86), the $S(t)$-term drops out from the equation for the radiation energy far away from the source, where the wave is traveling freely.

The calculation leading to the quadrupole formula is of course the complete analog of the calculation in (24.45) leading to the electric dipole in electromagnetism. Indeed, inserting (24.45) into (24.3) we find directly the vector potential expressed in terms of the electric dipole moment (24.46):

$$
A(x, t) = -\frac{1}{cr} d(t') | t' = t - \frac{r}{c}. \tag{24.183}
$$
Note that there is also quadrupole radiation from an electromagnetic source, but it gives only a higher-order correction carrying two more powers $d/\lambda$ where $\lambda$ is the wavelength and $d$ is the diameter of the source.

The above-calculated field components $\phi^{ij}(x, t)$ are sufficient to determine the energy carried off by the wave. Let $n \equiv k/|k|$ be the spatial direction of the $k$-vector. According to Eq. (20.25), the energy-momentum tensor of the field in linear approximation is, in the Hilbert gauge $\partial_a \phi^{ab} = 0$,

$$T^{ab} = \frac{1}{8\kappa} \left[ 2\partial^a \phi^{cd} \partial^b \phi_{cd} - \partial^a \phi \partial^b \phi - \eta^{ab} \left( \partial_c \phi \partial^c \phi_{de} - \frac{1}{2} \partial_c \phi \partial^c \phi \right) \right]. \quad (24.184)$$

This formula shows that a term proportional to $\delta_{ij}$ gives the energy current density along the $n^i$ direction. Separating $\phi^{ab}$ into space and time parts, we can write

$$\frac{d^2 E}{dtd\Omega r^2} = c^2 T^{0i} n^i = \frac{c^2}{8\kappa} \left[ 2\partial^0 \phi^{kl} \partial^i \phi_{kl} - 4\partial^0 \phi^{k0} \partial^i \phi^{k0} + \partial^0 \phi^{00} \partial^i \phi^{00} + \partial^0 \phi^{kk} \partial^i \phi^{ij} - \partial^0 \phi^{kk} \partial^i \phi^{jk} \right] n^i. \quad (24.185)$$

We now express the fields on the right-hand side in terms of time derivatives of the purely spatial field components

$$\phi^{kl}(x, t) = -\frac{\kappa}{4\pi r c} Q^{kl}(t - r/c). \quad (24.187)$$

which are spatially traceless and have no 00 component: $\phi^{kk} = 0$, $\phi^{00} = 0$ [recall (24.86)]. We also note that in the asymptotic field $\phi^{ab}(x, t)$ of Eq. (24.168) we can forget all space derivatives of the prefactor $1/r$ since they give nonleading $1/r^2$, $1/r^3$, \ldots contributions. We do have to maintain, however, derivatives with respect to the $r$ arising from the retarded time argument [recall the approximation (24.31)]. Hence we can approximate, to leading order in $1/r$,

$$\partial^i \phi^{ab} \approx -\partial^0 \phi^{ab} \partial^i r = \partial^i \phi^{ab} n^i. \quad (24.188)$$

Applying $\partial^i$ to $\phi^{kl}$ brings the first term in (24.186) to the form $2\tilde{\omega}^{kl}$. The other terms can be reduced in a similar way by making use of the Hilbert gauge condition. First we have

$$\partial_0 \phi^{k0} = -\partial_j \phi^{ki} = \partial^0 \phi^{k0} n^i, \quad (24.189)$$

second

$$\partial_0 \phi^{00} = -\partial_j \phi^{j0} = \partial_0 \phi^{j0} n_i = -\partial_j \phi^{ij} n_j = \phi^{ij} n_i n_j \quad (24.190)$$
and third

\[ \partial^i \phi^{00} = \partial_0 \phi^{00} n^i. \]  

(24.191)

In this way, the energy current becomes, to leading order in \( 1/r \),

\[ \frac{d^2 E}{dtd\Omega r^2} = c^2 T^0 n_i = \frac{1}{4\kappa} \left[ \dot{\phi}_{kl}^2 - 2\dot{\phi}_{kl} \dot{\phi}_{km} n^m + \frac{1}{2} \dot{\phi}^{kl} \dot{\phi}^{mn} n_k n_m n_r \right]. \]  

(24.192)

Comparison with (24.93) shows that due to the tracelessness of \( \phi_{ij} \), this can be written as

\[ \frac{d^2 E}{dtd\Omega r^2} = c^2 T^0 n_i = \frac{1}{4\kappa} \Lambda_{ij,kl}(\hat{n}) \dot{\phi}_{ij} \dot{\phi}_{kl}, \]  

(24.193)

where we have used the asymptotic tracelessness of the radiation field \( \phi_{ij} \).

Inserting (24.187) into (24.192), we obtain the rate of energy emitted into the solid angle \( d\Omega \):

\[ \frac{d^2 E}{dtd\Omega r^2} = c^2 T^0 n_i = \frac{1}{4\kappa} \Lambda_{ij,kl}(\hat{n}) \dot{\phi}_{ij} \dot{\phi}_{kl}, \]  

(24.194)

\[ \frac{d^2 E}{dtd\Omega r^2} = c^2 T^0 n_i = \frac{1}{4\kappa} \Lambda_{ij,kl}(\hat{n}) \dot{\phi}_{ij} \dot{\phi}_{kl}, \]  

(24.195)

We therefore write the energy loss of the quadrupole per solid angle \( d\Omega \) as

\[ \frac{d^2 E}{dtd\Omega r^2} = c^2 T^0 n_i = \frac{1}{4\kappa} \Lambda_{ij,kl}(\hat{n}) \dot{\phi}_{ij} \dot{\phi}_{kl}, \]  

(24.196)

This has a complicated angular dependence. The integral over all directions, however, is easily found using the angular averages of products of unit vectors

\[ \langle \hat{k}_i \hat{k}_j \rangle = \frac{1}{3} \delta_{ij}, \quad \langle \hat{k}_i \hat{k}_j \hat{k}_k \hat{k}_m \rangle = \frac{1}{15} \left( \delta^{ij} \delta^{kl} + \delta^{ik} \delta^{jl} + \delta^{il} \delta^{jk} \right). \]  

(24.197)

The tensor structure on the right-hand sides follow directly from the rotational symmetry. The normalizations are found by a contracting of the indices, using \( \hat{k}_i \hat{k}_i = 1 \). Inserting this into (24.91) we find

\[ \langle \Lambda^{ijkl}(\hat{k}) \rangle = \frac{2}{5} \left[ \frac{1}{2} \left( \delta^{ik} \delta^{jl} + \delta^{il} \delta^{jk} \right) - \frac{1}{3} \delta^{ij} \delta^{kl} \right]. \]  

(24.198)

Now the angular integral is straightforward, yielding

\[ \frac{dE}{dt} = \frac{1}{2} \frac{\kappa}{8\pi c^5} \left( \dot{Q}^{kl} \right)^2 = \frac{G}{5c^5} \left( \dot{Q}^{kl} \right)^2. \]  

(24.199)
For comparison, recall the analogous formulas in the electromagnetic case where the direction-dependent energy loss of dipole radiation is

\[
\frac{d^2 E}{dt d\omega} = r^2 (\mathbf{E} \times \mathbf{B}) \cdot \mathbf{n} = \frac{1}{8\pi c^3} |\mathbf{n} \times (\mathbf{n} \times \mathbf{d})|^2,
\]

(24.200)

and the total radiated power:

\[
\frac{dE}{dt} = \frac{1}{3c^3} \ddot{d}^2
\]

(24.201)

A useful general formula for the radiated power can be derived by rewriting Eq. (24.167) as

\[
\phi_{ab}(\mathbf{x}, t) = -\frac{2\kappa}{4\pi} \int d^3 x' \frac{1}{|\mathbf{x} - \mathbf{x}'|} \int \frac{d\omega}{2\pi} m_{ab}(\mathbf{x}', \omega) e^{-i\omega t + i\omega|\mathbf{x} - \mathbf{x}'|/c},
\]

(24.202)

where

\[
m_{ab}(\mathbf{x}', \omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} T_{ab}(\mathbf{x}', t)
\]

(24.203)

are the Fourier components of the energy-momentum tensor. Approximating for large \( r = |\mathbf{x}| \)

\[
e^{i\omega|\mathbf{x} - \mathbf{x}'|/c} \approx \frac{e^{i\omega/c}}{r} e^{-i\omega nx'/c}
\]

(24.204)

leaving an \( \mathbf{x}' \)-dependence only in the sensitive phase factor, and setting \( \omega nx'/c = kx' \), where \( k \) is the wave vector of the observed gravitational wave, we find

\[
\phi_{ab}(\mathbf{x}, t) = -\frac{2\kappa e^{i\omega/c}}{4\pi} \int d\omega \frac{1}{2\pi} m_{ab}(\mathbf{k}, \omega),
\]

(24.205)

where \( \hat{T}_{ab}(\mathbf{k}, \omega) \) is the Fourier transform in spacetime:

\[
\hat{T}_{ab}(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} dt \int d^3 x e^{i\omega t - i\mathbf{k}\mathbf{x}} \hat{T}_{ab}(\mathbf{x}, t).
\]

(24.206)

We can also write

\[
\phi_{ab}(\mathbf{x}, t) = -\frac{2\kappa e^{i\omega/c}}{4\pi} \hat{T}_{ab}(\mathbf{k}, t).
\]

(24.207)

At a point \( \mathbf{x} \) far away from the source, the spherically radiated field (24.205) looks like a passing plane wave with \( e^{i\omega/c} \approx e^{ikx} \). The formula (24.193) yields for the radiated energy per unit angle and time:

\[
\frac{dE}{dt d\Omega} = \frac{\kappa}{16\pi} \Lambda_{ijkl}(\hat{k}) \left[ \partial_t \hat{T}_{ij}(\mathbf{k}, t) \right]^* \left[ \partial_t \hat{T}_{kl}(\mathbf{k}, t) \right].
\]

(24.208)
This can also be written in a fully covariant form as

\[
\frac{dE}{dtd\Omega} = \frac{\kappa}{16\pi^2} \left\{ \left| \partial_t T^m_{a} b(k, t) \right|^2 - \frac{1}{2} \left| \partial_t T^m_{a b}(k, t) \right|^2 \right\}.
\] (24.209)

as follows directly by inserting (24.93) and the momentum-space version of the energy-momentum conservation law (24.170): \( \hat{k}_i T^{ia}(k, \omega) = \hat{T}_{0a}(k, \omega) \), by analogy with the simpler treatment of current terms (24.41).

For long wavelengths, the spatial components of the energy-momentum tensor have a negligible dependence on \( k \): \( m T_{ij}(k, t) \approx m T_{ij}(k = 0, t) = \int d^3x \hat{T}_{ij}(x, t) \).

Then we can perform the integral over all angles in (24.208) using the angular average (24.198), yielding

\[
\frac{dE}{dt} = \frac{\kappa}{4\pi^2} \left\{ \frac{1}{5} \left[ \left| \partial_t \int d^3x \hat{T}_{ij}^m(x, t) \right|^2 - \frac{1}{3} \left| \partial_t \int d^3x \hat{T}_{0i}^m(x, t) \right|^2 \right] \right\}.
\] (24.210)

In the quadrupole approximation, Eq. (24.175) tells us that

\[
\int d^3x \left[ \frac{m T_{ij}(x, t)}{2} - \frac{1}{3} \delta_{ij} \hat{T}_{0k}^m(x, t) \right] \approx \frac{1}{2c^2} \ddot{Q}^{ij}(t),
\] (24.211)

and (24.210) becomes

\[
\frac{dE}{dt} = \frac{\kappa}{8\pi^2} \left[ \ddot{Q}^{ij}(t) \right]^2 = \frac{G}{5c^3} \left[ \ddot{Q}^{ij}(t) \right]^2, \quad \frac{G}{c^3} \approx 3.6 \times 10^{52} \text{W},
\] (24.212)

in agreement with (24.199).

For a single nonrelativistic point particle moving along the orbit \( \vec{x}(t) \), the energy-momentum tensor has the spatial components

\[
m T_{ij}(x, t) = \frac{M}{c} \dot{x}^i(t) \dot{x}^j(t) \delta^{(3)}(x - \vec{x}(t)).
\] (24.213)

Then Eq. (24.210) becomes

\[
\frac{dE}{dt} = \frac{\kappa}{4\pi^2} \left[ \frac{2M^2}{5} \left| \partial_t \left( \dot{x}^i \dot{x}^j - \frac{1}{3} \delta^{ij} \vec{x}^k \vec{x}^k \right) \right|^2 \right].
\] (24.214)

The reader may wonder about the translational invariance of this formula, since a quadrupole moment possesses a reference point. Consider therefore the second moment of the energy-momentum tensor after a translation

\[
\int d^3x (x_i - a_i)(x_j - a_j) T^{00}(x, t) = \int d^3x x_i x_j T_{00}(x, t)
- a_i \int d^3x x_j T_{00}(x, t) - a_j \int d^3x x_i T_{00}(x, t) + a_i a_j \int d^3x T_{00}(x, t). \quad (24.215)
\]
We now observe that the last term is time independent because of energy conservation. The other two terms in the second line may be rewritten using the conservation law $\partial_a T^{ab}(x, t)$ and a partial integration as follows:

$$\partial_0^2 \int d^3x \int d^3x \tilde{T}^{00}(x, t) = -\partial_0 \int d^3x \int d^3x \tilde{T}^{00}(x, t) = \int d^3x \int d^3x \tilde{T}^{jk}(x, t)$$

the last zero following from momentum conservation. Thus $Q_{ij}(t)$ changes at most by a linear function of $t$, and formula (24.212) is indeed independent of the choice of the reference point for calculating the quadrupole moment.

For checking the dimensions of the above equations the following list of dimensions is useful:

$$[A] = \hbar, \quad [T^{ab}] = \hbar \text{ cm}^4, \quad [E] = \hbar \text{ sec}, \quad [p] = \hbar \text{ cm}, \quad [Q^{ij}] = \hbar \text{ sec}, \quad [\dddot{Q}^{ij}] = \hbar \text{ sec}^2,$$

$$[\kappa] = \text{cm}^2/\hbar, \quad [G] = \text{cm}^5/\hbar \text{ sec}^5, \quad [c^5/G] = \frac{\hbar}{\text{sec}^2} \approx 3.6 \times 10^{52} \text{ W}. \quad (24.217)$$

Note that $\dddot{Q}^{ij}$ has the dimension of a power.

Together with the energy, also the total angular momentum

$$L_i(t) = \epsilon_{ijk} \int d^3x x^j \tilde{T}^{0k}(x, t)$$

of the gravitational system decreases. This happens at a rate

$$\frac{dL_k}{dt} = \frac{2\kappa}{8\pi} \frac{1}{5c^2} \epsilon_{klm} \dddot{Q}^{ij}(t) Q^{lm}(t) = \frac{2G}{5c^5} \epsilon_{klm} \dddot{Q}^{ij}(t) \dddot{Q}^{lm}(t). \quad (24.219)$$

### 24.9 Simple Models for Sources of Gravitational Radiation

Let us calculate the radiated power for a few typical radiating systems. One distinguishes oscillating and bursting systems.

#### 24.9.1 Vibrating Quadrupole

Imagine two equal masses $M$ oscillating at the ends of a spring (see Fig. 24.3). The masses have the time-dependent positive

$$z = \pm \left( \frac{d}{2} + a \sin \omega t \right). \quad (24.220)$$

Assuming $a \ll d$, we may approximate

$$z^2 \sim \frac{d^2}{4} + ad \sin \omega t. \quad (24.221)$$

---

The quadrupole moment is therefore

\[ Q_{ij}(t) = \left( 1 + \frac{4a}{d} \sin \omega t \right) q_{ij} \]  

(24.222)

where

\[ Q_{ij}(0) = \frac{Md^2}{6} \begin{pmatrix} -1 & -1 \\ -1 & 2 \end{pmatrix}. \]  

(24.223)

and

\[ \ddot{Q}_{11}(t) = \ddot{Q}_{22}(t) = \frac{1}{2} \ddot{Q}_{22}(t) = -\frac{2M\omega^3}{3} d \sin \omega t. \]  

(24.224)

According to Eq. (24.194), the angular distribution for a diagonal quadrupole moment is

\[ \frac{d^2E}{dtd\Omega} = \frac{1}{4(4\pi)^2 c^2} \left\{ \dddot{Q}_{11}^2 + \dddot{Q}_{22}^2 + \dddot{Q}_{33}^2 - 2 \left[ (\dddot{Q}_{11}n_1)^2 + (\dddot{Q}_{22}n_2)^2 + (\dddot{Q}_{33}n_3)^2 \right] \right\} \]  

(24.225)

Introducing spherical angles

\[ n_1 = \sin \theta \cos \phi, \quad n_2 = \sin \theta \sin \phi, \quad n_3 = \cos \theta. \]  

(24.226)

as indicated in Fig. 24.3, the curly brackets in (24.225) become for \( Q_{11} = Q_{22} \)

\[ \left\{ 2 \dddot{Q}_{11}^2 + \dddot{Q}_{33}^2 - 2(\dddot{Q}_{11}^2 \sin^2 \theta + \dddot{Q}_{33}^2 \cos^2 \theta) + \frac{1}{2} (\dddot{Q}_{11}^2 \sin^4 \theta + \dddot{Q}_{33}^2 \cos^4 \theta) \right\}. \]  

(24.227)

For \( Q_{33} = -2Q_{11} \), this reduces to

\[ \frac{\dddot{Q}_{11}}{2} \frac{9}{2} \sin^4 \theta. \]  

(24.228)
The rate of energy radiation per solid angle \(d\Omega\) is then
\[
\frac{d^2E}{dtd\Omega} = \frac{1}{2c^2} \kappa \frac{M^2}{(4\pi)^2} \omega^3 a \sin \omega t)^2. \tag{24.229}
\]
The radiation is maximal in the direction of the equation and vanishes in the pole directions of oscillator.

Integrating (24.229) over all angles give the total omitted power
\[
\frac{dE}{dt} = \frac{1}{2c^2} \kappa \frac{M^2}{(4\pi)^2} \omega^3 a \sin \omega t)^2 \frac{32}{15\pi} = \frac{8G}{15c^3} M^2 \omega^3 a \sin \omega t)^2 \tag{24.230}
\]
whose temporal average is
\[
\frac{dE}{dt} = \frac{8G}{15c^3} M^2 \omega^6 a^2 c^2. \tag{24.231}
\]
The rate of radiation damping is defined as
\[
\gamma_{\text{rad}} \equiv \frac{1}{t_{\text{rad}}} \equiv \frac{1}{E \frac{dE}{dt}}, \tag{24.232}
\]
where \(t_{\text{rad}}\) is the damping time. Since the kinetic energy of each mass is \(M\omega^2 a^2/2\), we obtain
\[
\gamma_{\text{rad}} = \frac{46}{15c^5} Md^2 \omega^4. \tag{24.233}
\]
The formula estimates the damping rate of any linearly oscillating system of two masses. The linear character of the oscillation is important since for a spherically symmetric pulsating star, there is no gravitational radiation at all.

Vibrational radiation may emerge from nova exploitations at an early stage. These arise if a star circles around a white dwarf and transfers matter to him. After some time, the matter becomes large enough to explode.

This explosion causes vibrations in the white dwarf with frequencies 0.01 to 1 Hz. The energy released in a nova explosion is typically \(10^{45}\) erg, of which 10% could be deposited in vibrations, which send out gravitational radiation.

### 24.9.2 Two Rotating Masses

If the two masses in the previous example rotate around the \(z\)-axis (see Fig. 24.4) the quadrupole moment (24.223) in the \(xy\)-plane becomes
\[
Q_{ij}(t) = \frac{Md^2}{4} \begin{pmatrix}
1 - 3 \cos 2\omega t & -3 \sin 2\omega t \\
-3 \sin 2\omega & 1 + 3 \cos 2\omega t
\end{pmatrix}, \tag{24.234}
\]
where \(M\) is the reduced mass
\[
M \equiv \frac{M_1 M_2}{M_1 + M_2}, \tag{24.235}
\]
\[\text{H. Kleinert, GRAVITY WITH TORSION}\]
and $\omega$ is given by the third Kepler law:

$$\omega = \sqrt{\frac{G(M_1 + M_2)}{r^3}}. \quad (24.236)$$

The third time derivatives of the quadrupole moments are therefore

$$\ddot{Q}_{ij}(t) = -6 Md^2 \omega^3 \begin{pmatrix} -\sin 2\omega t & -\cos 2\omega t \\ 0 & 0 \\ -\cos 2\omega t & \sin 2\omega t \end{pmatrix}. \quad (24.237)$$

Inserting these into (24.225), integrating over all angles, and averaging over all times yields the total omitted power

$$\frac{dE}{dt} = \frac{8G^5}{5c^5} M^2 d^4 \omega^6. \quad (24.238)$$

Inserting $\omega$ from (24.236), this becomes

$$\frac{dE}{dt} = \frac{32G^4}{5c^5 d^5} (M_1 M_2)^2 (M_1 + M_2). \quad (24.239)$$

The total energy of the binary system is

$$E = \frac{1}{2} \frac{M_1 M_2}{M_1 + M_2} d^2 \omega^2 = \frac{1}{2} \frac{GM_1 M_2}{d}, \quad (24.240)$$

implying a rate of radiation loss is

$$\gamma = \frac{64G^4}{5c^5 d^5} M_1 M_2 (M_1 + M_2). \quad (24.241)$$

Since $E$ is inverse proportional to the distance $d$ by (24.240), the distance between the masses decreases as follows:

$$\frac{\dot{d}}{d} = -\frac{64G^3}{5c^5 d^4} M_1 M_2 (M_1 + M_2). \quad (24.242)$$

By Eq. (24.236) this implies that the frequency increases at a rate

$$\frac{\dot{\omega}}{\omega} = -\frac{3}{2} \frac{\dot{d}}{d}. \quad (24.243)$$
Due to the smallness of the gravitational constant, the power radiated by planetary systems is extremely small. The earth orbiting around the sun emits only 200 W, the Jupiter 5300 W. For narrow double stars, the power can increase to $10^{30}$ W and more, for doubleneutron stars for which $d$ can be quite small, it can reach to $10^{45}$ W.

Table 24.1 shows various astronomical objects and the gravitational amplitudes which can arrive from them at the earth.

**Table 24.1** Binary systems as sources of gravitational radiation (from D. H. Douglass and V. B. Braginsky *Gravitational-radiation Experiments*, in S. W. Hawking and W. Israel *General Relativity* (Cambridge University Press, Cambridge, 1979). The binary PSR 1913+16 emits radiation at multiples of $70 \times 10^{-6}$ Hz due to the large eccentricity of the orbit.

<table>
<thead>
<tr>
<th>System</th>
<th>Masses $(M_\odot)$</th>
<th>Distance (pc)</th>
<th>Wave freq. $(10^{-6} \text{Hz})$</th>
<th>Luminosity at earth $(10^{30} \text{erg/s})$</th>
<th>Flux at earth $(10^{-22})$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eclipsing binaries</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>τ Boo</td>
<td>1.0, 0.5</td>
<td>11.7</td>
<td>86</td>
<td>1.1</td>
<td>68.0</td>
</tr>
<tr>
<td>μ Sco</td>
<td>12, 12</td>
<td>109</td>
<td>16</td>
<td>51</td>
<td>38.0</td>
</tr>
<tr>
<td>V Pup</td>
<td>16.5, 9.7</td>
<td>520</td>
<td>16</td>
<td>59</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Cataclysmic binaries (novas)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AM CVn</td>
<td>1.0, 0.041</td>
<td>100</td>
<td>1900</td>
<td>300</td>
<td>240</td>
</tr>
<tr>
<td>WZ Sge</td>
<td>1.5, 0.12</td>
<td>75</td>
<td>410</td>
<td>24</td>
<td>37</td>
</tr>
<tr>
<td>SS Cyg</td>
<td>0.97, 0.83</td>
<td>30</td>
<td>84</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td><strong>Binary X-ray sources (black holes or neutron stars)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyg $X_1$</td>
<td>30,6</td>
<td>2500</td>
<td>4.1</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>PSR 1913+16</td>
<td>1.4, 1.4</td>
<td>5000</td>
<td>70</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>140</td>
<td>2.9</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>210</td>
<td>5.8</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Formula (24.243) has been used as an indirect evidence for the existence of gravitational radiation. In 1974, Hulse and Tayler searched for pulsars (rotating neutron stars emitting radio pulses) with the Arecibo telescope and found an object whose emitted radio frequency is periodically modulated. The modulation is attributed to the Doppler shift caused by the orbital motion around an undetected companion. A careful analysis of the modulation allowed to derive the eccentricity and the rate of the perihelion precession of the binary object (see Table 24.2 for details). A particularly interesting effect was the slow-down of the orbital motion. If $\tau$ denotes the period, one finds

\[
\dot{\tau} \approx -2.4184(9) \times 10^{12} \text{ sec per sec.}
\]  

(24.244)

The observed shift of the time of periastron passage is plotted in Fig. 24.7. The values of the masses were deduced from the perihel precession and the time delay of...
Table 24.2 Some observed parameters of PSR 1913+16 (Table taken from Ref. [1]).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>21,000 ly</td>
</tr>
<tr>
<td>Pulsar period (nominal)</td>
<td>59.02999792988 ms</td>
</tr>
<tr>
<td>Semi-major axis</td>
<td>1,950,100 km</td>
</tr>
<tr>
<td>Eccentricity</td>
<td>0.617131 ± 0.000003</td>
</tr>
<tr>
<td>Orbital period</td>
<td>27,907 ± 0.00002 s</td>
</tr>
<tr>
<td>Rate of precession of periastron</td>
<td>(4.22263 ± 0.0003)°/y</td>
</tr>
<tr>
<td>Amplitude of time-dilation factor</td>
<td>0.0044 ± 0.0001</td>
</tr>
<tr>
<td>Rate of decrease of orbital period</td>
<td>2.4184(9) × 10^{-12} s/s = 0.0000765 s/y</td>
</tr>
<tr>
<td>Rate of decrease of semimajor axis</td>
<td>3.5 m/y</td>
</tr>
<tr>
<td>Calculated lifetime (to final inspiral)</td>
<td>300,000,000 y</td>
</tr>
<tr>
<td>Diameter of each neutron star</td>
<td>20 km</td>
</tr>
<tr>
<td>Periastron separation</td>
<td>746,600 km</td>
</tr>
<tr>
<td>Apastron separation</td>
<td>3,153,600 km</td>
</tr>
<tr>
<td>Velocity of stars at periastron in CM frame</td>
<td>450 km/sec</td>
</tr>
<tr>
<td>Velocity of stars at apstro in CM frame</td>
<td>110 km/sec</td>
</tr>
<tr>
<td>Rate of precession of spin axis</td>
<td>?</td>
</tr>
</tbody>
</table>

Figure 24.5 Gravitational Amplitudes arriving at the earth from possible sources. LIGO denotes the Laser Interferometer Gravitational Wave Observatory (see http://www.ligo.caltech.edu). LISA is the Laser Interferometer Space Antenna, a joint three-spacecraft mission of ESA and NASA http://www.esa.int/esaSC/120376_index_0_m.html, where the time-dependence of the distance of two objects $5 \times 10^6$ km apart will be monitored. WDB denotes the regime of white dwarf binaries.

the signal passing the companion. The first depends only on $M_1 + M_2$, the $M_1$ and $M_2$ in a different combination. From these data, one deduces

$$M_1 \sim (1.4414 \pm 0.0002) M_\odot \quad \text{for pulsar}, \quad (24.245)$$

$$M_2 \sim (1.3867 \pm 0.0002) M_\odot \quad \text{for companion}. \quad (24.246)$$
Using these and the projected semi-major axis in Table 24.2 one finds a rate of change of the period
\[ \dot{\tau}_{\text{theor}} \sim -2.38 \times 10^{-12} \text{ sec per sec} \] (24.247)
in good agreement with the observed rate in Table 24.2.

The properties of binary objects containing a pulsar can be studied so well that the approximation (24.239) for the radiated power need several corrections [2]. Consider two masses Figure 24.8 orbiting around the common center-of-mass. If \( d \) denotes the distance of the two masses, the distances from the center-of-mass are \( d_1 = M_2d/(M_1 + M_2) \), \( d_2 = M_1d/(M_1 + M_2) \). Denoting the reduced mass \( M_1M_2/(M_1 + M_2) \) as before by \( M \), the components of the quadrupole moment are

\[ Q_{ij}(t) = Md^2 \begin{pmatrix} \cos^2 \varphi & \sin \varphi \cos \varphi \\ \sin \varphi \cos \varphi & \sin^2 \varphi \end{pmatrix}. \] (24.248)

The orbit of a Kepler ellipse has the general form
\[ d = \frac{a(1 - e^2)}{1 + e \cos \varphi}, \] (24.249)
where \( a \) is the semi-major axis of the ellipse, \( e \equiv \sqrt{1 - b^2/a^2} \) is the eccentricity (\( b \) denotes the semi-minor axis), and the angular velocity is given by
\[ \dot{\varphi}(t) = d^{-2}(t)\sqrt{(M_1 + M_2)a(1 - e^2)} \] (24.250)
From this we derive immediately

\[ \ddot{Q}_{11} = P(1 + e \cos \varphi)^2(2 \sin 2 \varphi + 3e \sin \varphi \cos^2 \varphi) \] (24.251)
\[ \ddot{Q}_{22} = P(1 + e \cos \varphi)^2[2 \sin 2 \varphi + e \sin \varphi(1 + 3 \cos^2 \varphi)] \] (24.252)
\[ \ddot{Q}_{12} = -P(1 + e \cos \varphi)^2[2 \sin 2 \varphi - e \sin \varphi(1 - 3 \cos^2 \varphi)] = \ddot{Q}_{21}, \] (24.253)

**Figure 24.6** Two pulsars orbiting around each other.
24.9 Simple Models for Sources of Gravitational Radiation

Figure 24.7 Shift of time of periastron passage of PSR 1913+16 for each orbit caused by the shrinking of the Kepler orbits as a consequence of formula (24.256). Curve is from theory (see Ref. [1]).

Figure 24.8 Two masses in a Keplerian orbit around the common center-of-mass.

where $P$ is a power factor

$$P \equiv 2{\sqrt{G^3 M_1^2 M_2^2 (M_1 + M_2)} \over a^5(1 - e^2)^5} \quad (24.254)$$

Inserting (24.251)–(24.253) into (24.225), integrating over all angles, and averaging over all times yields the total omitted power

$$dE \over dt = {8G^4 M_1^2 M_2^2 (M_1 + M_2) \over 15c^5} {1 + e \cos \varphi(t)} \over a^5(1 - e^2)^5} \{12[1 + e \cos \varphi(t)]^2 + e^2 \sin^2 \varphi(t)\},$$

where $t$ on the right-hand side is retarded with respect to the left-hand side. For circular orbits where $e = 0$, $a = d/2$, this reduces properly to (24.239). Averaging over one period of the elliptical orbit yields

$$\left<{dE \over dt}\right> = {32G^4 M_1^2 M_2^2 (M_1 + M_2) \over 5c^5} {1 + e^2 \over a^5(1 - e^2)^7/2} \left(1 + {73 \over 24} e^2 + {37 \over 96} e^4\right),$$

(24.256)
With respect to a circular orbit of equal total energy, the power is enhanced by a factor
\[ f = 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \]
(24.257), which grows rapidly from \( f(e) = 1 \) at \( e = 0 \) to infinity at \( e = 1 \). For a full multipole analysis of the radiation see [2].

Due to the shrinking of the Kepler orbits implied by formula (24.256), the orbital time shrinks according to the elliptic generalization of Eq. (24.243), implying that the periastron is reached a few seconds earlier for each orbit. The time shift is plotted in Fig. 24.7.

The radiation properties of binary objects can be studied especially well if both stars are pulsars. Such an astronomical object has recently been found. One of the two pulsars rotates with a period of 23 milliseconds (PSR J0737-3039A) around its axis, the other with a period of 2.8 seconds (PSR J0737-3039B) [3] (see Fig. 24.6).

### 24.9.3 Particle Falling into Star

Among the bursting sources of gravitational radiation, the simplest one consists of a mass falling into a star as shown in Fig. 24.9. If the mass starts at \( z = \infty \), its velocity is
\[ \frac{1}{2} m \dot{z}^2 = \frac{G m M}{z}, \]
so that
\[ \dot{z} = -\frac{1}{z^{1/2}} (2GM)^{1/2}, \quad \ddot{z} = -\frac{G M}{z^2}, \quad \dot{\ddot{z}} = -\frac{(2GM)^{3/2}}{z^{7/2}}. \]
(24.259)

The triple time derivative of the quadrupole moment
\[ Q_{ij} = m \begin{pmatrix} -z^2 & 0 & 0 \\ 0 & -z^2 & 0 \\ 0 & 0 & 2z^2 \end{pmatrix} \]
(24.260)
is
\[ \dot{Q}_{ij} = m (6 \ddot{z} + 2z \dot{z}) \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \]
(24.261)
and thus, because of (24.259):
\[ \ddot{Q}_{ij} = m \frac{(2GM)^{3/2}}{z^{5/2}} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \]
(24.262)
Inserted into (24.199), this leads to the energy loss per second
\[ \frac{dE}{dt} = \frac{2G m^2}{15 c^5} (6 \ddot{z} + 2z \dot{z})^2 \]
(24.263)
From Eqs. (24.263) and (24.259) we deduce

$$\frac{dE}{dz} = \frac{1}{z^{9/2}} \frac{2Gm^2}{15c^2} (2GM)^{5/2}/dz.$$  \hfill(24.264)

The radiated energy from \( z = \infty \) to \( z = R \) is

$$\Delta E = \frac{1}{R^{7/2}} \frac{4Gm^2}{105c^5} (2GM)^{5/2}.$$  \hfill(24.265)

Obviously, the radiated energy increases with decreasing \( R \). Suppose the large object is a black hole. If we let the mass \( m \) fall down to the Schwarzschild radius \( R_s \)

$$R_s = \frac{2GM}{c^2},$$  \hfill(24.266)

we obtain

$$\Delta E = \frac{2}{105} mc^2 \frac{m}{M} \approx 0.019 mc^2 \frac{m}{M}.$$  \hfill(24.267)

If one takes relativistic effects into account, as the particle approaches \( R_s \), the number 0.019 changes to 0.0104\(^4\).

For a black hole of radius \( M \sim 10M_\odot \) with Schwarzschild \( R_s \sim 30km \), the total radiated energy is

$$\Delta E \sim 2 \times 10^{51}\text{erg}.$$  \hfill(24.268)

The radiated energy is mostly emitted at the end (see Fig. 24.9 for the curve resulting from a detailed analysis).

**Table 24.3 Typical Astrophysical Sources of Gravitational Radiation.** Distances have been selected large enough to yield approximately three events per year (from K. S. Thorne, *Gravitational Radiation*, in S. W. Hawking, and W. Israel, eds., *Three Hundred Years of Gravitation*, (Cambridge University Press, Cambridge, 1987).).

<table>
<thead>
<tr>
<th>Source</th>
<th>Frequency</th>
<th>Distance</th>
<th>Amplitude ((\kappa A))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Periodic sources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binaries</td>
<td>(10^{-4}) Hz</td>
<td>10 pc</td>
<td>(10^{-20})</td>
</tr>
<tr>
<td>Nova</td>
<td>(10^{-2}) to 1</td>
<td>500 pc</td>
<td>(10^{-22})</td>
</tr>
<tr>
<td>Spinning neutron star (Crab)</td>
<td>60</td>
<td>2 kpc</td>
<td>(&lt;10^{-24})</td>
</tr>
<tr>
<td><strong>Bursting sources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coalescence of binary</td>
<td>10 to (10^3)</td>
<td>100 MPC</td>
<td>(10^{-21})</td>
</tr>
<tr>
<td>Infall of star into 10(M_\odot) b.h.</td>
<td>(10^{-4})</td>
<td>10 Mpc</td>
<td>(10^{-21})</td>
</tr>
<tr>
<td>Supernova</td>
<td>(10^3)</td>
<td>10 kpc</td>
<td>(10^{-18})</td>
</tr>
<tr>
<td>Gravitational collapse of (10^4M_\odot) star</td>
<td>(10^{-1})</td>
<td>3 Gpc</td>
<td>(10^{-19})</td>
</tr>
</tbody>
</table>

Figure 24.9 Spectrum of the gravitational radiation emitted by a particle of mass $m$ falling radially into a black hole of mass $M$. The quantity $dE/d\omega$ gives the amount of energy radiated per unit frequency interval. The curve marked $l = 2$ corresponds to quadrupole radiation; the other curves ($l = 3, l = 4$) correspond to multipole radiation of higher order. Note that most of the radiation is emitted with frequencies below $\omega \approx 0.5c^3/GM$ [4].

Figure 24.10 Particle falling radially toward a mass.
24.9.4 Gas of Colliding Stars

A gas of stars treated as point-like objects has an energy-momentum tensor

\[ T^{ab} = \sum_n \frac{p_n^a p_n^b}{E_n^3} \delta^{(3)}(x - x(t)). \]  

(24.269)

For a gas with collisions at \( t = 0 \) where the stars change their velocities from \( v_n \) to \( \bar{v}_n \), the energy-momentum tensor is

\[ T^{ab}(x, t) = \sum_n \frac{p_n^a p_n^b}{E_n^3} \delta^{(3)}(x - v_n t) \Theta(-t) - \sum_n \frac{\bar{p}_n^a \bar{p}_n^b}{E_n^3} \delta^{(3)}(x - \bar{v}_n t) \Theta(t), \]  

(24.270)

where \( \Theta(t) \) is the Heaviside function defined in Eq. (22.148). Representing this as a Fourier integral

\[ \Theta(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{i\omega t}}{\omega - i\eta}, \]  

(24.271)

and the \( \delta \)-function as

\[ \delta^{(3)}(x - v) = \int \frac{d^3k}{(2\pi)^3} e^{i k x}, \]  

(24.272)

we see that

\[ T^{ab}(k, \omega) = -\frac{1}{(2\pi)^4} \sum_n \left( \frac{p_n^a p_n^b}{E_n} \frac{i}{\omega - v_n \cdot k - i\eta} - \frac{\bar{p}_n^a \bar{p}_n^b}{E_n} \frac{i}{\omega - \bar{v}_n \cdot k + i\eta} \right). \]  

(24.273)

Now we observe that

\[ E_n (\omega - v_n \cdot k) = p_n k, \]  

(24.274)

where \( k \) is the four-momentum of the emerging gravitational wave. Since the stars are nonrelativistic objects, we can further drop the \( i\eta \)'s. Thus

\[ T^{ab}(k, \omega) = -\frac{i}{(2\pi)^4} \sum_n \left( \frac{p_n^a p_n^b}{p_n k} - \frac{\bar{p}_n^a \bar{p}_n^b}{\bar{p}_n k} \right). \]  

(24.275)

Observe that \( k_a T^{ab} = \sum_n (p_n - \bar{p}_n) = 0 \), as it should by momentum conservation. Inserting this into (24.209) yields

\[ \frac{dE}{d\Omega} = \frac{\kappa}{8\pi^2} \int \frac{d\omega}{2\pi} \omega^2 \sum_{NN'} \frac{\sigma_N \sigma_{N'}}{(p_N k)(p_{N'} k)} \left[ (p_N p_{N'})^2 - \frac{1}{2} M_N^2 M_{N'}^2 \right]. \]  

(24.276)

where the index \( N \) runs over the particles before and after the collisions, with \( \sigma_N \) being \( \pm 1 \) for the two cases.

The frequency integral diverges linearly in \( \omega \). This is due to the fact that we have assumed an instantaneous change of momenta during the collisions. In actual collisions, the change takes place over some finite collision time \( \Delta t \) and the integral only contain frequencies up to \( 1/\Delta t \).

For nonrelativistic two-body scattering, the radiated energy is

\[ \frac{dE}{d\Omega} = \frac{2\kappa}{5\pi} \mu^2 \int \frac{d\omega}{2\pi} v^4 \sin^2 \theta, \]  

(24.277)

where \( \mu \) is the reduced mass. \( v \) the relative velocity, and \( \theta \) the scattering angle in the center-of-mass frame.
24.10 Orders of Magnitude of Different Radiation Sources

In order to have an idea as to how much energy can be radiated in various precesses, consider a massive steel rod of radius 1m, length d=20m, mass \( M \approx 4.9 \times 10^8 g \) (=490 tons, using the steel density 7.89g/cm\(^3\)). If the maximal quadropole radiation can be obtained by rotating it around an axis orthogonal to the rod with an angular velocity \( \omega \). The formula (24.199) yields the total emitted power

\[
\frac{dE}{dt} = \frac{2}{45c^3} M^2 l^4 \omega^6. \tag{24.278}
\]

The angular velocity is limited by the tensile strength, which is \( t \approx 3 \times 10^9 \) dyn/cm\(^2\). Thus one can maximally use an angular velocity

\[
\omega = \left( \frac{8t}{\rho l^2} \right)^{1/2} \approx 28 \frac{1}{\text{sec}}. \tag{24.279}
\]

This gives a radiated power

\[
\frac{dE}{dt} \approx 10^{-23} \text{erg/sec}. \tag{24.280}
\]

In order to have an idea how small this is we note that a single photon in the visible range has an energy

\[
hc \frac{2\pi}{4000A} \approx 10^{-27} 10^{10} \frac{2\pi}{4 \times 10^{-5}} \approx 1.5 \times 10^{-12} \text{erg}. \tag{24.281}
\]

Thus the radiated power corresponds to the emission of one visible photon in \( \approx 10^9 \) seconds or \( \approx 3000 \) years.

If one wants to have any observable effects it is therefore necessary to look for radiation emitted by large stellar objects. Consider a bunch of stars of total mass \( M \) distributed over a region of size \( R \). their velocity is of the order \( R/T \) where \( T \) is the time it takes for the masses to move from our side to the other. Their quadropole moment is of the order of

\[
Q \sim R^2 M. \tag{24.282}
\]

Hence we can estimate

\[
\ddot{Q} \sim \frac{R^2 M}{T^3} \sim M \left( \frac{R}{T} \right)^2 \frac{1}{T}. \tag{24.283}
\]

The right-hand side has the dimension energy per time. It gives an estimate for the internal power flow in the system. The radiated power is equal to

\[
\frac{dE}{dt} \sim \frac{1}{c^5/G} \left( \frac{R^2 M}{T^3} \right)^2. \tag{24.284}
\]
The denominator term $\frac{c^5}{G}$ is itself a power [compare (24.217)]

$$\frac{c^5}{G} \approx 3.63 \times 10^{50} \text{erg sec}.$$  \hspace{1cm} (24.285)

It is called the Planck power. Thus, radiated power is of the order of the internal power square divided by the Planck power $\frac{c^5}{G}$.

Note that if a system is to experience sizable radiation damping, it has to have internal power flow of the order of $\frac{c^5}{G}$. This is an immense power. It corresponds to a kinetic energy which is generated by burning up $2 \times 10^5$ solar masses per second via nuclear precesses.

A sizable radiation can reach us only from stellar catastrophies. A mass $m$ falling into a black hole of mass $M$ radiates off a total energy

$$E \approx 0.0104 \frac{m}{M} mc^2.$$  \hspace{1cm} (24.286)

As an example, consider a star with solar mass falling into a black hole whose mass is 10 times the solar mass. This gives

$$E \sim -2 \times 10^{51} \text{erg}.$$  \hspace{1cm} (24.287)

This energy is radiated during a time [5]

$$t \sim 10^{-3} \text{sec}.$$  \hspace{1cm} (24.288)

Hence the radiated power is of the order

$$\frac{dE}{dt} \approx \frac{m^2 c^5}{M^2 G}.$$  \hspace{1cm} (24.289)

The most likely sources of detectable gravitational radiation are supernova explosions. They have been estimated to emit up to $\sim 10^{54} \text{erg/sec}$.

The youngest known pulsar of the Crab Nebulue is believed to be a neutron star in fast rotation. Its gravitational radiation has been estimated to be of the order of $10^{38} \text{erg/sec}$, if it is deformed by 0.001 from a spherical shape. This radiation would be necessary to explain the observed slowdown of the rotation frequency by gravitational radiation damping. It is not clear, however, whether a neutron star can be deformed to such an extent.

Compare this with the radiation emitted from binary stars. For a circular orbit their kinetic energy is

$$E_{\text{kin}} = \frac{1}{2} \frac{M_1 M_2}{R}.$$  \hspace{1cm} (24.290)

The power output is

$$\frac{dE}{dt} = \frac{32 \mu^2 M^3}{5 R^5}.$$  \hspace{1cm} (24.291)
where \( \mu = M_1 m_2 / (M_1 + m_2) \), \( M = M_1 + M_2 \). The energy loss due to gravitational radiation will lead to the two stars spiraling together within a time

\[
 t_{\text{spiral}} = \frac{5}{256} \frac{R^4}{\mu M^2}. \tag{24.292}
\]

For the sun-jupiter system, this time is \( \approx 2.5 \times 10^{23} \) years. But for the binary system PSR 1913+16, the spiral time shrinks to \( 3 \times 10^8 \) years, as shown in Table 24.2, which makes it observable, as we have seen in Fig. 24.7.

This makes it almost impossible to observe this radiation damping of the Kepler orbits.

For neutron stars a few thousand km apart, the spiral times could shrink to the order of years or days, so it could be observable at least, in principle. Here the problem lies in the identification of the source.

### Notes and References


[4] Figure is taken from


Evanescent Properties of Torsion in Gravity

What additional physics is brought about by torsion? If the field action is of the Einstein-Cartan type (19.4), the consequences turn out to be practically unobservable. This remains true if higher powers of the curvature tensor are added. The field equation (17.130) determines the torsion by the field equation:

\[ S_{\mu\nu\lambda} = \frac{\kappa}{2} \left( \sum_{\mu\nu,\lambda} + \frac{1}{2} g_{\nu\lambda} \sum_{\mu\kappa} \cdot \kappa - \frac{1}{2} g_{\nu\lambda} \sum_{\mu\kappa} \cdot \kappa \right). \]  

(25.1)

Let us discuss the effect of this equation for fields of various spins.

25.1 Local Four-Fermion Interaction due to Torsion

A non-trivial effect of torsion can be derived for Dirac fields. The spin density of matter is, from (17.125),

\[ \sum_{\alpha\beta,\gamma} = -i \frac{1}{2} \bar{\psi}[\gamma_{\gamma}, \Sigma_{\alpha\beta}] \psi, \]

with \( \Sigma_{\alpha\beta} = (i/4)[\gamma_{\alpha}, \gamma_{\beta}] \). This can be written as

\[ \sum_{\alpha\beta,\gamma} = \frac{1}{2} \bar{\psi} \gamma_{\alpha\beta\gamma} \gamma_{\gamma} \psi = \frac{1}{2} \varepsilon_{\alpha\beta\gamma\lambda} \bar{\psi} \gamma_{\gamma} \gamma_{\lambda} \psi \]

(25.3)

with \( \gamma_{5} \equiv (1/4!)\varepsilon_{\alpha\beta\gamma\delta} \gamma^{\alpha} \gamma^{\beta} \gamma^{\gamma} \gamma^{\delta} \), where the brackets around the subscripts denote their complete antisymmetrization. Due to antisymmetry, the Palatini tensor divided by 2, the torsion, and the contortion tensor are all equal to \((\kappa/2) \sum_{\alpha\beta,\gamma} \)

\[ \frac{1}{2} S_{\alpha\beta,\gamma} = S_{\alpha\beta,\gamma} = K_{\alpha\beta,\gamma} = \frac{\kappa}{2} \sum_{\alpha\beta,\gamma} \]. \]

(25.4)

In Eq. (11.145) we have expressed the curvature tensor in terms of the Riemann curvature tensor plus the contortion. Two contractions give the corresponding decomposition of the scalar curvature

\[ R = \bar{R} + \bar{D}_{\mu} K_{\nu}^{\mu} - \bar{D}_{\nu} K_{\mu}^{\nu} + \left( K_{\mu}^{\mu\rho} K_{\nu}^{\nu} - K_{\nu}^{\mu\rho} K_{\mu}^{\nu} \right). \]

(25.5)
In the gravitational Einstein-Cartan action, $\bar{R}$ is integrated over the total invariant volume of the universe. The terms $\bar{D}_\mu K^\nu_{\nu\mu}$ and $\bar{D}_\nu K^{\nu\mu}_\mu$ produce irrelevant surface terms and can be ignored. The action can therefore be separated into a Hilbert-Einstein action

$$f_{\mathcal{A}} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \bar{R}, \quad (25.6)$$

plus a field torsion action

$$f_{\mathcal{A}}^S = \int d^4x \sqrt{-g} \bar{f}_S, \quad (25.7)$$

with a Lagrangian density

$$\bar{f}_S = -\frac{1}{2\kappa} \left( K^{\mu}{}_{\mu\rho} K^\rho_{\nu} - K^{\mu}{}_{\mu\rho} K^\rho_{\mu\nu} \right). \quad (25.8)$$

This can be rearranged to

$$\bar{f}_S = \frac{1}{2\kappa} S^\mu_{\mu\nu\lambda} K^{\lambda\nu\mu}, \quad (25.9)$$

where $S^\mu_{\mu\nu\lambda}$ is the Palatini tensor (25.31). As a cross check we differentiate this with respect to $K^{\lambda\nu\mu}$ and obtain

$$\frac{\partial \bar{f}_S}{\partial K^{\lambda\nu\mu}} = \frac{1}{2\kappa} S^\mu_{\mu\nu\lambda}, \quad (25.10)$$

in accordance with (15.17) and (15.54).

We now add to (25.9) the matter-torsion interaction Lagrangian density extracted from the Dirac action (17.15):

$$m_{\mathcal{L}} = \frac{1}{2\kappa} m^\mu_{\mu\nu\lambda} K^{\lambda\nu}, \quad (25.11)$$

Extremizing the combined torsion Lagrangian $\mathcal{L} = f_{\mathcal{A}}^S + m_{\mathcal{L}}$, we recover once more (25.4). Inserting this back into the total Lagrange density gives, at the extremum, the effective torsion Lagrangian

$$\mathcal{L}^{\text{eff}} = \kappa \frac{m^\mu_{\mu\nu\lambda}}{4} K^{\lambda\nu}, \quad (25.12)$$

or explicitly, with (25.3) (see [1])

$$\mathcal{L}^{\text{eff}} = \frac{3\kappa}{16} \bar{\psi}\gamma\gamma\gamma^5\bar{\psi}\gamma^\mu\gamma^\nu\gamma^5\psi. \quad (25.13)$$
Unfortunately, this interaction is too weak to be detectable by present-day experiments, and probably also for many generations to come. The interaction (25.13) will interfere with the weak interactions of nuclear $\beta$-decay which have the Lagrangian

$$\mathcal{L} = -\frac{G}{\sqrt{2}} \left[ \bar{p}\gamma^\lambda \left( 1 - \frac{g_A}{g_V} \gamma_5 \right) n \right] \left[ \bar{e}\gamma^\lambda (1 - \gamma_5) \nu \right] + c.c., \quad (25.14)$$

with the coupling constant

$$G = (1.14730 \pm 0.000641) \times 10^{-5} \text{GeV}^{-2}, \quad (25.15)$$

and the ratio

$$g_A/g_V = 1.255 \pm 0.006. \quad (25.16)$$

In the unifies theory of weak and electromagnetic interactions, the coupling constant (25.15) is given by

$$G \approx \frac{e^2}{m_W^2} \frac{4\pi\alpha}{m_W^2} \quad (25.17)$$

where

$$m_W = 80.423 \pm 0.039 \text{GeV} \quad (25.18)$$

is the mass of the charged vector mesons $W^\pm$. This shows that the torsion interaction (25.13) is smaller than the weak interaction by the immense factor [recall $m_p$ from Eq. (12.44)]

$$\frac{m_W^2}{m_p^2} \approx 4.34 \times 10^{-35}. \quad (25.19)$$

Thus any hope for a detection in the foreseeable future is an illusion.

An additional problem is that a four-fermion interaction such as (25.13) is not renormalizable, so that it cannot possibly be a fundamental interaction, but at best a phenomenological approximation to some more fundamental theory.

### 25.2 Scalar Fields

In Eq. (15.60) we have already noted that, as a consequence of their avanishing spin density, scalar fields do not give rise to torsion. This result contradicts our finding in Section 14.1.2 that classical particle trajectories are coupled to torsion, which makes them autoparallel rather than geodesic. Let us study this problem in more detail by deriving particle trajectories once more from field-theoretic equations.
25.2.1 Possible Cure

How can we remove the discrepancy with respect to the previous derivation of an autoparallel trajectory in Eq. (14.7), where it was found as a consequence of the multivalued mapping procedure? Apparently, the variational procedure of the metric which has led to the Einstein equation (15.61) must be modified to account for the torsion.

One cure is clearly if spin-one-half particles are the only sources of torsion, which is compatible with the present knowledge of the composition of fundamental particles. The torsion is completely antisymmetric so that it decouples from the classical equation of motion. (14.7).

Another solution may ultimately emerge from the following consideration. In the presence of torsion, the particle trajectory does not satisfy $\frac{\delta m}{\delta q^\mu(\tau)} = 0$ but, according to (14.37),

$$
\frac{\delta m}{\delta q^\mu(\tau)} = \frac{\partial L}{\partial q^\mu} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{q}^\mu} = -2S_{\mu\nu} \lambda^\nu \frac{\partial L}{\partial \dot{q}^\lambda}.
$$

(25.20)

For the Lagrangian in the action (11.11), reparametrized with the proper time $\tau = s/c$, the right-hand side becomes

$$
2S_{\mu\nu} \lambda^\nu \frac{\partial L}{\partial \dot{q}^\lambda} = -m 2S_{\mu\nu\lambda} \dot{q}^\nu(\tau) \dot{q}^\lambda(\tau).
$$

(25.21)

The autoparallel equation of motion would obviously be obtained if the energy-momentum tensor of a free spinless point particles would satisfy the covariant conservation law

$$
D^\nu T^{\mu\nu}(x) = 0,
$$

(25.22)

rather than (18.76).

In order to see how such a conservation law could be obtained recall its derivation: We calculated in Eq. (18.55) the change of the total action under Einstein transformations (11.79). For a point particle, this may be written explicitly as

$$
\delta_E A = \int d^4x \frac{\delta A}{\delta g_{\mu\nu}(x)} \delta_E g_{\mu\nu}(x) + \int d\tau \frac{\delta m}{\delta q^\mu(\tau)} \delta_E q^\mu(\tau).
$$

(25.23)

Now the last term does not vanish on autoparallel trajectories. However, if we change $\delta q(\tau)$ into the nonholonomic $\delta q(\tau)$ defined in Eq. (14.14) with the property (14.30), then the second term vanishes.

The vanishing of the first term in (25.23) for arbitrary infinitesimal $\delta_E g_{\mu\nu}(x)$ of Eq. (18.18) has produced the covariant conservation law (18.75) leading to autoparallel trajectories. It is interesting to realize that if we were to replace the Einstein transformation (18.18) in (25.23) by a transformation defined by

$$
\delta_E g_{\mu\nu}(x) = D_\mu \xi_\nu(x) + D_\nu \xi_\mu(x) = \bar{D}_\mu \xi_\nu(x) + \bar{D}_\nu \xi_\mu(x) - 4S^{\lambda}_{\mu\nu} \xi_\lambda(x),
$$

(25.24)
which looks like (11.79), but with the Riemann covariant derivative replaced by the full covariant derivative $D_\mu$, the variation would contain an extra term,
\[ \delta_E g_{\mu\nu}(x) = \delta_E g_{\mu\nu}(x) - 4S_{\mu\nu}^\lambda \xi_\lambda(x). \] (25.25)

The change of the field part of the action would become
\[ \delta_E A = -\frac{1}{2} \int d^4x \sqrt{-g} T^{\mu\nu}(x) \delta_E g_{\mu\nu} = -\int d^4x \sqrt{-g} T^{\mu\nu}(x) D_\nu \xi_\mu(x). \] (25.26)

Integrals over invariant expressions containing the covariant derivative $D_\mu$ can be integrated by parts according to a rule (15.38). If we neglect the surface terms we find
\[ \delta_E A = \int d^4x \sqrt{-g} D_\nu^* T^{\mu\nu}(x) \xi_\mu(x), \] (25.27)

where $D_\nu^* = D_\nu + 2S_{\nu\kappa}^\lambda$. From this Einstein transformation we would thus find the covariant conservation law of the energy-momentum tensor of a spinless point particle in a space with curvature and torsion $D_\nu^* T^{\mu\nu}(x) = 0$, which is precisely the law (25.22) corresponding to autoparallel trajectories.

The question arises whether the new conservation law (25.22) allows for the construction of an extension of Einstein’s field equation
\[ \bar{G}^{\mu\nu} = \kappa T^{\mu\nu} \] (25.28)
to spaces with torsion, where $\bar{G}^{\mu\nu}$ is the Einstein tensor formed from the Ricci tensor $\bar{R}_{\mu\nu} \equiv \bar{R}_{\lambda\mu\nu}^\lambda$ in Riemann spacetime (11.144). The minimal extension of (25.28) to spacetimes with torsion replaces the left-hand side by the Einstein-Cartan tensor $G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} \bar{g}_{\mu\nu} R^\tau \tau$ and the right-hand side by the canonical energy-momentum tensor $\Theta^{\mu\nu}$, and becomes
\[ G^{\mu\nu} = \kappa \Theta^{\mu\nu}. \] (25.29)

The Einstein-Cartan tensor $G^{\mu\nu}$ satisfies a Bianchi identity
\[ D_\nu^* G^{\mu\nu} + 2S_{\lambda\mu}^\kappa G^{\lambda\kappa} - \frac{1}{2} S_{\kappa}^{\lambda;\nu} R^{\mu\lambda}_{\nu\kappa} = 0, \] (25.30)
where $S^{\lambda;\nu}_\kappa$ is the Palatini tensor (15.45),
\[ S^{\lambda;\nu}_\kappa \equiv 2(S^{\lambda;\nu}_\kappa + \delta^{\nu}_\lambda S^{\kappa;\tau}_\kappa - \delta^{\kappa;\tau}_\nu S^{\lambda;\kappa}_\tau). \] (25.31)

It is then concluded that the energy-momentum tensor satisfies the conservation law
\[ D_\nu^* \Theta^{\mu\nu} + 2S_{\lambda\mu}^\kappa \Theta^{\lambda\kappa} - \frac{1}{2\kappa} S_{\kappa}^{\lambda;\nu} R^{\mu\lambda}_{\nu\kappa} = 0. \] (25.32)

For standard field theories of matter, this is indeed true if the Palatini tensor satisfies the second Einstein-Cartan field equation
\[ S^{\lambda;\nu}_\kappa = 0, \] (25.33)
where $\Sigma^{\lambda \kappa; \mu}$ is the canonical spin density of the matter fields. A spinless point particle contributes only to the first two terms in (25.32).

What tensor will stand on the left-hand side of the field equation (25.29) if the energy-momentum tensor satisfies the conservation law (25.22) instead of (18.76)? At present, we can give an answer [3] only for the case of a pure gradient torsion which has the general form [4]

$$S_{\mu \nu} = \frac{1}{2} (\delta_{\mu}^{\lambda} \partial_{\nu} \sigma - \delta_{\nu}^{\lambda} \partial_{\mu} \sigma).$$

(25.34)

Then we may simply replace (25.29) by

$$e^\sigma G_{\mu \nu} = \kappa T_{\mu \nu}.$$  

(25.35)

Note that for gradient torsion, $G^{\mu \nu}$ is symmetric as can be deduced from the fundamental identity (which expresses merely the fact that the Einstein-Cartan tensor $R_{\mu \nu \lambda}^{\kappa}$ is the covariant curl of the affine connection)

$$D^* S_{\mu \nu} = G_{\mu \nu} - G_{\nu \mu}.$$  

(25.36)

Indeed, inserting (25.34) into (25.31), we find the Palatini tensor

$$S_{\mu \lambda} = -2[\delta_{\mu}^{\kappa} \partial_{\lambda} \sigma - (\lambda \leftrightarrow \mu)].$$  

(25.37)

This has a vanishing covariant derivative

$$D^* S_{\mu \nu} = -2[D^* \mu \partial_{\nu} \sigma - D^* \nu \partial_{\mu} \sigma] = 2[S_{\mu \nu} \lambda \partial_{\lambda} \sigma - 2S_{\mu \lambda} \lambda \partial_{\nu} \sigma + 2S_{\nu \lambda} \lambda \partial_{\mu} \sigma],$$

(25.38)

since the terms on the right-hand side cancel after using (25.34) and $S_{\mu \lambda} \equiv S_{\mu} = -3\partial_{\mu} \sigma/2$. Now we insert (25.34) into the Bianchi identity (25.30), with the result

$$D^* G_{\lambda \nu} + \partial_{\lambda} \sigma G_{\kappa}^{\kappa} - \partial_{\nu} \sigma G_{\lambda}^{\nu} + 2\partial_{\nu} \sigma R_{\lambda}^{\nu} = 0.$$  

(25.39)

Inserting here $R_{\lambda \kappa} = G_{\lambda \kappa} - \frac{1}{2} g_{\lambda \kappa} G_{\nu \nu}$, this becomes

$$D^* G_{\lambda}^{\nu} + \partial_{\nu} \sigma G_{\lambda}^{\nu} = 0.$$  

(25.40)

Thus we find the Bianchi identity

$$D^* (e^\sigma G_{\lambda}^{\nu}) = 0.$$  

(25.41)

This makes the left-hand side of the new field equation (25.35) compatible with the autoparallel covariant conservation law (25.22).

If torsion exists in space and if it is not of the gradient type, there is, so far, no way of reconciling the classical result with field theory. The field theory is not able to render autoparallel trajectories for scalar particles. This remains an important problem of gravity with torsion.

There is, however, a simple resolution of this problem. If torsion has fermions such as quarks and leptons as the only sources, the torsion tensor is completely antisymmetric. In this case all autoparallels are automatically geodesics and the conflict disappears.
25.2 Scalar Fields

25.2.2 Electromagnetism

The most disturbing problem with torsion is that it is impossible to couple electromagnetism to it without destroying gauge invariance. As mentioned in the Preface, this was first observed by Schrödinger in the 1930s, causing him to derive upper bounds for the photon mass from experimental observations. The present upper bound is

\[ m_\gamma < 3 \times 10^{-27} \text{eV}. \] (25.42)

In order to be invariant under the usual electromagnetic gauge transformations

\[ A_\mu \rightarrow A_\mu + \partial_\mu \Lambda, \] (25.43)

the electromagnetic action

\[ A_{\text{em}} = -\frac{1}{4} \int d^4 x \sqrt{-g} F_{\mu \nu} F^{\mu \nu}, \] (25.44)

must contain the same field strengths

\[ F_{\mu \nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \] (25.45)

in spaces with curvature and torsion as in flat space. Gauge invariance permits replacing the derivatives \( \partial_\mu \) in \( F_{\mu \nu} \) by the Riemann covariant derivatives \( \bar{D}_\nu \), since the Christoffel symbols drop out in front of the antisymmetric tensor \( F_{\mu \nu} \):

\[ \bar{D}_\mu A_\nu - \bar{D}_\nu A_\mu = \partial_\mu A_\nu - \Gamma_{\mu \nu}^\lambda A_\lambda - \partial_\nu A_\mu + \bar{\Gamma}_{\nu \mu}^\lambda A_\lambda = \partial_\mu A_\nu - \partial_\nu A_\mu - 2 S_{\mu \nu} A_\lambda. \] (25.46)

Thus the field strength (25.45) is gauge invariant under electromagnetic, Einstein, and local Lorentz transformations.

For the full covariant derivative \( D_\mu \), however, the replacement introduces an additional term

\[ D_\mu A_\nu - D_\nu A_\mu = \partial_\mu A_\nu - \Gamma_{\mu \nu}^\lambda A_\lambda - \partial_\nu A_\mu + \Gamma_{\nu \mu}^\lambda A_\lambda = \partial_\mu A_\nu - \partial_\nu A_\mu - 2 S_{\mu \nu} A_\lambda, \] (25.47)

which destroys the gauge invariance. Inserting this covariant curl into the action (25.44) would lead to massive photons in any region with nonvanishing torsion.

As far as physical observations are concerned, the problem is really academic. We have seen before in the discussion of fermions, that torsion gives only forces at extremely short distances of the order of Planck length. Since no conceivable experiment can invade into such a short distance, the mass of the photons would remain undetectable. Moreover, at such short distances from matter, the properties of photons would much stronger be modified by electromagnetic dispersion and absorption than by any conceivable gravitational torsion field. Only if torsion could propagate, for which we see no possible mechanism, a non-gauge-invariant coupling would have fatal consequences. In fact, the experimental upper bound (25.42) is in complete agreement with the absence of torsion in the cosmos.
25.3 Compatibility Problems of Gravity with Torsion and Electroweak Interactions

Apart from the problem to find autoparallel classical trajectories for scalar particles and the violation of gauge invariance in the field theory of gravity with torsion, there are other difficulties. These arise on purely theoretical grounds and would be observable only if torsion could propagate, which it does not in Einstein-Cartan theory or any conceivable extension of it.

As we just saw, the electromagnetic field cannot couple minimally to torsion since this would destroy gauge invariance [5].

Massive vector bosons, on the other hand, such as the $\rho$-meson, whose wave function has a large amplitude in a state of a quark and an antiquark in an $s$-wave spin triplet channel, should certainly couple to torsion via their quark content.

By analogy with photons, the fundamental action describing electroweak processes should contain no minimally coupled torsion in the gradient terms of the bare vector bosons $W$ and $Z$. However, these particles acquire a mass via the Meissner-Higgs effect which makes them essentially composite particles, their fields being a mixture of the original massless vector fields and the Higgs fields. By analogy with the massive $\rho$-vector field, we could expect that also the massive electroweak vector fields couple to torsion, and the question arises how the Meissner-Higgs effect is capable of generating such a coupling.

25.3.1 Solution for Gradient Torsion

In general, we do not know the answer to the last problem. We can only give an answer if the torsion is of the gradient type

$$S_{\mu\nu}^\lambda(x) = \frac{1}{2} \left[ \delta_\mu^\lambda \partial_\nu \sigma(x) - \delta_\nu^\lambda \partial_\mu \sigma(x) \right]. \quad (25.48)$$

It is immediately obvious that a minimally-coupled scalar Higgs field with an action

$$A[\phi] = \int d^4x \sqrt{-g} \left( \frac{1}{2} g^{\mu\nu} |\nabla_\mu \phi \nabla_\nu \phi| - \frac{m^2}{2} |\phi|^2 - \frac{\lambda}{4} |\phi|^4 \right) \quad (25.49)$$

cannot equip a previously uncoupled massless vector field with a torsion coupling. For simplicity, we consider only a simple Ginzburg-Landau-type theory with a complex field to avoid inessential complications. As usual, $g = \det g_{\mu\nu}$ denotes the determinant of the metric $g_{\mu\nu}(x)$, and $\nabla_\mu$ is the electromagnetic covariant derivatives $\nabla_\mu = \partial_\mu - ieV_\mu$. The square mass is negative, so that the Higgs field has a nonzero expectation value with $|\phi|^2 = -m^2/\lambda$. From the derivative term, the vector field acquires a mass term $e^2 |\phi|^2 V_\mu V^\mu/2$, leading to the free part of the vector boson action

$$A[V] = \int d^4x \sqrt{-g} \left( -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{e^2 m^2}{2\lambda} V_\mu V^\mu \right), \quad (25.50)$$

H. Kleinert, GRAVITY WITH TORSION
where $F_{\mu\nu}$ the covariant curl $F_{\mu\nu} \equiv \partial_{\mu}V_{\nu} - \partial_{\nu}V_{\mu}$. Of course, the covariant curl of the nonabelian electroweak vector bosons would also have self-interactions, which can however be ignored in the present discussion since we are only interested in the free-particle propagation.

Since the Meissner-Higgs effect creates the mass of the vector bosons by mixing the uncoupled bare vector boson with the scalar Higgs field, it is obvious that the massive vector bosons can couple to torsion only if the scalar Higgs field has such a coupling. Indeed, it has recently been emphasized [6, 7] that, contrary to common belief [8], trajectories of scalar particles should experience a torsion force. This conclusion was reached by a careful reinvestigation of the geometric properties of the variational procedure of the action. Taking into account the fact that in the presence of torsion parallelograms exhibit a closure failure, the variational procedure required a modification of this procedure [6, 9, 10] which led to the conclusion that scalar particles should move along autoparallel trajectories rather than geodesic ones as derived from a minimally coupled scalar field action [8]. The modification of the variational procedure was suggested to us by the close analogy of spaces with torsion with crystals containing defects [11].

### 25.3.2 New Scalar Product

In the textbook [6] it has been pointed out that there exists a consistent Schrödinger formulation for a particle in a space with torsion if this has the restricted gradient form (25.48) or if it is completely antisymmetric. The Schrödinger equation is driven by the Laplace operator $g^\mu\nu D_\mu D_\nu$, where $D_\mu$ is the covariant derivative involving the full affine connection $\Gamma^{\lambda}_{\mu\nu}$, including torsion. It differs from the Laplace-Beltrami operator in torsion-free spaces $\Delta \equiv \sqrt{|g|}^{-1} \partial_\mu \sqrt{|g|} g^{\mu\nu} \partial_\nu$ by a term $-2S^\nu\lambda \partial_\nu = -3(\partial^\nu \sigma)\partial_\nu$. This operator, however, is hermitian only in a scalar product which contains a factor $e^{-3\sigma}$ [12]. In the case of totally antisymmetric torsion, the two Laplace operators are equal and the original scalar product ensures hermiticity and thus unitarity of time evolution. Such a torsion drops also out from the classical equation of motion, so that autoparallel and geodesic trajectories coincide. For this reason we shall continue the discussion only for gradient torsion.

The gradient torsion has the advantage that it can be incorporated into the classical action of a scalar point particle in such a way that the modification of the variational procedure found in [9, 10] becomes superfluous. The modified action reads for a massive particle [13]

$$A[x] = -mc \int d\tau e^{\sigma(x)} \sqrt{g_{\mu\nu}(x)} \dot{x}^\mu \dot{x}^\nu = -mc \int ds e^{\sigma(x(s))},$$

(25.51)

where $\tau$ is an arbitrary parameter and $s$ the proper time. From the Euler-Lagrange equation we find that for $\tau = s$, the Lagrangian under the integral is a constant of motion, whose value is, moreover, fixed by the mass shell constraint

$$L = e^{\sigma(x)} \sqrt{g_{\mu\nu}(x)} \dot{x}^\mu \dot{x}^\nu \equiv 1, \quad \tau = s.$$

(25.52)
The necessity of a factor $e^{-3\sigma(x)}$ in the scalar product discovered in [6] became the basis of a series of studies in general relativity [14, 15]. In the latter work, the action of a relativistic free scalar field $\phi$ was found to be

$$A[\phi] = \int d^4x \sqrt{-g} e^{-3\sigma} \left( \frac{1}{2} g^{\mu\nu} |\nabla_\mu \phi \nabla_\nu \phi| - \frac{m^2}{2} |\phi|^2 e^{-2\sigma} \right).$$  

(25.53)

The associated Euler-Lagrange equation is

$$D_\mu D^\mu \phi + m^2 e^{-2\sigma(x)} \phi = 0,$$

(25.54)

whose eikonal approximation $\phi(x) \approx e^{iE(x)}$ yields the following equation for the phase $E(x)$ [15]:

$$e^{2\sigma(x)} g^{\mu\nu}(x) [\partial_\mu E(x)] [\partial_\nu E(x)] = m^2.$$  

(25.55)

Since $\partial_\mu E$ is the momentum of the particle, the replacement $\partial_\mu E \rightarrow m \dot{x}_\mu$ shows that the eikonal equation (25.55) guarantees the constancy of the Lagrangian (25.52), thus describing autoparallel trajectories.

### 25.3.3 Self-Interacting Higgs Field

Apart from the factor $e^{-3\sigma(x)}$ accompanying the volume integral, the $\sigma$-field couples to the scalar field like a dilaton, the power of $e^{-\sigma}$ being determined by the dimension of the associated term. If we therefore add to the free-field action (25.53) a quartic self-interaction to have a Meissner-Higgs effect, this self-interaction will not carry an extra factor $e^{-\sigma}$, so that the proper Higgs action in the presence of gradient torsion reads

$$A[\phi] = \int d^4x \sqrt{-g} e^{-3\sigma} \left( \frac{1}{2} g^{\mu\nu} |\nabla_\mu \phi \nabla_\nu \phi| - \frac{m^2}{2} |\phi|^2 e^{-2\sigma} - \frac{\lambda}{4} |\phi|^2 \right).$$  

(25.56)

If $m^2$ is negative, and the torsion depends only weakly on spacetime, the Higgs field has a smooth vacuum expectation value

$$|\phi|^2 = -\frac{m^2}{\lambda} e^{-2\sigma}.$$  

(25.57)

The smoothness of the torsion field is required over a length scale of the Compton wavelength of the Higgs particle, i.e. over a distance of the order $1/20 \text{GeV} \approx 10^{-15}$ cm. For a torsion field of gravitational origin, this smoothness will certainly be guaranteed. From the gradient term in (25.56) we then extract in the gauge $\phi = \text{real}$ the mass term of the vector bosons

$$\int d^4x \sqrt{-g} e^{-3\sigma} \frac{1}{2} m_V^2 e^{-2\sigma(x)} V^\mu V_\mu,$$

(25.58)

where

$$m_V^2 = \frac{e^2}{\lambda} m^2, \quad m^2 < 0.$$  

(25.59)
Taking the physical scalar product in the presence of torsion into account, we obtain for the massive vector bosons the free-field action

\[ \mathcal{A}[V] = \int d^4x \sqrt{-g} e^{-3\sigma} \left( -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + m^2 V_\mu V^\mu \right). \]  

(25.60)

The appearance of the factor \( e^{-2\sigma} \) in the mass term guarantees again the same autoparallel trajectories in the eikonal approximation as for spinless particles in the action (25.53).

Note that the scalar product factor \( e^{-3\sigma(x)} \) implies a coupling to torsion also for the massless vector bosons which is fully compatible with gauge invariance. Due to the symmetry between Z-boson and photon, this factor must be present also in the electromagnetic action.

Notes and References


[8] F.W. Hehl, Phys. Lett. A 36, 225 (1971). See also Section 7 of [7].


[12] See Section 11.4 in Ref. [6]. Note that the normalization of the $\sigma$-field is normalized differently from the present one by a factor $2/3$. There we introduced $\delta$ via the relation $S_{\mu\nu}^\nu = \partial_\mu \sigma$, whereas here $S_{\mu\nu}^\nu = (3/2) \partial_\mu \sigma$.


At this point it is useful to remind the reader of an alternative theory of gravity proposed by Einstein in the thirties, the so-called theory of teleparallelism. As remarked in the Preface, this theory was inspired by Cartan’s work in 1922 and a subsequent letter communication between Einstein and Cartan [1]. In this theory, spacetime is assumed to be generated from flat space by assuming the multivalued local Lorentz tranformations $\Lambda^{\mu}_{\alpha}(x)$ in the basis tetrads $e^{a}_{\mu}(x)$ of Eq. (17.40) to be absent, so that $e^{a}_{\mu}(x)$ becomes single-valued and coincides with the vierbein field $h^{a}_{\mu}(x)$ with $a = \alpha$. It follows from Eq. (11.129) that the Riemann-Cartan curvature tensor vanishes identically:

$$R_{\mu\nu\lambda}^{\kappa} \equiv 0 \quad (\text{in teleparallel spacetime}). \quad (26.1)$$

This property has the pleasant consequence that it allows for the definition of parallel vector fields in all spacetime. Hence the name teleparallelism. Since the vanishing of the Riemann-Cartan curvature tensor is caused by by commuting derivatives in (11.129), the vanishing of $R_{\mu\nu\lambda}^{\kappa}$ may be considered as a Biachi identity of teleparallel spacetime. Since $\Lambda^{\mu}_{\alpha}(x)$ is single-valued, we may go at each point to a Lorentz frame where $\Lambda^{\mu}_{\alpha}(x) = 1$ and $\Gamma^{\alpha\beta\gamma} = 0$. Then we deduce from Eq. (17.63) that in this gauge the affine connection $\Gamma^{\mu}_{\nu\lambda}$ coincides with the quantities $h^{\nu}_{\mu\lambda}$ defined in Eq. (17.62):

$$\Gamma^{\mu}_{\nu\lambda} = h^{\nu}_{\mu\lambda} \equiv h^{\alpha}_{\mu}\frac{\partial}{\partial \mu}h^{\alpha}_{\nu} \equiv -h^{\alpha}_{\nu}\frac{\partial}{\partial \nu}h^{\alpha}_{\mu}, \quad (26.2)$$

and the torsion tensor reduces to the object of anholonomy (17.69):

$$S_{\mu\nu}^{\lambda} = \frac{1}{2} \left( h^{\alpha}_{\mu}\frac{\partial}{\partial \mu}h^{\alpha}_{\nu} - h^{\alpha}_{\nu}\frac{\partial}{\partial \nu}h^{\alpha}_{\mu} \right) \equiv -\frac{1}{2} \left( h^{\alpha}_{\nu}\frac{\partial}{\partial \nu}h^{\alpha}_{\mu} - h^{\alpha}_{\mu}\frac{\partial}{\partial \mu}h^{\alpha}_{\nu} \right). \quad (26.3)$$

Recalling the decomposition of the Riemann-Cartan curvature tensor (11.145), the vanishing of $R_{\mu\nu\lambda}^{\kappa}$ implies further that the Riemann curvature tensor can be expressed in terms of the kontortion tensor as

$$-\bar{R}_{\mu\nu\lambda}^{\kappa} = \bar{D}_{\mu}K_{\nu\lambda}^{\kappa} - \bar{D}_{\nu}K_{\mu\lambda}^{\kappa} - \left( K_{\mu\lambda}^{\rho}K_{\nu\rho}^{\kappa} - K_{\nu\lambda}^{\rho}K_{\mu\rho}^{\kappa} \right). \quad (26.4)$$
For the scalar curvature this becomes
\[
\bar{R} = \bar{R}_{\mu\nu\lambda}^\mu g^{\nu\lambda} = -\bar{D}_\mu K_{\nu}^{\nu\mu} + \bar{D}_\nu K_{\mu}^{\nu\mu} + \left( K_{\mu}^{\nu\rho} K_{\nu\rho\mu} - K_{\nu}^{\nu\rho} K_{\nu\rho\mu} \right),
\]  
(26.5)
or, expressing everything in terms of the torsion tensor using (11.118) and \( K_{\mu}^{\nu\rho} = 2S^{\nu} \) [recall (15.40)]:
\[
\bar{R} = -4\bar{D}_\mu S^{\mu} + \left( S_{\mu\nu\lambda} S^{\mu\nu\lambda} + 2S_{\mu\nu\lambda} S^{\mu\lambda\nu} - 4S^{\rho} S_{\rho} \right).
\]  
(26.6)
Thus, in a teleparallel spacetime, the Einstein action (15.8) can be replaced by
\[
\mathcal{A}_{E,S} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \left[ -4\bar{D}_\mu S^{\mu} + \left( S_{\mu\nu\lambda} S^{\mu\nu\lambda} + 2S_{\mu\nu\lambda} S^{\mu\lambda\nu} - 4S^{\rho} S_{\rho} \right) \right].
\]  
(26.7)
Integrating this by parts yields the action
\[
\mathcal{A}_{E,S} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \left( S_{\mu\nu\lambda} S^{\mu\nu\lambda} + 2S_{\mu\nu\lambda} S^{\mu\lambda\nu} - 4S^{\rho} S_{\rho} \right),
\]  
(26.8)
where we have dropped the pure surface term
\[
\mathcal{A}_{E,S,\text{surface}} = \frac{1}{2\kappa} \int d^4x \, 4\partial_\mu \left( \sqrt{-g} S^{\mu} \right),
\]  
(26.9)
which does not contribute to the field equations.

It is useful to decompose the Lagrangian density into irreducible parts of \( S_{\mu\nu\lambda} \) constructed with the help of Young tableaux [2]. An obvious irreducible part of \( S_{\mu\nu\lambda} \) is the vector \( S_{\mu} = S_{\mu\nu} \) of Eq. (15.40). It is associated with the mixed Young tableau:
\[
\mu \nu \lambda : S_{\mu\nu}. 
\]  
(26.10)
A second irreducible part is the totally symmetric combination:
\[
\mu \nu \lambda : t_{\mu\nu\lambda} = \frac{1}{2} \left( S_{\mu\lambda\nu} + S_{\mu\nu\lambda} \right) + \frac{1}{6} \left( g_{\mu\lambda\nu} S_{\nu} + g_{\mu\nu\lambda} S_{\nu} - 2g_{\nu\lambda\nu} S_{\mu} \right).
\]  
(26.11)
The third is an axial vector arising from the totally antisymmetric combination:
\[
\mu \nu \lambda : a^{\mu} = \frac{1}{6} e^{\mu\nu\lambda\kappa} S_{\nu\lambda\kappa}
\]  
(26.12)
where \( e^{\mu\nu\lambda\kappa} \) is the covariant Levi-Civita tensor (11A.1). The torsion can be recovered from these tensors and vectors as follows:
\[
S_{\mu\nu\lambda} = \frac{2}{3} \left( t_{\mu\nu\lambda} - t_{\nu\mu\lambda} \right) - \frac{1}{3} \left( g_{\mu\lambda\nu} S_{\nu} - g_{\nu\lambda\nu} S_{\mu} \right) - e_{\mu\nu\lambda\kappa} a^{\kappa}.
\]  
(26.13)
The three invariants in the action (26.8) can be expressed in terms of the irreducible invariants \( t_{\mu\nu\lambda}^{\mu\nu\lambda} \), \( S_{\mu}^{\mu} \), and \( a_{\mu}a^{\mu} \) as follows:

\[
S_{\mu\nu\lambda}S^{\mu\nu\lambda} = \frac{4}{3} t_{\mu\nu\lambda}^{\mu\nu\lambda} + \frac{2}{3} S_{\mu}^{\mu} - 6a_{\mu}a^{\mu}, \quad S_{\mu\nu\lambda}S^{\mu\lambda\nu} = \frac{2}{3} t_{\mu\nu\lambda}^{\mu\nu\lambda} + \frac{1}{3} S_{\mu}^{\mu} + 6a_{\mu}a^{\mu}.
\]

(26.14)

The inverse relations are

\[
t_{\mu\nu\lambda}^{\mu\nu\lambda} = \frac{1}{2} S_{\mu\nu\lambda}S^{\mu\nu\lambda} + \frac{1}{2} S_{\mu\nu\lambda}S^{\nu\lambda\mu} - \frac{1}{2} S_{\mu}^{\mu}, \quad a_{\mu}a^{\mu} = -\frac{1}{18} (S_{\mu\nu\lambda}S^{\mu\nu\lambda} - 2S_{\mu\nu\lambda}S^{\nu\lambda\mu}).
\]

(26.15)

Thus we can rewrite the action (26.8) as

\[
\mathcal{A}_{E,S} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \left( \frac{8}{3} t_{\mu\nu\lambda}^{\mu\nu\lambda} - \frac{8}{3} S_{\mu}^{\mu} + 6a_{\mu}a^{\mu} \right).
\]

(26.16)

All results of Einstein’s theory can be rederived from this action in Einstein-Cartan spacetime with \( R_{\mu\nu\lambda\kappa} = 0 \), in which the vierbein fields \( h_{\alpha}^{\mu} \) are four teleparallel vector fields.

This reformulation of Einstein gravity would only become interesting if future experiments were to discover deviations from Einstein’s theory. Since \( S_{\mu\nu\lambda} \) is a tensor field, the action (26.16) does not necessarily have to contain the three invariants in the specific combination implied by Eq. (26.6). Any other combination is invariant [2]:

\[
\mathcal{A}_{E,S} = \int d^4x \sqrt{-g} \mathcal{L}_{S} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \left( \gamma_1 t_{\mu\nu\lambda}^{\mu\nu\lambda} + \gamma_2 S_{\mu}^{\mu} + \gamma_3 a_{\mu}a^{\mu} \right).
\]

(26.17)

Of course, the three parameters \( \gamma_i \) are not completely free. There will be one constraint between them fixed by Newton’s law, so that the generalized theory has two free parameters. These can be fixed, for example, by the post-Newtonain expansion of the gravitational field around a mass point.

Expressed in terms of the original invariants in (26.17), the generalized action (26.17) reads

\[
\mathcal{A}_{S} = \int d^4x \sqrt{-g} \mathcal{L}_{S} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} \left( \sigma_1 S_{\mu\nu\lambda}S^{\mu\nu\lambda} + \sigma_2 S_{\mu\nu\lambda}S^{\nu\lambda\mu} + \sigma_3 S_{\mu}^{\mu} \right),
\]

(26.18)

with

\[
\sigma_1 = \frac{1}{2} \gamma_1 - \frac{1}{6} \gamma_3, \quad \sigma_2 = \frac{1}{2} \gamma_1 + \frac{1}{3} \gamma_3, \quad \sigma_3 = \gamma_2 - \frac{1}{3} \gamma_1.
\]

(26.19)

For the sake of deriving the equations of motion there exists several convenient forms of writing the action (26.18) to be distinguished by superscripts for better reference. The first is

\[
\mathcal{A}^{(1)}_{S} = \int d^4x \sqrt{-g} \mathcal{L}^{(1)} = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} S_{\mu\nu\lambda} P^{\mu\nu\lambda\rho\sigma\lambda'} S_{\rho\sigma\lambda'},
\]

(26.20)
where the tensor $P^{\mu\nu\lambda,\mu'\nu'\lambda'}$ is a combination of contravariant metric tensors:

$$P^{\mu\nu\lambda,\mu'\nu'\lambda'} = \sigma_1 g^{\mu\nu} g^{\nu\lambda'} g^{\lambda\lambda'} + \sigma_2 g^{\mu\nu'} g^{\nu\lambda'} g^{\lambda\lambda'} + \sigma_3 g^{\mu\nu'} g^{\nu\lambda} g^{\lambda\lambda'}. \quad (26.21)$$

The second form is

$$\mathcal{A}^{(2)}_S = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} S_{\mu\nu\lambda} F^{\mu\nu\lambda} \quad (26.22)$$

with

$$F^{\mu\nu\lambda} = \frac{\kappa}{2} \frac{\partial L}{\partial S_{\mu\nu\lambda}} = P^{[\mu\nu\lambda,\mu'\nu'\lambda']} S_{\mu'\nu'\lambda'}$$

$$= \sigma_1 S^{\mu\nu\lambda} + \frac{\sigma_2}{2} (S^{\mu\nu} - S^{\nu\lambda\mu}) + \frac{\sigma_3}{2} (g^{\mu\lambda} S^{\nu\mu} - g^{\nu\lambda} S^{\mu\nu}) = - F^{\mu\nu\lambda}. \quad (26.23)$$

The indices in brackets are antisymmetrized, as usual. This tensor may also be expressed as a modification of (26.13):

$$F^{\mu\nu\lambda} = \gamma_1 (t^{\mu\nu\lambda} - t^{\nu\mu\lambda}) - \gamma_2 (g^{\mu\lambda} S^{\nu\mu} - g^{\nu\lambda} S^{\mu\nu}) - \frac{\gamma_3}{3} \frac{\partial S^{\mu\nu\lambda \kappa}}{\partial \kappa} a_{\kappa} = - F^{\mu\nu\lambda}. \quad (26.24)$$

Using the asymmetry of $F^{\mu\nu\lambda}$ we can replace the torsion tensor in (26.22) by $h_{\alpha\lambda} \partial_{\mu} h^{\alpha\nu}$ via formula (26.3), and rewrite the action (26.22) also as

$$\mathcal{A}^{(3)}_S = \int d^4x \sqrt{-g} \mathcal{L}^{(3)}_S = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} h_{\alpha\lambda} \partial_{\mu} h^{\alpha\nu} F^{\mu\nu\lambda}, \quad (26.25)$$

or as

$$\mathcal{A}^{(4)}_S = \int d^4x \sqrt{-g} \mathcal{L}^{(4)}_S = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} h_{\alpha\lambda} \partial_{\mu} h^{\alpha\nu} P^{[\mu\nu\lambda,\mu'\nu'\lambda']} h_{\alpha'\lambda'} \partial_{\mu'} h^{\alpha'\nu'}, \quad (26.26)$$

where $P^{[\mu\nu\lambda,\mu'\nu'\lambda']}$ is the tensor (26.21) antisymmetrized in $\mu\nu$ and $\mu'\nu'$.

In order to have a better comparison with Einstein’s theory, we shall add to the above action also the Einstein action multiplied with a parameter $\gamma_0$, thus working with the Lagrangian density of the gravitational field

$$\mathcal{A} = \int d^4x \sqrt{-g} \mathcal{L} \equiv \gamma_0 \mathcal{A}_E + \mathcal{A}_S = \int d^4x \sqrt{-g} \left( \gamma_0 \mathcal{L}_E + \mathcal{L}_S \right)$$

$$= \int d^4x \sqrt{-g} \left[ -\frac{\gamma_0}{2\kappa} \bar{\mathcal{R}} - \frac{1}{2\kappa} \left( \gamma_1 t_{\mu\nu\lambda} t^{\mu\nu\lambda} + \gamma_2 S_{\mu\nu} S^{\mu\nu} + \gamma_3 a_{\mu} a^{\mu} \right) \right], \quad (26.27)$$

where the subscripts $E$ and $S$ indicate Einstein’s terms and their possible torsion corrections, respectively. The parametrization involves now a redundant parameter due to the equality of $\mathcal{L}_E$ and $\mathcal{L}_c$ for $\gamma_1 = -8/3, \gamma_2 = 8/3, \gamma_3 = -6$.

According to Eq. (17.134), the energy-momentum tensor of the gravitational field is obtained by varying the field action with respect to $h^{\alpha\nu}$ and multiplying the result with $h^{\alpha\nu}$:

$$\Theta^{\mu\nu} = -\frac{1}{\sqrt{-g}} h^{\alpha\nu} \delta \mathcal{A}/\delta h^{\alpha\nu} \quad (26.28)$$

H. Kleinert, GRAVITY WITH TORSION
Due to the absence of the object of anholonomy in the teleparallel formulation, the spin contribution analogous to (17.147) is absent, so that the result is here the canonical energy-momentum tensor, not the symmetric one. Applying formula (26.28) to the Einstein term in the action (26.27), which depends only on the metric tensor $g_{\mu\nu}$, we may evaluate the functional derivative on the right-hand side as $-\gamma_0 \sqrt{-g}^{-1} \delta \mathcal{A}_E / \delta g_{\mu\nu}$, and obtain, according to Eq. (17.139), $-\kappa^{-1}$ times the Einstein tensor $\tilde{G}^{\mu\nu}$. For this reason it will be convenient to use the left-hand side of (26.28) to define a generalized Einstein tensor $\tilde{G}^{\mu\nu}$ in the teleparallel Einstein-Cartan space by

$$f\Theta^{\mu\nu} \equiv -\frac{1}{\kappa} \tilde{G}^{\mu\nu},$$

(26.29)

whose Einstein contribution is $-\gamma_0 \kappa^{-1} \tilde{G}^{\mu\nu}$. The remaining action $\mathcal{A}_S$ adds to $\kappa^{-1} \tilde{G}^{\mu\nu}$ the correction

$$\frac{1}{\kappa} \Delta G^{\mu\nu} = \frac{1}{\sqrt{-g}} h^{\alpha\mu} \left[ \frac{\partial f_{\mathcal{L}S}}{\partial h^{\alpha\nu}} - \partial_{\sigma} \frac{f_{\mathcal{L}S}}{\partial h^{\alpha\nu}} \right].$$

(26.30)

One contribution to this is trivially found from the expression (26.25) of the action. By forming the derivatives with respect to the explicit $h^{\alpha\nu}$-terms, we obtain

$$\frac{1}{\kappa} \Delta^{(3)} G^{\mu\nu} = -\frac{1}{\kappa} \left( \Gamma_{\sigma\lambda\mu} F^{\sigma\lambda\nu} - h^{\alpha\mu} \frac{1}{\sqrt{-g}} \partial_{\sigma} \sqrt{-g} h_{\alpha\lambda} F^{\sigma\nu\lambda} \right).$$

(26.31)

The prefactor $1/2$ in (26.25) is cancelled by the fact that $S_{\mu\nu\lambda}$ is contained once more in $F^{\mu\nu\lambda}$ in a symmetric fashion, as is manifest in the expression (26.26) of the same action. Performing the derivative in the second term of (26.31) using (12.159) in the form

$$\frac{1}{\sqrt{-g}} \partial_{\nu} \sqrt{-g} = \tilde{\Gamma}^{\mu\nu} = \Gamma_{\mu\nu}^{\mu} - K_{\mu\nu}^{\mu} = \Gamma_{\mu\nu}^{\mu} + 2S_{\nu},$$

(26.32)

and exploiting further the antisymmetry of $F^{\sigma\lambda\nu}$, we may also write

$$\frac{1}{\kappa} \Delta^{(3)} G^{\mu\nu} = -\frac{1}{\kappa} \left( S_{\lambda\mu}^{\mu} F^{\sigma\lambda\nu} + S_{\lambda\mu}^{\nu} F^{\sigma\lambda\mu} - D_{\sigma} F^{\sigma\nu\mu} - 2S_{\nu} F^{\sigma\nu\mu} \right).$$

(26.33)

A further contribution to $\kappa^{-1} \Delta G^{\mu\nu}$ term comes from the dependence of the tensor $F^{\mu\nu\lambda\mu'\nu'}$ in the action (26.20) on the inverse metric $g_{\alpha\beta} = h_{\alpha\mu} h_{\beta\nu}$. The derivative rule (26.30) yields the contribution

$$\frac{1}{\kappa} \Delta^{(1)} G^{\mu\nu} \equiv -\frac{2}{\sqrt{-g}} \frac{\delta \mathcal{A}_S^{(1)}}{\delta g^{\mu\nu}},$$

(26.34)

which contains the three terms

$$\frac{1}{\kappa} \Delta^{(1)} G^{\mu\nu} = -\frac{1}{\kappa} \left[ \sigma_1 \left( 2S_{\lambda\nu}^{\mu} S^{\lambda\nu\lambda\kappa} + S_{\nu\lambda}^{\mu} S^{\lambda\nu\kappa} \right) + \sigma_2 \left( S_{\lambda\nu}^{\mu} S^{\nu\lambda\kappa} - S_{\nu\lambda}^{\mu} S_{\lambda\nu}^{\lambda\kappa} \right) - \sigma_3 \left( S^{\mu} S^{\nu} + S^{\lambda\mu} S_{\lambda}^{\nu} + S^{\lambda\nu} S_{\lambda}^{\mu} \right) \right].$$

(26.35)
A final contribution is due to the variation of $\sqrt{-g}$ in the action (26.18) [compare (17.137)], which adds to the energy-momentum tensor (26.28) a further term $-g^{\mu\nu}E_S$, i.e., to $\kappa^{-1}G^{\mu\nu}$ a further term

$$\frac{1}{\kappa} \Delta^{(4)}G^{\mu\nu} = g^{\mu\nu}L_S. \quad (26.36)$$

In order to add all terms we define several tensors:

$$H^{\mu\nu}_1 \equiv S^{\mu\lambda\kappa} S^{\nu}_{\lambda\kappa} = H_1^{T\mu\nu},$$
$$H^{\mu\nu}_2 \equiv S^{\mu\lambda\kappa} S^{\nu}_{\kappa\lambda} = H_2^{T\mu\nu},$$
$$H^{\mu\nu}_3 \equiv S^{\lambda\nu\mu} S^{\kappa\nu}_{\mu\kappa} = H_3^{T\mu\nu},$$
$$H^{\mu\nu}_4 \equiv S^{\mu\lambda\kappa} S^{\nu}_{\kappa\lambda} \neq H_4^{T\mu\nu},$$
$$H^{\mu\nu}_5 \equiv S^{\mu\nu} S^{\nu\mu} = H_5^{T\mu\nu},$$
$$H^{\mu\nu}_6 \equiv S^{\lambda\nu\mu} S^{\lambda}_{\mu\nu} \neq H_6^{T\mu\nu}. \quad (26.37)$$

Using these and the decomposition (26.23), we rewrite the first two terms in the parentheses of (26.33) as

$$S_{\sigma\lambda\mu} F^{\sigma\lambda\nu} (\mu \leftrightarrow \nu) = 2\sigma_1 H_3^{\mu\nu} - \sigma_2 (H_4^{\mu\nu} + H_4^{T\mu\nu}) + \frac{\sigma_3}{2} (H_6^{\mu\nu} + H_6^{T\mu\nu}). \quad (26.38)$$

The brackets on the right-hand side of (26.35) become

$$[\sigma_1 (2H_1^{\mu\nu} + H_3^{\mu\nu}) + \sigma_2 \left( H_2^{\mu\nu} - H_4^{\mu\nu} - H_4^{T\mu\nu} \right) + \sigma_3 \left( H_5^{\mu\nu} + H_6^{\mu\nu} + H_6^{T\mu\nu} \right)] \right). \quad (26.39)$$

Adding all contributions and also the energy-momentum tensor of matter, we obtain the field equation

$$\frac{1}{\kappa} \tilde{G}_{\mu\nu} = \frac{1}{\kappa} \tilde{G}_{\mu\nu} + \frac{1}{\kappa} \left[ D_\lambda F^{\lambda\nu}_{\mu} + 2S_\lambda F^{\lambda\nu}_{\mu} + H_{\mu\nu} \right] + g_{\mu\nu} L_S = \tilde{m}_{\mu\nu}, \quad (26.40)$$

where

$$H_{\mu\nu} = H_{\nu\mu} = 2S^{\mu\rho}_{\nu\sigma} F_{\nu^\rho} - S^{\sigma\rho}_{\nu\mu} F^{\sigma\rho}_{\mu}. \quad (26.41)$$

As a check we note that the two terms in the tensor (26.41) have the decomposition

$$S^{\mu\rho}_{\nu\sigma} F_{\nu^\rho} = \sigma_1 H_1^{\mu\nu} + \frac{\sigma_2}{2} (H_2^{\mu\nu} - H_4^{\mu\nu}) + \frac{\sigma_3}{2} (H_5^{\mu\nu} + H_6^{\mu\nu}), \quad (26.42)$$
$$S^{\sigma\rho}_{\nu\mu} F^{\sigma\rho}_{\mu} = \sigma_1 H_3^{\mu\nu} - \sigma_2 H_4^{\mu\nu} + \sigma_3 H_6^{\mu\nu}, \quad (26.43)$$

so that they combine to

$$H^{\mu\nu} = \sigma_1 (2H_1^{\mu\nu} - H_3^{\mu\nu}) + \sigma_2 H_2^{\mu\nu} + \sigma_3 H_5^{\mu\nu}, \quad (26.44)$$

which is the same as the difference between (26.38) and (26.39). Note that $H^{\mu\nu}$ is a symmetric tensor.

For a mass point at the origin, the energy-momentum tensor $\tilde{m}_{\mu\nu}$ is symmetric and equal to the symmetric energy-momentum tensor $\tilde{T}_{\mu\nu}$, which has the simple form (20.47):

$$\tilde{m}_{\mu\nu} = m \delta_{\mu}^{\nu} \delta_0^3(x). \quad (26.45)$$

H. Kleinert, GRAVITY WITH TORSION
26.1 Schwarzschild Solution

A spherically symmetric solution in empty space up to a point source at the origin is obtained setting $\Theta^\nu_{\mu\nu} = 0$ in Eq. (26.40) and inserting the spatially isotropic ansatz for the vierbein field

$$h^\alpha_\mu = \begin{pmatrix} \sqrt{H(r'')} & 0 & 0 & 0 \\ 0 & \sqrt{J(r'')} & 0 & 0 \\ 0 & 0 & \sqrt{J(r'')} & 0 \\ 0 & 0 & 0 & \sqrt{J(r'')} \end{pmatrix}.$$  \hfill (26.46)

The vierbein field $h^\alpha_\mu$ is given by the inverse of this matrix. The metric $g_{\mu\nu} = \eta_{\alpha\beta} h^\alpha_\mu h^\beta_\nu$ yields the invariant length (22.10), by construction.

The affine connection (26.2) has only a few nonzero matrix elements. The matrices $\Gamma^\nu_\mu$ and $\Gamma^\nu_3\gamma$ are zero, $\Gamma^\nu_\theta$ has only one nonzero element $\cot \theta$, and $\Gamma^\nu_\rho$ has only diagonal element:

$$\Gamma^\nu_\rho = \begin{pmatrix} \dot{h}(r'')/2 & 0 & 0 & 0 \\ 0 & j(r'') & 0 & 0 \\ 0 & 0 & 1/r'' & 0 \\ 0 & 0 & 0 & 1/r'' \end{pmatrix},$$  \hfill (26.47)

where $h = \log H$, $j = \log J$, and a dot denotes derivatives with respect to $r''$, as before in Eq. (22.36) and Subsection 22.2.2. Multiplying the field equation (26.40) by $\kappa$, we obtain as before in (22.31)–(22.35) only nonzero diagonal elements, extending the results of the Einstein theory to

$$\bar{G}^\nu_\rho = \frac{C}{2} \left\{ (1-2\epsilon) \ddot{h} + \frac{2}{r''} \left[ \epsilon \ddot{h} + (1-2\epsilon) \dot{j} \right] + \frac{\epsilon}{4} \dot{h}^2 + \frac{\epsilon}{2} \dot{h} \dot{j} + \frac{1-4\epsilon}{4} j^2 \right\}.$$  \hfill (26.48)

$$\bar{G}^{(0)} = \frac{C}{2} \left\{ (1-2\epsilon) \ddot{h} + \dot{j} + \frac{1}{r''} \left[ (1-2\epsilon) \dot{h} + \dot{j} \right] + \frac{1-3\epsilon}{2} \dot{h}^2 + \epsilon \dot{h} \dot{j} \right\},$$  \hfill (26.49)

$$\bar{G}^{(2)} = \frac{C}{2} \left\{ (1-2\epsilon) \ddot{h} + \dot{j} - \frac{1}{r''} \left[ (1-2\epsilon) \dot{h} + \dot{j} \right] + \frac{1-4\epsilon}{2} \dot{h}^2 - (1-3\epsilon) \dot{h} \dot{j} - \frac{1}{2} j^2 \right\}$$  \hfill (26.50)

where

$$C \equiv \gamma_0 - \gamma_2 - \gamma_1/4, \quad \epsilon \equiv -(\gamma_1 + \gamma_2)/4C.$$  \hfill (26.51)

Einstein’s expressions (22.33)–(22.35) are recovered by either setting $\gamma_0 = 1$ and $\epsilon = 0$, or by setting $\gamma_0 = 0$ and $\gamma_1 = 4/3, \gamma_2 = -4/3$. Note that $\gamma_3$ does not contribute to the Schwarzschild metric.

From (26.48)–(26.50) we find the spherical components (22.32) of the Einstein tensor

$$\bar{G}^r_r = \frac{C}{2J} \left\{ \frac{2}{r''} \left[ (1-2\epsilon) \ddot{h} + \dot{j} \right] + (1-2\epsilon) \dot{h} \dot{j} + \frac{\epsilon}{2} \dot{j}^2 + \frac{1}{2} j^2 \right\},$$  \hfill (26.52)

$$\bar{G}^\phi_\phi = \bar{G}^\theta_\theta = \frac{1}{2} \bar{G}^r_r + \frac{C}{2J} \left[ (1-2\epsilon) \ddot{h} + \dot{j} + \frac{1-7\epsilon}{2} \dot{h}^2 - \frac{1-4\epsilon}{2} \dot{h} \dot{j} + \frac{1}{4} j^2 \right].$$  \hfill (26.53)
rather than (26.52) and (26.53).

The vanishing of the combination $\bar{G}_r^r + \bar{G}_\theta^\theta$ yields now the differential equation extending (22.58)

$$\left(1 - 2\epsilon\right)\ddot{h} + \dot{j} + \frac{3}{r^2} \left[(1 - 2\epsilon)\dot{h} + \dot{j}\right] + \frac{1}{2} \left(\dot{h} + \dot{j}\right) \left[(1 - 2\epsilon)\dot{h} + \dot{j}\right] = 0, \tag{26.54}$$

which can be rewritten as

$$\frac{d}{dr^n} \left\{r^{n3}[1 - 2\epsilon]\dot{h} + \dot{j}\right\} + \frac{r^{n3}}{2} \left(\dot{h} + \dot{j}\right) [(1 - 2\epsilon)\dot{h} + \dot{j}] = 0, \tag{26.55}$$

and solved by [compare (22.60)]

$$\left(1 - 2\epsilon\right)\dot{h} + \dot{j} = \frac{1}{\sqrt{HJ}} \frac{c_1^2}{r^{n3}}, \tag{26.56}$$

where $c_1^2$ is a constant of integration. From the vanishing of $G^t_t + 2G^\theta^\theta - 3G^r_r$ (except at the mass point at the origin) we obtain

$$\frac{d}{dr^n} \left\{r^{n2}[(1 - 3\epsilon)\dot{h} + 2\epsilon\dot{j}]\right\} + \frac{1}{2r^{n2}} \left(\dot{h} + \dot{j}\right) [(1 - 3\epsilon)\dot{h} + 2\epsilon\dot{j}] = 0, \tag{26.57}$$

which is solved by

$$(1 - 3\epsilon)\dot{h} + 2\epsilon\dot{j} = \frac{1}{\sqrt{HJ}} \frac{c_2}{r^{n2}}, \tag{26.58}$$

where $c_2$ is a second constant of integration.

The combination $(1 - 5\epsilon)(1 - 2\epsilon)\left(\dot{h} + \dot{j}\right) = \frac{1}{\sqrt{HJ}} \left[(1 - 5\epsilon)\frac{c_1^2}{r^{n4}} + 2\epsilon\frac{c_2}{r^{n2}}\right], \tag{26.59}$$

or

$$\sqrt{HJ} = \frac{1}{2(1 - \epsilon)(1 - 4\epsilon)} \left[(1 - 5\epsilon)\frac{c_1^2}{r^{n4}} + 2\epsilon\frac{c_2}{r^{n2}}\right], \tag{26.60}$$

from which we find

$$\sqrt{HJ} = \left[1 - \frac{1}{4} \frac{(1 - 5\epsilon)c_1^2}{r^{n2}} - \frac{c_2}{r^{n2}}\right] \left[1 + \frac{a_+}{2r^{n2}}\right] \left[1 - \frac{a_-}{2r^{n2}}\right], \tag{26.61}$$

where $c_1^2 \equiv \frac{c_1^2}{(1 - \epsilon)(1 - 4\epsilon)}$, $c_2 \equiv \frac{c_2}{(1 - \epsilon)(1 - 4\epsilon)}$, and

$$a_+ = \sqrt{(1 - 5\epsilon)c_1^2 + \epsilon^2 c_2^2} + \epsilon c_2. \tag{26.62}$$

Anticipating that the generalization of the previous relation $c_2 = 2c_1$ is now $\bar{c}_2 = 2\bar{c}_1$, we obtain

$$a_+ = \left[\sqrt{(1 - \epsilon)(1 - 4\epsilon)} + 2\epsilon\right] \bar{c}_1. \tag{26.63}$$

H. Kleinert, GRAVITY WITH TORSION
Inserting this into (26.58) yields the differential equation

\[(1 - 3\epsilon)\dot{h} + 2\epsilon\dot{j} = \frac{2}{a_+ + a_-} \left( \frac{1}{1 - a_-/2r^\eta} - \frac{1}{1 + a_+/2r^\eta} \right) \frac{c_2}{r^{\eta^2}}. \tag{26.64}\]

This is integrated to

\[H^{1-3\epsilon} j^{2\epsilon} = \left[ \frac{1 - a_-/2r^\eta}{1 + a_+/2r^\eta} \right]^\nu, \quad \nu \equiv \frac{2c_2}{a_+ + a_-} = 2\sqrt{(1 - \epsilon)(1 - 4\epsilon)}, \tag{26.65}\]

which generalizes (22.67). As before in (22.65), the third constant of integration has been set equal to unity to ensure that the metric is Minkowsian at infinity.

Together with (26.61) we obtain

\[H = \frac{(1 - a_-/2r^\eta)^{\nu-4\epsilon/(1-5\epsilon)}}{(1 + a_+/2r^\eta)^{\nu+4\epsilon/(1-5\epsilon)}}, \quad J = \frac{(1 + a_+/2r^\eta)^{2+\nu-6\epsilon/(1-5\epsilon)}}{(1 - a_-/2r^\eta)^{-2+\nu+6\epsilon/(1-5\epsilon)}}, \tag{26.66}\]

so that

\[HJ = \left(1 + a_+/2r^\eta\right)^2 \left(1 - a_-/2r^\eta\right)^2. \tag{26.67}\]

Defining the powers

\[p_\pm = \sqrt{(1 - \epsilon)(1 - 4\epsilon) \pm 2\epsilon} \left(1 - 5\epsilon\right)^{-2\epsilon} \left(1 - \epsilon \right)^{+2\epsilon} = \frac{\bar{c}_1}{a_\pm}, \tag{26.68}\]

we can write

\[H = \frac{(1 - a_-/2r^\eta)^{2p_-}}{(1 + a_+/2r^\eta)^{2p_+}}, \quad J = \frac{(1 - a_-/2r^\eta)^{2(1-p_-)}}{(1 + a_+/2r^\eta)^{-2(1+p_+)}}. \tag{26.69}\]

The condition that the asymptotic behavior

\[H = 1 - \frac{p_+a_+ + p_-a_-}{r^\eta} + \ldots \tag{26.70}\]

yields Newton’s law if

\[p_+a_+ + p_-a_- = 2\bar{c}_1 = \frac{2MG}{c^2}, \tag{26.71}\]

[compare (22.54)].

The parameter \(\epsilon\) of the generalized theory changes the post-Newtonian approximation of Einstein’s theory. Expanding \(H\) and \(J\) to higher orders in \(1/r^\eta\), we obtain

\[H = 1 - \frac{2GM}{c^2r^\eta} + \frac{1}{4} \left( a_+^2p_+ - a_-^2p_- + 2a_+^2p_-^2 + 4a_+a_-p_+p_- + 2a_-^2p_+^2 \right) \left( \frac{GM}{c^2r^\eta} \right)^2 + \ldots \]

\[= 1 - \frac{2GM}{c^2r^\eta} + 2 \left(1 - \frac{\epsilon}{2}\right) \left( \frac{GM}{c^2r^\eta} \right)^2 + \ldots, \tag{26.72}\]
and

\[ J = 1 + \left( a_+ - a_- + a_+ p_+ + a_- p_- \right) \frac{GM}{c^2 r''} + \frac{1}{4} \left[ a_+^2 + a_-^2 - 4a_+ a_- (1 + p_+ - p_- - p_+ p_- \right) \]

\[-3(a_+^2 p_- - a_-^2 p_+) + 2(a_+^2 p_-^2 + 2a_+^2 p_+^2) \left( \frac{GM}{c^2 r''} \right)^2 + \ldots \]

\[= 1 + (1 - 2\epsilon) \frac{2GM}{c^2 r''} + 2 \left[ \frac{3}{4} (1 - 3\epsilon + 8\epsilon^2/3) \right] \left( \frac{GM}{c^2 r''} \right)^2 + \ldots . \] (26.73)

By comparison with (22.54) we identify the post-Newtonian parameters

\[ \beta = 1 - \frac{1}{2} \epsilon, \quad \gamma = 1 - 2\epsilon, \quad \delta = \frac{3}{4} (1 - 3\epsilon + 8\epsilon^2/3). \] (26.74)

Let us finally verify that the two constants of integration \( \tilde{c}_1 \) and \( \tilde{c}_2 \) are related by \( \tilde{c}_2 = 2\tilde{c}_1 \). Combining (26.56) with (26.58), we find

\[ \dot{h} = \frac{1}{\sqrt{H J}} \left( \frac{\tilde{c}_2}{r''^2} - 2\epsilon \frac{\tilde{c}_1^2}{r''^3} \right), \quad \dot{j} = \frac{1}{\sqrt{H J}} \left[ (1 - 3\epsilon) \frac{\tilde{c}_1^2}{r''^3} - (1 - 2\epsilon) \frac{\tilde{c}_2}{r''^2} \right], \] (26.75)

which generalize (22.62) and (22.63). Inserting these equations into \( \tilde{G}_{r''} = 0 \) of Eq. (26.52), we obtain

\[ \tilde{G}_{r''} = -\frac{C}{2J H J} (1 - \epsilon)(1 - 4\epsilon) \left( 2\epsilon_1^2 - \frac{1}{2} \epsilon_2^2 \right). \] (26.76)

which generalizes (22.70) and vanishes for \( \tilde{c}_2 = 2\tilde{c}_1 \), as anticipated.

It remains to fix the parameter \( C \) in Eqs. (26.48)–(26.50). For this we go to large-\( r'' \) where according to Eqs. (26.72) and (26.73)

\[ H = -\frac{2GM}{c^2 r''} + \ldots, \quad J = 1 + (1 - 2\epsilon) \frac{2GM}{c^2 r''} + \ldots, \] (26.77)

and consider the field equation Eq. (26.40) for \( \mu = \nu = t \)

\[ \frac{1}{\kappa} \tilde{G}_{t''} = M \epsilon \delta(x''). \] (26.78)

The point source changes the vanishing expression (22.61) into

\[ \frac{C}{J} \frac{d}{dr''} (r''^2 \dot{h}) + \frac{1}{2r''^2} \dot{h} (\dot{h} + \dot{j}) = \kappa M c \frac{\delta(r'')}{4\pi}, \] (26.79)

where we have replaced \( \delta^{(3)}(x'') \rightarrow \delta(r'')/4\pi r''^2 \). Here we recall that the Laplace equation for the Coulomb potential reads

\[ -\Delta \frac{1}{r} = 4\pi \delta^{(3)}(x). \] (26.80)
In spherical coordinates, this becomes
\[ -\frac{1}{r^2} \partial_r r^2 \partial_r \frac{1}{r} = \frac{1}{r^2} \delta(r), \]  
(26.81)
as is easily verified by integrating this equation from zero to a small nonzero radius \( r_0 \) and performing on the left-hand side an integration by parts. As a consequence, the solution (26.58) of the homogenous Eq. (26.57) for \( r'' \neq 0 \) acquires at \( r'' = 0 \) an inhomogenous part
\[ \frac{C}{J} \frac{d}{d\varpi} (r'^2 \dot{h}) + \frac{1}{2r'^2} \dot{h} (\dot{h} + j) = \frac{1}{\sqrt{HJ}} \frac{C c^2}{J} \delta(r''). \]  
(26.82)

Since \( H \) and \( J \) have unit values at the origin, comparison with (26.79) fixes \( C c^2 = \kappa M c^2 / 4\pi = 2G_NM/c^2 \) [recall (28.38)], or
\[ C c^2 = \bar{C} c^2 (1 - \epsilon)(1 - 4\epsilon) = C c^2 \frac{G_NM}{c^2}. \]  
(26.83)

Inserting \( \bar{C} = GM/c^2 \) from (26.71), the constant \( C \) must satisfy
\[ C(1 - \epsilon)(1 - 4\epsilon) = (c_0 - c_2 - c_1/4)(1 - \epsilon)(1 - 4\epsilon) = 1. \]  
(26.84)

### 26.2 Planetary Motion in Teleparallel Schwarzschild Metric

Consider first the geodesic trajectories. Inserting the Riemann connection Eq. (22.17) into (22.72), we obtain the differential equation
\[ \ddot{r}'' + \frac{H'}{r'' H} \dot{r}'' \dot{x}'' = 0, \]  
(26.85)
\[ \ddot{x}'' + \frac{J'}{2r'' J} [2x'' (x'' \cdot \dot{x}'') - x'' \dot{x}''] + \frac{c^2 H'}{2r'' J} x'' \dot{t}^2 = 0. \]  
(26.86)
where the dot indicates differentiation with respect to the proper time \( \tau \).

By rewriting the first equation as
\[ \ddot{r}'' + \frac{H'}{H} \dot{r}'' = \ddot{r}'' + \frac{H}{H} \dot{r}'' = 0, \]  
(26.87)
and using Eq. (22.12) to replace \( H(r'') \rightarrow B(r), H'(r'') \rightarrow B'(r)\sqrt{J/A}, \hat{H} = \hat{B} \), we recover the previous equation (22.73). The second equation may be multiplied by \( x \) and rewritten in a rotationally invariant form as
\[ x'' \cdot x'' + \frac{J'}{2r'' J} [2(x'' \cdot x'')^2 - r^2 \dot{x}^2] + \frac{c^2 r'' H'}{2J} \dot{t}^2 = 0. \]  
(26.88)
For comparison, let us calculate the autoparallel trajectories. Replacing the Riemann connection in Eq. (22.72) by the Riemann-Cartan connection [see (26.47)], we find the equations of motion

\[
\dddot{t} + \frac{H'}{2H} \dot{t} \dddot{x} = 0, \quad \dddot{x} + \frac{J'}{2r''J} \dot{x} \dddot{x} = 0.
\]  
(26.89)

The first can be rewritten as

\[
\dddot{t} + \frac{H'}{2H} \dot{r} \dddot{r} = 0,
\]  
(26.90)

the second implies that

\[
\dot{x} \dddot{x} + \frac{J'}{2r''J} (\dot{x} \dddot{x})^2 = 0,
\]  
(26.91)

or

\[
\dddot{r} + \left( \frac{J'}{2J} - \frac{1}{r''} \right) r''^2 - r'' (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) = 0.
\]  
(26.92)

Notes and References

[1] Einstein’s original German papers, in particular
A. Einstein, Math. Ann. 102, 685 (1939),
and their translations can be downloaded from
http://www.lrz-muenchen.de/~aunzicker/ae1930.html.

Emerging Gravity

The natural length scale of gravitational physics is the Planck length \( l_P \approx 1.616 \times 10^{-33} \text{ cm} \) formed in Eq. (28.39) from combinations of Newton’s gravitational constant \( G_N \), the light velocity \( c \), and Planck’s constant \( \hbar \). It is the Compton wavelength \( l_P \equiv \hbar/m_P c \) associated with the Planck mass \( m_P \approx 2.177 \times 10^{-5} \text{ g} = 1.22 \times 10^{22} \text{ MeV}/c^2 \). The Planck length is an extremely small quantity which presently lies beyond any experimental resolution, and will probably be so in the not too distant future. Particle accelerators are presently able to probe distances which are still 10 orders of magnitude larger than \( l_P \). Considering the fast growing costs of accelerators with energy, it is unimaginable, that they will get close to the Planck length for many generations to come. This length may therefore be considered as the shortest length accessible to experimental physics. Thus it makes no physical sense to produce theories which predict properties of the universe at smaller length scales. Since the times of Galileo Galilei, such theories fall into the realm of philosophy or even religion. The history of science shows us that nature has always surprised us with new discoveries as observations invaded into shorter and shorter distances. So far, all theories in the past which claimed for a while to be theories of everything have been falsified by such discoveries.

The history of theoretical physics is full of such examples of exotic theories. The presently most popular example is string theory. Its main strength lies in making predictions for the trans-Planckian regime down to zero length. In the experimentally accessible energy range, these theories require spacetime dimensions and predict particles which are not found in nature. In particular, the assumption of a string representing fundamental particles makes only sense if there are overtones, which any string must have. In the string model these overtones lie all in the inaccessible Planck regime. It is thus unclear how one can believe any of their predictions in this regime.

One of the most important features of string theories is that they predict the validity of Lorentz invariance at all energies in the trans-Planckian regime. In this chapter we would like to point out that if one is willing to spend time with speculations, an entirely different scenario is possible.
27.1 World Crystal

Let us suppose, just for the fun of it, that we live in a world crystal with a lattice constant of the order of the Planck length [1]. Up to now we would have been unable to notice this. And this would remain so for a long time to come. None of the present-day relativistic physical laws would be observably violated. The gravitational forces and their geometric description would arise from variants of the plastic forces in this world crystal. The observed curvature of spacetime would be just a signal of the presence of disclinations in the crystal. Matter would be sources of disclinations [2].

For simplicity, we shall present such a construction only for a system without torsion. If the world crystal is distorted by an infinitesimal displacement field

\[ x^\mu \rightarrow x'^\mu = x^\mu + u^\mu(x), \]  

(27.1)

it has a strain energy

\[ A = \mu \int d^4 x \left( \partial_\mu u_\nu + \partial_\nu u_\mu \right)^2, \]  

(27.2)

where \( \mu \) is some elastic constant. We assume the second possible elastic constant, the Poisson ratio, to be zero. If the distortions are partly plastic, the world crystal contains defects defined by Volterra surfaces, where crystalline sections have been cut out. The displacement field is multivalued, and the action (27.2) is the analog of the magnetic action (4.84) in the presence of a current loop. In order to do field theory with this action, we have to make the displacement field single-valued with the help of \( \delta \)-functions which describe the jumps across the Volterra surfaces, by complete analogy with the magnetic energy (4.85):

\[ A = \mu \int d^4 x \left( u_\mu^\nu - u_\mu^p\nu \right)^2, \]  

(27.3)

where \( u_\mu^\nu \equiv (\partial_\mu u_\nu + \partial_\nu u_\mu)/2 \) is the elastic strain tensor, and \( u_\mu^p\nu \) the plastic strain tensor [compare Eq. (9.71)] describing the Volterra surfaces via \( \delta \)-functions on these surfaces. As explained in Section 9.11, the plastic strain tensor is a gauge field of plastic deformations. The energy density is invariant under the single-valued defect gauge transformations [the continuum limit of (9.85)]

\[ u_\mu^p\nu \rightarrow u_\mu^p\nu + (\partial_\mu \lambda_\nu + \partial_\nu \lambda_\mu)/2, \quad u_\mu \rightarrow u_\mu + \lambda_\mu. \]  

(27.4)

Physically, they express the fact that defects are not affected by elastic distortions of the crystal. Only multivalued gauge functions \( \lambda_\mu \) change the defect content in \( u_\mu^p\nu \).

We now rewrite the action (27.5) in a canonical form [the analog of (4.86)] by introducing an auxiliary symmetric stress tensor field \( \sigma_{\mu\nu} \) as

\[ A = \int d^3 x \left[ \frac{1}{4\mu} \sigma_{\mu\nu} \sigma^{\mu\nu} + i \sigma^{\mu\nu} (u_\mu^p - u_\mu^p) \right]. \]  

(27.5)
After a partial integration and extremization in $u_\mu$, the second term in the action yields the equation

$$\partial_\nu \sigma^{\mu \nu} = 0. \quad (27.6)$$

This may be guaranteed identically, as a Bianchi identity, by an ansatz

$$\sigma^{\mu \nu} = \epsilon^\mu_\kappa \lambda \sigma \epsilon^\nu_\kappa \lambda' \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'}. \quad (27.7)$$

The field $\chi_{\sigma \sigma'}$ plays the role of an elastic gauge field. It is the analog of the vector potential $A(x)$ in Eq. (4.88). Inserting (27.7) into (27.5), we obtain the analog of (4.89):

$$A = \int d^4 x \left\{ \frac{1}{4 \mu} \left[ \epsilon^{\mu \kappa \lambda \sigma} \epsilon^{\nu \kappa \lambda' \sigma} \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'} \right]^2 + i \epsilon^{\nu \kappa \lambda \sigma} \epsilon^{\nu' \kappa \lambda' \sigma'} \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'} u_{\mu \nu} \right\}. \quad (27.8)$$

A further partial integration brings this to the form

$$A = \int d^4 x \left\{ \frac{1}{4 \mu} \left[ \epsilon^{\mu \kappa \lambda \sigma} \epsilon^{\nu \kappa \lambda' \sigma} \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'} \right]^2 + i \chi_{\sigma \sigma'} \left[ \epsilon^{\sigma \kappa \lambda \nu} \epsilon^{\sigma' \kappa \lambda' \mu} \partial_\lambda \partial_{\lambda'} u_{\mu \nu} \right] \right\}, \quad (27.9)$$

which is the analog of the double-gauge theory (4.90). The action is now double-gauge theory invariant under the defect gauge transformation (27.4), and under stress gauge transformations

$$\chi_{\sigma \tau} \rightarrow \chi_{\sigma \tau} + \partial_\sigma \Lambda_\tau + \partial_\tau \Lambda_\sigma. \quad (27.10)$$

The action can further be rewritten as

$$A = \int d^4 x \left\{ \frac{1}{4 \mu} \left[ \epsilon^{\mu \kappa \lambda \sigma} \epsilon^{\nu \kappa \lambda' \sigma} \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'} \right]^2 + i \chi_{\sigma \sigma'} \left[ \epsilon^{\sigma \kappa \lambda \nu} \epsilon^{\sigma' \kappa \lambda' \mu} \partial_\lambda \partial_{\lambda'} u_{\mu \nu} \right] \right\}, \quad (27.11)$$

where $\eta_{\mu \nu}$ is the four-dimensional extension of the defect density $\eta_{\mu \nu}$ in Eq. (12.34) [the analog of the magnetic current (4.91)]:

$$\eta_{\mu \nu} = \epsilon^\mu_\kappa \lambda \sigma \epsilon^\nu_\kappa \lambda' \partial_\lambda \partial_{\lambda'} u_{\mu \nu}. \quad (27.12)$$

This is invariant under defect gauge transformations (27.4), and satisfies the conservation law

$$\partial_\nu \eta^{\mu \nu} = 0. \quad (27.13)$$

We may now replace $u_{\mu \nu}$ by half the metric field $g_{\mu \nu}$ in (12.21) and, recalling Eq. (12.31), we recognize the tensor $\eta_{\mu \nu}$ as the Einstein tensor associated with the metric tensor $g_{\mu \nu}$.

Let us eliminate the stress gauge field from the action (27.11). For this we use the identity (1A.23) for the product of two Levi-Civita tensors, and rewrite the stress field (27.7) as

$$\sigma^{\mu \nu} = \epsilon^\mu_\kappa \lambda \sigma \epsilon^\nu_\kappa \lambda' \partial_\lambda \partial_{\lambda'} \chi_{\sigma \sigma'}$$

$$= -(\partial^2 \chi_{\mu \nu} + \partial_\mu \partial_\nu \chi^\lambda - \partial_\mu \partial_\lambda \chi^\nu - \partial_\nu \partial_\lambda \chi^\mu) + \eta_{\mu \nu} (\partial^2 \chi^\lambda - \partial_\lambda \partial_\kappa \chi^\kappa). \quad (27.14)$$
Introducing the field $\phi_{\mu}^{\nu} \equiv \chi_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \chi_{\lambda}^{\lambda}$, and going to the Hilbert gauge $\partial^{\mu} \phi_{\mu}^{\nu} = 0$, the stress tensor reduces to

$$\sigma_{\mu\nu} = -\partial^{2} \phi_{\mu\nu},$$

and the action of an arbitrary distribution of defects becomes

$$A = \int d^{4}x \left\{ \frac{1}{4\mu} \partial^{2} \phi^{\mu\nu} \partial^{2} \phi_{\mu\nu} + i \phi_{\mu}^{\nu} (\eta_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \eta_{\lambda}^{\lambda}) \right\}. \quad (27.15)$$

Extremization with respect to the field $\phi^{\mu\nu}$ yields the interaction of an arbitrary distribution of defects [the analog of (4.92)]:

$$A = \mu \int d^{4}x (\eta_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \eta_{\lambda}^{\lambda}) \frac{1}{(\partial^{2})^{2}} (\eta_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \eta_{\lambda}^{\lambda}). \quad (27.16)$$

This is not the Einstein action for a Riemann spacetime. It would be so if the derivatives $\partial^{2}$ in (27.31) were replaced by $\partial$. Then the Green function of $(\partial^{2})^{2}$ would be replaced by the Green function of $-\partial^{2}$. An index rearrangement would lead to the interaction

$$A = \mu \int d^{4}x (\eta_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \eta_{\lambda}^{\lambda}) \frac{1}{-\partial^{2}} \eta_{\mu}^{\nu}. \quad (27.17)$$

The defect tensor $\eta_{\mu\nu}$ is composed of the plastic gauge fields $u^p_{\mu\nu}$ in the same way as the stress tensor is in terms of the stress gauge field in Eq. (27.14), i.e.:

$$\eta_{\mu\nu} = \epsilon_{\mu}^{\kappa\lambda} \epsilon_{\nu}^{\lambda'} \partial_{\lambda} \partial_{\lambda'} u^p_{\sigma\tau}. \quad (27.18)$$

If we introduce the auxiliary field $w^p_{\mu\nu} \equiv u^p_{\mu\nu} - \frac{1}{2} \delta_{\mu}^{\nu} u^p_{\lambda\lambda}$, and chose the Hilbert gauge $\partial^{\mu} w^p_{\mu\nu} = 0$, the defect density reduces to

$$\eta_{\mu\nu} = -\partial^{2} w^p_{\mu\nu}, \quad \eta_{\mu}^{\nu} - \frac{1}{2} \delta_{\mu}^{\nu} \eta_{\lambda}^{\lambda} = -\partial^{2} u^p_{\mu\nu}. \quad (27.19)$$

and the interaction (27.18) of an arbitrary distribution of defects would become

$$A = \mu \int d^{4}x u^p_{\mu\nu}(x) \eta_{\mu\nu}(x). \quad (27.20)$$

This coincides with the linearized Einstein action

$$A = \frac{1}{2\kappa} \int d^{4}x \sqrt{-g} \bar{R} \quad (27.21)$$

where $\kappa$ is the gravitational constant. Indeed, in the linear approximation $g_{\mu}^{\nu} = \delta_{\mu}^{\nu} + h_{\mu}^{\nu}$ with $|h_{\mu}^{\nu}| \ll 1$, where the Christoffel symbols can be approximated by

$$\bar{\Gamma}_{\mu\nu}^{\lambda} \approx \frac{1}{2} \left( \partial_{\nu} h_{\lambda\mu} + \partial_{\mu} h_{\lambda\nu} - \partial_{\lambda} h_{\mu\nu} \right). \quad (27.22)$$
and the Riemann curvature tensor becomes
\[ \tilde{R}_{\mu\nu\lambda\kappa} \approx \frac{1}{2} \left[ \partial_\mu \partial_\lambda h_{\nu\kappa} - \partial_\nu \partial_\kappa h_{\mu\lambda} - (\mu \leftrightarrow \nu) \right], \tag{27.24} \]
as can be seen directly from Eq. (11.150). This gives the Ricci tensor
\[ \tilde{R}_{\mu\kappa} \approx \frac{1}{2} \left( \partial_\mu \partial_\lambda h_{\lambda\kappa} + \partial_\kappa \partial_\lambda h_{\mu\lambda} - \partial_\mu \partial_\kappa h - \partial^2 h_{\mu\kappa} \right), \tag{27.25} \]
where \( h \equiv h^{\lambda}_{\lambda} \) is the trace of the tensor \( h_{\mu\nu} \). The associated scalar curvature is
\[ \tilde{\mathcal{R}} \approx -\left( \partial^2 h - \partial_\mu \partial_\nu h^{\mu\nu} \right). \tag{27.26} \]
In combination with (27.25) we obtain the Einstein tensor
\[ \tilde{G}_{\mu\kappa} = \tilde{R}_{\mu\kappa} - \frac{1}{2} g_{\mu\kappa} \tilde{\mathcal{R}} \tag{27.27} \]
\[ \approx -\frac{1}{2} \left( \partial^2 h_{\mu\kappa} + \partial_\mu \partial_\kappa h - \partial_\mu \partial^2 h^{\kappa}_{\lambda} - \partial_\kappa \partial^2 h^{\mu}_{\lambda} \right) + \frac{1}{2} \eta_{\mu\kappa} \left( \partial^2 h - \partial_\mu \partial_\nu h^{\mu\nu} \right). \]
This can be written as a four-dimensional version of a double curl
\[ \tilde{G}_{\mu\kappa} = \frac{1}{2} \epsilon_{\mu\delta}^{\phantom{\mu\delta}\nu\lambda} \epsilon_{\kappa\delta\sigma}^{\phantom{\kappa\delta\sigma}\tau} \partial_\nu \partial_\sigma h_{\lambda\tau}, \tag{27.28} \]
as can be verified using the identity (1A.23).

Thus the Einstein-Hilbert action has the linear approximation
\[ \mathcal{A} \approx \frac{1}{4\kappa} \int d^4 x \ h_{\mu\nu} \tilde{G}^{\mu\nu}. \tag{27.29} \]
Recalling the previously established identifications of plastic field and defect density with metric and Einstein tensor, respectively, the interaction between defects (27.21) is indeed the linearized version of the Einstein-Hilbert action (27.22), if we identify the constant \( \mu \) with \( 1/4\kappa \).

The world crystal with the elastic energy (27.5) does not lead to this action. It must be modified. A first modification is to introduce two more derivatives and assign to the crystal the higher-gradient elastic energy
\[ \mathcal{A}' = \mu \int d^4 x \ \left[ \partial(u_{\mu\nu} - u^{\mu}_{\mu\nu}) \right]^2. \tag{27.30} \]
This removes one power of \(-\partial^2\) from the denominator in the interaction (27.17).

I order to obtain the correct contractions in (27.18), we must replace the action (27.31) by
\[ \mathcal{A} = \int d^4 x \left[ -\frac{1}{4\kappa} \left( \phi^{\mu\nu} \partial_\phi^{\mu\nu} - \frac{1}{2} \phi^{\mu}_{\mu} \partial^2 \phi^{\nu}_{\nu} \right) + i \phi^{\nu}_{\nu} (\eta^{\mu}_{\mu} - \frac{1}{2} \delta^{\mu}_{\nu} \eta^{\lambda}_{\lambda}) \right]. \tag{27.31} \]
This, in turn, follows from an interaction energy
\[ \mathcal{A} = \mu \int d^4 x \ \left\{ [\partial(u_{\mu\nu} - u^{\mu}_{\mu\nu})]^2 - \frac{1}{4} [\partial(u^{\mu}_{\mu} - u^{\mu}_{\mu\nu})]^2 \right\}. \tag{27.32} \]
Thus we have shown that defects in the world crystal create a Riemannian space-time with a Euclidean action of the Einstein type.
27.2 Gravity Emerging from Fluctuations of Matter and Radiation in Closed Friedmann Universe

In 1967 Sacharov put forward an interesting idea [3, 4] that the geometry does not possess a dynamics of its own, but that the stiffness of spacetime could be entirely due to the vacuum fluctuations of the fundamental fields in the universe (scalar, vector, tensor, spinor). These give rise to an Einstein action proportional to $R$, but also a cosmological term without $R$, and to all possible higher powers of $R_{\mu\nu\lambda\kappa}$ contracted to scalars such as $R^2$, $R_{\mu\nu}R^{\mu\nu}$, $R_{\mu\nu\lambda\kappa}R^{\mu\nu\lambda\kappa}$, $R^3$, . . . . The lowest three coefficients diverge in the ultraviolet, but if all fluctuating field stem from a renormalizable quantum field theory, all infinities can be subtracted to leave a finite value to be fixed by experiment.

The cosmological term without $R$ changes the gravitational action (15.8) to

$$f A = -\frac{1}{2\kappa} \int d^4x \sqrt{-g} (R + 2\lambda).$$

(27.33)

The constant $\lambda$ is the so-called cosmological constant. It changes the energy-momentum tensor of the gravitational field from $-(1/\kappa)G_{\mu\nu}$ to $-(1/\kappa) \left(G_{\mu\nu} - \lambda g_{\mu\nu}\right)$.

When calculating the effect of fluctuations for any of the fundamental fields one finds a contribution to the cosmological constant corresponding to an action density

$$\Lambda \equiv \frac{\lambda}{\kappa} = \frac{\lambda c^3}{8\pi G_N}$$

(27.34)

which is of the order of $\pm \hbar/l_P^4$, where $l_P$ is the Planck length. For bosons, the sign is positive, for fermions negative, reflecting the filling of all negative-energy states in the vacuum.

A constant of this size is much larger than the present experimental estimate. In the literature one usually finds estimates for the dimensionless quantity

$$\Omega_{\lambda 0} \equiv \frac{\lambda c^2}{3H_0^2}$$

(27.35)

where $H_0$ is the Hubble constant which parametrizes the expansion velocity of the universe as a function of the distance $r$ from us by Hubble’s law:

$$v = H_0 r.$$  

(27.36)

The inverse of $H_0$ is roughly equal to the lifetime of the universe

$$H_0^{-1} \approx 14 \times 10^9 \text{ years}.$$  

(27.37)

Present fits to distant supernovae and other cosmological data yield the estimate [5]

$$\Omega_{\lambda 0} \equiv 0.68 \pm 0.10.$$  

(27.38)
As a result, the experimental number for $\Lambda$ is

$$\Lambda = \Omega_{\lambda} \frac{8H_0^2}{c^2} \frac{l_p^4}{8\pi} \approx 10^{-12} \bar{h} l_p^4.$$  \hfill (27.39)

Such a small prefactor in front of the “natural” action density $\bar{h}/l_p^4$ can only arise from an almost perfect cancellation of the contributions of boson and fermion fields. This cancellation is the main reason why some people postulate the existence of a broken supersymmetry in the universe, in which every boson has a fermionic counterpart. So far, the known particle spectra show no trace of such a symmetry. Thus there is need to explain it by some other not yet understood mechanism.

A mechanical model for Sacharov’s idea of emerging gravity would be an infinitely thin plastic bag filled with water. The bag represents the geometry which does not have any dynamics of its own. All its movements are controlled by the dynamics of the water contained in it.

Sacharov’s idea is very appealing. Unfortunately, the calculation of the emerging gravitational action requires the knowledge of what are all elementary fields in nature. It is questionable whether this will ever be available. Moreover, it is unclear what role is played by the vacuum fluctuations of composite particles, such as the many elements in the periodic system. Do gold and lead contribute to the vacuum energy?

**Notes and References**


Cosmology with Fluctuating Boson Field

Let us calculate the change of the equations of motion in a slowly evolving Friedmann universe caused by fluctuating bosonic matter and radiation [1]. The contribution to the cosmological constant will be ignored since we do not discuss the negative contributions of the fermion fields. The equation of state turns out to have the form $P = \rho/3$ for radiation with arbitrary coupling.

An important quantum effect in a Friedmann universe is that of particle creation [6, 7, 8]. It is a dynamic effect and depends sensitively on the speed of evolution. This will be ignored here, assuming the evolution to be sufficiently slow, to get pure finite-size effects (see e.g. [9, 10]).

For simplicity, we shall consider only a closed Friedmann universe. There the difference between the field energy in flat and curved universes bears close resemblance with the so-called Casimir energy of that field. Recall that this arises from the difference between the spectral sum in the closed universe and the spectral integral in the flat space. This difference can be calculated with the help of the Euler-Maclaurin formula. Casimir was the first to observe that such a difference arises in a flat space-time if electromagnetic fields are enclosed between conducting plates. It gives rise to an attraction called Casimir effect [3], and discussed extensively in the literature (see the reviews in [4, 5]). Until now, Casimir energies have been investigated mainly for zero temperature [5, 6, 7, 8, 11, 12]. In Ref. [1] it was derived for any temperature.

For the description of both fluctuating matter and radiation we consider generically a free massive scalar field coupled to gravity. If the mass is set equal to zero, the result can be applied to fluctuating radiation. The free energy of the scalar field is obtained by performing the functional integral over the Fourier components of free field [13, 14]. In the static universe, there is no problem to do this since the oscillator frequencies are time independent.

The regularization of the infinite sums is performed using various standard methods existing in the literature [5, 6, 7, 17, 18, 19, 20, 21]. The expressions for the energy density and the pressure follow from the free energy by the thermodynamic rules.
28.0.1 Robertson-Walker-Friedmann Universe

The simplest model of a universe is due to Friedmann set up in a Robertson-Walker background geometry.

As an idealization of the observed density of matter, the universe is assumed to be isotropic and homogeneous. Then it is convenient to describe it by a coordinate frame in which the metric is rotationally invariant. To account for the expansion, we have to allow for an explicit time dependence of the spatial part of the metric. In the spatial part, we choose coordinates which participate in the expansion. They can be imagined as being attached to the gas particles in a homogenized universe. Then the time passing at each coordinate point is the proper time. In this context it is the so-called cosmic standard time to be denoted by \( t \). We imagine an observer sitting at a coordinate point with \( \frac{dx^i}{dt} = 0 \), and \( t \) is measured by counting the number of orbits of an electron around a proton in a hydrogen atom, starting from the moment of the big bang (forgetting the fact that in the early time of the universe, the atom does not really exist).

**Geometry**

With this time calibration, the component \( g_{00} \) of the metric tensor is identically equal to unity

\[
g_{00}(x) \equiv 1, \quad (28.1)
\]

such that at a fixed coordinate point, the proper time coincides with the coordinate time, \( d\tau = dt \). Moreover, since all clocks in space follow the same prescription, there is no mixing between time and space coordinates, a property called time orthogonality, so that

\[
g_{0i}(x) \equiv 0. \quad (28.2)
\]

As a consequence, the Christoffel symbol \( \bar{\Gamma}^{\mu}_{00} \) [recall (11.22)] vanishes identically:

\[
\bar{\Gamma}^{\mu}_{00} \equiv \frac{1}{2} g^{\mu\nu} (\partial_0 g_{0\nu} + \partial_0 g_{0\nu} - g_{00}) \equiv 0. \quad (28.3)
\]

This is the mathematical way of expressing the fact that a particle sitting at a coordinate point, which has \( \frac{dx^i}{dt} = 0 \), and thus \( \frac{dx^\mu}{dt} = u^\mu = (c, 0, 0, 0) \), experiences no acceleration

\[
\frac{du^\mu}{d\tau} = -\bar{\Gamma}^{\mu}_{00} c^2 = 0. \quad (28.4)
\]

The coordinates themselves are trivially comoving.

Under the above condition, the invariant distance has the general form

\[
ds^2 = c^2 dt^2 - (3) g_{ij}(x) dx^i dx^j. \quad (28.5)
\]
We now impose the spatial isotropy upon the spatial metric \( g_{ij} \). Denote the spatial length element by \( dl \), so that
\[
dl^2 = (^{(3)} g_{ij}(x) dx^i dx^j). \tag{28.6}
\]
The isotropy and homogeneity of space is most easily expressed by considering the spatial curvature \((^{(3)} R_{ijkl})\) calculated from the spatial metric \((^{(3)} g_{ij}(x))\). The space corresponds to a spherical surface. If its radius is \( a \), the curvature tensor is, according to Eq. (12.202),
\[
^{(3)} R_{ijkl}(x) = \frac{1}{a^2} \left[ (^{(3)} g_{il}(x) (^{(3)} g_{jk}(x) - (^{(3)} g_{ik}(x) (^{(3)} g_{jl}(x) \right]. \tag{28.7}
\]
The derivation of this expression in Subsection 12.7.4 was based on the assumption of a spherical space whose curvature \( K \equiv 1/a^2 \) is positive. If we allow also for hyperbolic and parabolic spaces with negative and vanishing curvature, and characterize these by a constant
\[
k = \begin{cases} 
1 & \text{spherical} \\
0 & \text{parabolic} \\
-1 & \text{hyperbolic}
\end{cases} \text{ universe,} \tag{28.8}
\]
then the prefactor \( 1/a^2 \) in (28.7) is replaced by \( K \equiv k/a^2 \). For \( k = -1 \) and 0, the space has an open topology and an infinite total volume.

The Ricci tensor and curvature scalar are in these three cases [compare (12.203) and (12.196)]
\[
^{(3)} R_{il} = k \frac{2}{a^2} g_{il}(x), \quad ^{(3)} R = k \frac{6}{a^2}. \tag{28.9}
\]

By construction it is obvious that for \( k = 1 \), the three-dimensional space has a closed topology and a finite spatial volume which is equal to the surface of a sphere of radius \( a \) in four dimensions
\[
^4 S^a = 2\pi^2 a^3. \tag{28.10}
\]
A circle in this space has maximal radius \( a \) and a maximal circumference \( 2\pi a \). A sphere with radius \( r_0 < a \) has a volume
\[
^{(3)} V_{r_0}^a = \int_0^{2\pi} d\varphi \int_0^\pi d\theta \int_0^r dr \frac{r^2}{\sqrt{1-r^2/a^2}} \tag{28.11}
\]
\[
= 4\pi \left( \frac{a^3}{2} \arcsin \frac{r_0}{a} - \frac{a^2 r_0}{2} \sqrt{1 - \frac{r_0^2}{a^2}} \right).
\]

For small \( r_0 \), the curvature is irrelevant and the volume depends on \( r_0 \) like the usual volume of a sphere in three dimensions:
\[
^{(3)} V_{r_0}^a \approx V_{r_0} = \frac{4\pi}{3} r_0^2. \tag{28.12}
\]

For \( r_0 \to a \), however, \( ^{(3)} V_{r_0}^a \) approaches a saturation volume \( 2\pi a^3 \).

The analogous expressions for negative and zero curvature are obvious.

H. Kleinert, GRAVITY WITH TORSION
Robertson-Walker Metric

In spherical coordinates, the four-dimensional invariant distance (28.5) defines the Robertson-Walker metric.

\[ ds^2 = c^2 dt^2 - dl^2 \]  
\[ dl^2 = \frac{dr^2}{1 - kr^2/a^2} + r^2(d\theta^2 + \sin^2 \theta d\varphi^2). \]  
\[ (28.13) \]
\[ (28.14) \]

It will be convenient to introduce, instead of \( r \), the angle \( \alpha \) on the surface of the four-sphere, such that

\[ r = a \sin \alpha. \]  
\[ (28.15) \]

Then the metric has the four-dimensional angular form

\[ ds^2 = c^2 dt^2 - a^2(t)[d\alpha^2 + f^2(\alpha)(d\theta^2 + \sin^2 \theta d\varphi^2)]. \]  
\[ (28.16) \]

where for spherical, parabolic, and hyperbolic spaces, \( f(\alpha) \) is equal to

\[ f(\alpha) = \begin{cases} 
\sin \alpha & k = 1, \\
\alpha & k = 0, \\
\sinh \alpha & k = -1. 
\end{cases} \]  
\[ (28.17) \]

In order to have maximal symmetry, it is useful to absorb \( a(t) \) into the time and define a new timelike variable \( \eta \) by

\[ c dt = a(\eta) d\eta \]  
\[ (28.18) \]

so that the invariant distance is measured by

\[ ds^2 = a^2(\eta)[d\eta^2 - d\alpha^2 - f^2(\alpha)(d\theta^2 + \sin^2 \theta d\varphi^2)]. \]  
\[ (28.19) \]

Then the metric is simply

\[ g_{\mu\nu} = a^2(\eta) \begin{pmatrix} 1 & -1 & -f^2(\alpha) & -f^2(\alpha) \sin^2 \theta \\
\end{pmatrix} \]  
\[ (28.20) \]

and the Christoffel symbols become

\[ \Gamma^0_{00} = \frac{a^2}{a}, \quad \Gamma^0_{0i} = 0, \quad \Gamma^0_{0i} = 0, \quad \Gamma^0_{0i} = \frac{a^2}{a} \delta^i_j, \quad \Gamma^0_{ij} = -\frac{a^2}{a^2} g_{ij}, \quad \Gamma^0_{ij} = 0, \]  
\[ (28.21) \]

where the subscripts denote derivatives with respect to the corresponding variables:

\[ a_\eta \equiv \frac{da}{d\eta} = \frac{a}{c} \frac{dt}{d\eta} = \frac{a}{c} a_\tau. \]  
\[ (28.22) \]
We now calculate the 00-component of the Ricci tensor:

\[
R_{00} = \partial_\mu \Gamma^\mu_{00} - \partial_0 \Gamma_{\mu0}^\mu - \Gamma^\mu_{0\nu} \Gamma_{00}^\nu + \Gamma^\nu_{0\mu} \Gamma_{00}^\mu. \tag{28.23}
\]

Inserting the Christoffel symbols (28.21) we find

\[
\partial_\mu \Gamma^\mu_{00} - \partial_0 \Gamma_{\mu0}^\mu = -3 \frac{d}{d\eta} \frac{a_\eta}{a} = -3 \frac{1}{a^2} \left( a_{\eta\eta} a - a_\eta^2 \right),
\]

\[
\Gamma^\mu_{0\nu} \Gamma_{00}^\nu = \Gamma^\mu_{00} \Gamma^\nu_{00} + \Gamma^\nu_{0i} \Gamma_{0i}^\mu + \Gamma^\mu_{0i} \Gamma_{0i}^\nu + \Gamma^\nu_{0i} \Gamma_{0i}^\mu = \left( \frac{a_\eta}{a} \right)^2 + 3 \left( \frac{a_\eta}{a} \right)^2, \tag{28.24}
\]

\[
\Gamma^\mu_{0\nu} \Gamma_{00}^\nu = \Gamma^\mu_{00} \Gamma^\nu_{00} + \Gamma^\nu_{0i} \Gamma_{0i}^\mu + \Gamma^\mu_{0i} \Gamma_{0i}^\nu + \Gamma^\nu_{0i} \Gamma_{0i}^\mu = \left( \frac{a_\eta}{a} \right)^2 + 3 \left( \frac{a_\eta}{a} \right)^2, \tag{28.25}
\]

so that

\[
R_{00} = -\frac{3}{a^2} \left( a_{\eta\eta} a - a_\eta^2 \right), \quad R^0_0 = g^{00} R_{00} = -\frac{3}{a^2} \left( a_{\eta\eta} a - a_\eta^2 \right). \tag{28.27}
\]

The other components can be determined by relating them to the three-dimensional curvature tensor \(^{(3)} R_{ij}\) which has the simple form (28.7). So we calculate

\[
R_{ij} = R_{\muij}^\mu = R_{kij}^k + R_{0ij}^0
\]

\[
= \left( ^{(3)} R_{ij} - \Gamma_{i0}^0 \Gamma_{j0}^0 + \Gamma_{ij}^0 \Gamma_{k0}^k + R_{0ij}^0 \right). \tag{28.28}
\]

Inserting

\[
R_{0ij}^0 = \partial_0 \Gamma_{ij}^0 - \partial_i \Gamma_{0j}^0 - \partial_j \Gamma_{0i}^0 - \Gamma_{0i}^0 \Gamma_{0j}^0 + \Gamma_{ij}^0 \Gamma_{00}^0 + \Gamma_{ij}^0 \Gamma_{00}^0,
\]

\[
\left( ^{(3)} R_{ij} = k \frac{2}{a^2} g_{ij} \right) \tag{28.30}
\]

and the above Christoffel symbols (28.21) gives

\[
R_{ij} = -\frac{1}{a^2} \left( 2ka^2 + a_\eta^2 + aa_{\eta\eta} \right) g_{ij} \tag{28.31}
\]

and thus a curvature scalar

\[
R = g^{00} R_{00} + g^{ij} R_{ij} = -\frac{1}{a^2} \left[ 3 a_\eta^2 \left( a_{\eta\eta} a - a_\eta^2 \right) \right] - \frac{3}{a^4} \left( 2ka^2 + a_\eta^2 + aa_{\eta\eta} \right)
\]

\[
= \frac{-6}{a^3} \left( a_{\eta\eta} + ka \right). \tag{28.32}
\]

In \( D' \) space dimensions, the result is

\[
R = g^{00} R_{00} + g^{ij} R_{ij} = -\frac{1}{a^2} \left[ \frac{D'}{a^2} \left( a_{\eta\eta} a - a_\eta^2 \right) \right] - \frac{D'}{a^4} \left( (D' - 1)ka^2 + a_\eta^2 + aa_{\eta\eta} \right)
\]

\[
= \frac{-D'}{a^3} \left[ 2a_\eta + k(D' - 1)a \right], \tag{28.33}
\]
while
\[
R^2_{\mu\nu} = R_{0}^0 R_{0}^0 + R_{ij}^i R_{ij}^j \\
= \frac{D'^2}{a^8} (aa_{\eta\eta} - a_\eta^2)^2 + \frac{D'}{a^8} [(D' - 1)ka^2 + a_\eta^2 + aa_{\eta\eta}]^2. \tag{28.34}
\]

For a static universe in \( D = D' + 1 \) spacetime dimensions, we obtain
\[
R^2 = \frac{k^2}{a^4} D'^2 (D' - 1)^2, \quad R^2_{\mu\nu} = \frac{k^2}{a^4} D'^2 (D' - 1)^2. \tag{28.35}
\]

**Action and Field Equation**

In the absence of matter, the *Einstein-Hilbert action* of the gravitational field is
\[
\mathcal{A} = \int d^4 x \sqrt{-g} \mathcal{L} = -\frac{1}{2\kappa} \int d^4 x \sqrt{-g} (R + 2\lambda), \tag{28.36}
\]
were \( \kappa \) is related to Newton’s gravitational constant
\[
G_N \approx 6.673 \cdot 10^{-8} \text{ cm}^3 \text{g}^{-1} \text{s}^{-2} \tag{28.37}
\]
by
\[
\frac{1}{\kappa} = \frac{c^3}{8\pi G_N}. \tag{28.38}
\]

A natural length scale of gravitational physics is the *Planck length*, which can be formed from combinations of Newton’s gravitational constant (28.37), the light velocity \( c \approx 3 \times 10^{10} \text{ cm/s} \), and Planck’s constant \( \hbar \approx 1.05459 \times 10^{-27} \):
\[
l_P = \left( \frac{c^3}{G_N\hbar} \right)^{-1/2} \approx 1.615 \times 10^{-33} \text{ cm}. \tag{28.39}
\]

It is the Compton wavelength \( l_P \equiv \hbar/m_Pc \) associated with the Planck mass
\[
m_P = \left( \frac{c\hbar}{G_N} \right)^{1/2} \approx 2.177 \times 10^{-5} \text{g} = 1.22 \times 10^{22} \text{MeV}/c^2. \tag{28.40}
\]

The constant \( 1/\kappa \) in the action (28.36) can be expressed in terms of the Planck length as
\[
\frac{1}{\kappa} = \frac{\hbar}{8\pi l_P^2}. \tag{28.41}
\]

If we add to (28.36) a matter action and vary to combined action with respect to the metric \( g_{\mu\nu} \), we obtain the *Einstein equation*
\[
\frac{1}{\kappa} \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - \lambda g_{\mu\nu} \right) = T_{\mu\nu}, \tag{28.42}
\]
where $T_{\mu\nu}$ is the energy-momentum tensor of matter. The constant $\lambda$ is called the cosmological constant.

It is believed to arise from the zero-point oscillations of all quantum fields in the universe. A single field contributes to the Lagrangian density $\mathcal{L}$ in (28.36) a term $-\Lambda \equiv -\lambda/\kappa$ which is typically of the order of $\hbar/l_P^4$. For bosons, the sign is positive, for fermions negative, reflecting the filling of all negative-energy in the vacuum. A constant of this size is much larger than the present experimental estimate. In the literature one usually finds estimates for the dimensionless quantity

$$\Omega_{\lambda 0} \equiv \frac{\lambda c^2}{3H_0^2}$$

where $H_0$ is the Hubble constant, whose inverse is roughly the lifetime of the universe

$$H_0^{-1} \approx 14 \times 10^9 \text{ years.}$$

(28.44)

It governs the speed of expansion of the universe as a function of the distance from us (Hubble’s law). Present fits to distant supernovae and other cosmological data yield the estimate [2]

$$\Omega_{\lambda 0} \equiv 0.68 \pm 0.10.$$  

(28.45)

As a result, the experimental number for $\Lambda$ is

$$\Lambda = \frac{\lambda}{\kappa} = \Omega_{\lambda 0} \frac{3H_0^2 l_P^2}{c^2} \approx 10^{-122} \frac{\hbar}{l_P^4}. $$

(28.46)

Such a small prefactor can only arise from an almost perfect cancellation of the contributions of boson and fermion fields. This cancellation is the main reason for postulating a broken supersymmetry in the universe, in which every boson has a fermionic counterpart. So far, the known particle spectra show no trace of such a symmetry. Thus there is need to explain it by some other not yet understood mechanism.

The simplest model of the universe governed by the action (28.36) is called Friedmann model or Friedmann universe.

### 28.0.2 Friedmann Model

Inserting (28.27) and (28.32) into the 00-component of the Einstein equation (28.42), we obtain the equation for the energy

$$\frac{3}{a^2} \left( a_t^2 + ka^2 \right) - \lambda = \kappa T_0^0.$$  

(28.47)

In terms of the cosmic standard time $t$, the general equation reads

$$3 \left[ \left( \frac{a_t}{a} \right)^2 + k c^2 \right] - \lambda c^2 = c^2 \kappa T_0^0. $$

(28.48)
The simplest Friedmann model is based on an energy-momentum tensor $T^0_0$ of an ideal pressureless gas of mass density $\rho$:

$$T^\nu_\mu = c\rho u^\nu u_\mu,$$ (28.49)

where $u^\mu$ are the four-vectors of the unit vectors of the particle velocities $u^\mu = (\gamma, \gamma v/c)$ with $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$. In a gas is assumed to be at rest in our comoving coordinates, so that only the $T^0_0$-component is nonzero:

$$T^0_0 = c\rho,$$ (28.50)

This component is invariant under the time transformation (28.18).

As a fortunate accident, this component of the Einstein equation has no $a_m a^m$ term. Thus we may simply study the first-order differential equation

$$\frac{3}{a^2} (a_\eta^2 + ka^2) - \lambda = c\kappa \rho.$$ (28.51)

Since the total volume of the universe is $2\pi a^3$, we can express $\rho$ in terms of the total mass $M$ as follows

$$\rho = \frac{M}{2\pi^2 a^3}.$$ (28.52)

In this way we arrive at the differential equation

$$\frac{3}{a^2} (a_\eta^2 + ka^2) - \lambda = \frac{\kappa Mc}{2\pi^2} = \frac{4G_N M}{\pi c^2 a^3},$$ (28.53)

This equation of motion can also be obtained in another way. We express the action (28.36) in terms of $a(\eta)$ using the equation (28.65) for $R$. We use the volume (28.10) and the relation (28.18) to rewrite the integration measure as

$$\int d^4x \sqrt{-g} = \int dt (4) S^a = 2\pi^2 \int d\eta a^4(\eta),$$ (28.54)

so that

$$\mathcal{J} = \frac{2\pi^2}{2\kappa} \int d\eta \left[ 6a(a_m + ka) - 2\lambda a^4 \right] = \frac{2\pi^2}{\kappa} \int d\eta \left[ -3a_\eta^2 + 3ka^2 - \lambda a^4 \right].$$ (28.55)

The second expression arises from the first by a partial integration and ignoring the boundary terms which do not influence the equation of motion. The above matter is described by the action

$$m_{\mathcal{A}} = -\int d^4x \sqrt{-\bar{g}} \rho = -2\pi^2 \int d\eta \frac{Mc}{2\pi^2} a(\eta).$$ (28.56)

Variation in $a$ yields the Euler-Lagrange equation

$$\frac{1}{\kappa} \left[ 6(a_m + ka) - 4a^3\lambda \right] - \frac{Mc}{2\pi^2} = 0.$$ (28.57)
Multiplying this with \( a_\eta \), and integration over \( \eta \) yields the pseudo-energy conservation law

\[
3(a_\eta^2 + ka^2) - a^4\lambda - \frac{\kappa Mc}{2\pi^2}a = \text{const},
\]

(28.58)
in agreement with (28.53) for a vanishing pseudo-energy.

This equation may also be written as

\[
a_\eta^2 + ka^2 - a_{\text{max}} a - \frac{\lambda}{3} a^4 = 0,
\]

(28.59)

where

\[
a_{\text{max}} \equiv \frac{\kappa M c}{6\pi^2} = \frac{4G_N M}{3\pi c^2}.
\]

(28.60)

This looks like the energy conservation law for a point particle of mass 2 in an effective potential of the universe

\[
V_{\text{univ}}(a) = ka^2 - a_{\text{max}} a - \frac{\lambda}{3} a^4,
\]

(28.61)
at zero total energy. The potential is plotted for the spherical case \( k = 1 \) in Fig. 28.1.

![Figure 28.1](image)

**Figure 28.1** Potential of closed Friedman universe as a function of the reduced radius \( a/a_{\text{max}} \) for \( \lambda a_{\text{max}}^2 = 0.1 \). Note the metastable minimum which leads to a possible solution \( a \equiv a_0 \). A tunneling process to the right leads to an expanding universe.

Friedmann neglects the cosmological constant and considers the equation

\[
a_\eta^2 + ka^2 - a_{\text{max}} a = 0.
\]

(28.62)
The solution of the differential equation for this trajectory is found by direct integration. Assuming \( k = 1 \) we obtain

\[
\eta = \int \frac{da}{\sqrt{-V_{\text{univ}}(a)}} = \int \frac{da}{\sqrt{-(a - a_{\text{max}}/2)^2 + a_{\text{max}}^2/4}} = -\arccos \frac{2a}{a_{\text{max}}}. \quad (28.63)
\]
With the initial condition \(a(0) = 0\), this implies

\[
a(\eta) = \frac{a_{\text{max}}}{2} (1 - \cos \eta)
\]  

(28.64)

Integrating Eq. (28.18), we find the relation between \(\eta\) and the physical (=proper) time

\[
t = \frac{1}{c} \int d\eta a(\eta) = \frac{a_{\text{max}}}{2c} \int d\eta (1 - \cos \eta) = \frac{a_{\text{max}}}{2c} (\eta - \sin \eta).
\]  

(28.65)

The solution \(a(t)\) is the cycloid pictured in Fig. 28.2. The radius of the universe bounces periodically with period \(t_0 = \pi a_{\text{max}}/c\) from zero to \(a_{\text{max}}\). Thus it emerges from a big bang, expands with a decreasing expansion velocity due to the gravitational attraction, and recontracts to a point.

\[\text{Figure 28.2 Radius of universe as a function of time in Friedman universe, measured in terms of the period } t_0 \equiv \pi a_{\text{max}}/c. \text{ (solid curve=closest, dashed curve=hyperbolic, dotted curve =parabolic). The curve for the closed universe is a cycloid.}\]

Certainly, for high densities the solution is inapplicable since the pressureless ideal gas approximation (28.50) breaks down.

Consider now the case of negative curvature with \(k = -1\). Then the differential equation (28.62) reads

\[
a_{\eta}^2 - a^2 - a_{\text{max}}a - \frac{\lambda}{3}a^4 = 0,
\]  

(28.66)

In order to compare the curves we shall again introduce a mass parameter \(M\) and rewrite the density as in (28.52), although \(M\) has no longer the meaning of the total mass of the universe (which is now infinite). The solution for \(\lambda = 0\) is now

\[
a(\eta) = \frac{a_{\text{max}}}{2} (\cosh \eta - 1)
\]  

(28.67)

\[
t = \frac{a_{\text{max}}}{2c} (\sinh \eta - \eta).
\]  

(28.68)

The solution is again depicted in Fig. 28.2. After a big bang, the universe expands forever, although with decreasing speed, due to the gravitational pull. The quantity \(a_{\text{max}}\) is no longer the maximal radius nor is \(t_0\) the period.
Consider finally the parabolic case $k = 0$, where the equation of motion (28.62) reads
\[ a^2_\eta - a_{\text{max}} a - \frac{\lambda}{3} a^4 = 0, \] (28.69)
where $M$ is a mass parameter defined as before in the negative curvature case. Now the solution for $\lambda = 0$ is simply
\[ \eta = 2 \left( \frac{a}{a_{\text{max}}} \right), \] (28.70)
which is inverted to
\[ a(\eta) = a_{\text{max}} \frac{\eta^2}{4}. \] (28.71)
Now we solve (28.18) with
\[ t = \frac{a_{\text{max}}}{12c} \eta^3, \] (28.72)
so that
\[ a(t) = \left( \frac{9}{4} a_{\text{max}} \right)^{1/3} (ct)^{2/3}. \] (28.73)
This solution is simply the continuation of leading term in the previous two solutions to large $t$.

28.0.3 Free Scalar Field in Slowly Evolving Robertson-Walker-Friedmann Universe

The action of a real scalar field on the background of an arbitrary gravitational field is [7, 8]
\[ \mathcal{A} = \int d^{D+1}x \sqrt{-g} \left[ \frac{1}{2} g^{\mu \nu} \partial_\mu \varphi \partial_\nu \varphi - V(\varphi) \right]. \] (28.74)
For a massive scalar field the most general harmonic potential $V(\varphi)$ is
\[ V(\varphi) = \frac{\xi R \varphi^2}{2} + \frac{M^2 \varphi^2}{2} \] (28.75)
where $R$ is the scalar curvature of spacetime, $M$ is a mass of scalar field $\varphi$. The parameter $\xi$ is usually assumed to lie in the interval [8] $0 \leq \xi \leq \frac{1}{\xi}$. Note that we use natural units with $c = \hbar = k_B = 1$, where $k_B$ is the Boltzmann constant.

A free particle has $\xi = 0$ [14]. Without additional work we treat here the general case of arbitrary $\xi$. The value $\xi = (D - 1)/4D$ makes the massless scalar action conformally invariant.
The free energy of the system with the temperature $T \equiv 1/\beta$ is

$$F = -T \ln Z$$

(28.76)

where the partition function $Z$ is given by the functional integral

$$Z = \int D\varphi e^{-\mathcal{A}^e}$$

(28.77)

where $\mathcal{A}^e$ is the euclidean action obtained from the Lorentzian one (28.74) by the substitution $t \rightarrow -i\tau$. The functional integral is performed over all fields $\varphi$ which are periodic in the imaginary time $\tau$ with period $\beta$. For the static metric with $g_{0i} = 0$, the action reads

$$\mathcal{A}^e = \frac{1}{2} \int d\tau d^3x \sqrt{\gamma} a^3 \left[ (\partial_\tau \varphi)^2 + \frac{1}{a^2} \gamma^{ij} (\partial_i \varphi) (\partial_j \varphi) + M^2 \varphi^2 \right],$$

(28.78)

where we have abbreviated the spatial part of the metric $(^3g)_{ij}$ by $\gamma_{ij}$, and introduced the effective mass

$$M^2 = \frac{6k\xi}{a^2} + m^2.$$

(28.79)

We shall the Friedmann universe to evolve so slowly that it can be assumed to be static for the purpose of our calculation. Later in Appendix 28C we shall specify precisely the condition under which the slowness assumption is valid. Under this assumption, we can calculate the path integral (28.77) by expanding the scalar field $\varphi(x)$ into eigenfunctions of the time-independent three-dimensional Laplace-Beltrami operator $\Delta^{(3)}$

$$\Delta^{(3)} = \frac{1}{\sqrt{\gamma}} \frac{\partial}{\partial x^i} \left( \sqrt{\gamma} \gamma^{ij} \frac{\partial}{\partial x^j} \right).$$

(28.80)

These are given by the spherical harmonics $\Phi_{nlm}$

$$\Delta^{(3)} \Phi_{nlm} = -(n^2 - 1)\Phi_{nlm}.$$  

(28.81)

The quantum numbers $nlm$ have the same ranges (and meaning) as in the hydrogen spectrum: $n = 1, 2, 3, \ldots$ is the principal quantum number, $l$ the quantum number of total angular momentum, running from $l = 0$ to $n - 1$, an $m$ is the azimuthal quantum number running from $m = -l$ to $l$. The functions $\Phi_{nlm}$ can be expressed in terms of the usual spherical harmonics $Y_{lm}$ as follows:

$$\Phi_{nlm} \equiv \Pi_{n}^{l}(r)Y_{lm}(\theta, \phi),$$

(28.82)

where $\Pi_{n}^{l}$ are the “Fock” harmonics [29];

$$\Pi_{n}^{l}(r) = \frac{\sin^l r}{d(cos r)^{l+1}} d^{l+1} \cos nr.$$  

(28.83)
The orthonormality reaction is
\[
\int d^3x \sqrt{\gamma} \Phi^*_{nlm}(x) \Phi_{nlm'}(x) = \delta_{nn'} \delta_{ll'} \delta_{mm'},
\] (28.84)

where \( \Phi^*_{nlm} \) is the complex conjugate of \( \Phi_{nlm} \). By partial integration and the eigenvalue equation (28.81) we find the integral:
\[
\int d^3x \sqrt{\gamma} \gamma^{i\beta} \partial_\beta \Phi_{nlm} \partial_\beta \Phi_{nlm'} = -(n^2 - 1) \delta_{nn'} \delta_{ll'} \delta_{mm'}.
\] (28.85)

More information about the eigenfunctions in the case of the hyperbolic and flat 3-spaces of constant curvature can be found in Refs. [7, 8, 14, 25, 26].

The fields are now expanded into the eigenfunctions as follows
\[
\varphi(x) = \frac{1}{2} \sum_{nlm} [a_{nlm}(\tau) \Phi_{nlm}(x) + c.c.],
\] (28.86)

with time-dependent coefficients satisfying periodic boundary conditions to account for a finite temperature:
\[
a_{nlm}(\tau = 0) = a_{nlm}(\tau = \beta) = 0.
\] (28.87)

Substituting (28.86) into (28.78) and using the orthogonality relations (28.84), we find the diagonalized spectral expansion of the euclidean action
\[
\mathcal{A}^e = \sum_{nlm} \frac{1}{2} \int d\tau [\dot{a}_{nlm}^2 + \omega_n^2 |a_{nlm}|^2],
\] (28.88)

where the dot denotes the differentiation with respect to \( \tau \), and
\[
\omega_n^2 = M^2 + (n^2 - 1)/a^2, \quad n = 1, 2, 3 \ldots.
\] (28.89)

define the eigenfrequencies, which read more explicitly
\[
\omega_n^2 = \frac{6\xi}{a^2} + m^2 + \frac{n^2 - 1}{a^2} = m^2 + \frac{n^2 + (6\xi - 1)}{a^2}.
\] (28.90)

The decomposition has reduced the functional integral for the partition function (28.77) to a product of simple path integrals of harmonic oscillators. It is then easy to calculate the total free energy [13, 14]
\[
F = \sum_{n=1}^{\infty} n^2 \frac{\omega_n}{2} + T \sum_{n=1}^{\infty} n^2 \ln \left( 1 - e^{-\omega_n/T} \right).
\] (28.91)

The factor \( n^2 \) in the sums counts the degeneracy of the eigenvalues in the isotropic space. The first term is the divergent zero-point energy. In a realistic theory of the
universe, the divergence must be canceled by the presence of an equal number of Fermi fields.

In the case of an open universe with \( k = 0 \) or \(-1\), the sums in (28.91) are replaced by spectral integrals.

We shall also consider the free energy in a general \( D \)-dimensional spacetime, where the eigenvalues (28.81) are

\[
\Delta_2^{(D-1)}\Phi_{nm} = -(n-1)(n + D - 3)\Phi_{nm}.
\]  

(28.92)

The \( D - 2 \)-dimensional quantum number \( m \) runs through different values. The \( D \)-dimensional free energy is therefore

\[
F^{(D)} = \sum_{n=1}^{\infty} d_n^{(D)} \frac{\omega_n^{(D)}}{2} + T \sum_{n=1}^{\infty} d_n^{(D)} \ln \left( 1 - e^{-\omega_n^{(D)}/T} \right),
\]  

(28.94)

with

\[
\omega_n^{(D)2} \equiv M^2 + \frac{(n-1)(n + D - 3)}{a^2}.
\]  

(28.95)

For \( D = 4 \), this agrees with (28.91), and for \( D = 3 \) it reduces to

\[
F^{(3)} = \sum_{n=1}^{\infty} (2n - 1) \frac{\omega_n^{(3)}}{2} + T \sum_{n=1}^{\infty} (2n - 1) \ln \left( 1 - e^{-\omega_n^{(3)}/T} \right).
\]  

(28.96)

Replacing \( n \) by \( l + 1 \), where \( l = 0, 1, 2, \ldots \) are the usual angular momenta on the surface of a sphere in three dimensions, we obtain a more familiar-looking expression for the free energy of a particle on a sphere of radius \( a \).

### 28.0.4 Zero Temperature Effective Energy

Let us first calculate the contribution at zero temperature

\[
F^{(D)} = \sum_{n=1}^{\infty} d_n^{(D)} \frac{\omega_n^{(D)}}{2} = \frac{1}{2} \sum_{n=1}^{\infty} d_n^{(D)} \sqrt{m^2 + \frac{1}{a^2} [(n-1)(n + D - 3) + 6\xi]}. 
\]  

(28.97)

Such divergent sums can be calculated by analytic regularization [16] in various ways. One is based on rewriting the square-root as

\[
\sqrt{A} = \frac{1}{\Gamma(-1/2)} \int_{0}^{\infty} \frac{d\tau}{\tau^{1/2}} e^{-\tau A}, \quad \Gamma(-1/2) = -2\sqrt{\pi},
\]  

(28.98)

and obtain

\[
F^{(D)} = -\frac{1}{4\sqrt{\pi a}} \int_{0}^{\infty} \frac{d\tau}{\tau^{3/2}} e^{-\tau M^2 a^2} \sum_{n=1}^{\infty} d_n^{(D)} e^{-[(n-1)(n + D - 3) + 6\xi]e^{-\tau\tau}}
\]  

(28.99)
The simplest results are obtained for $m = 0$, where

$$F^{(D)} = -\frac{1}{4\sqrt{\pi} \alpha} C^{(D)},$$  \hspace{1cm} (28.100)

with

$$C^{(D)} \equiv \int_{0}^{\infty} \frac{d\tau}{\tau^{3/2}} c^{(D)}(\tau), \quad c^{(D)}(\tau) \equiv \sum_{n=1}^{\infty} d^{(D)}_n e^{-[(n-1)(n+D-3)+6\xi]\tau},$$  \hspace{1cm} (28.101)

Let us calculate $c^{(D)}(\tau)$. First for $D = 4$ and $\xi = 1/6$, where

$$c^{(4)}(\tau) = \sum_{n=1}^{\infty} n^2 e^{-n^2\tau} = -\partial_{\tau} h(\tau) = -\partial_{\tau} \sum_{n=1}^{\infty} e^{-n^2\tau},$$  \hspace{1cm} (28.102)

so that, after a partial integration,

$$C^{(4)} \equiv -\frac{3}{2} \int_{0}^{\infty} \frac{d\tau}{\tau^{3/2}} \sum_{n=1}^{\infty} e^{-n^2\tau}.$$  \hspace{1cm} (28.103)

Then we use Poisson’s summation formula [14, 17, 18, 19],

$$\sum_{n=-\infty}^{\infty} f(n) = \int_{-\infty}^{\infty} d\nu \sum_{k=-\infty}^{\infty} e^{2\pi i n k} f(\nu)$$  \hspace{1cm} (28.104)

to rewrite

$$h(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} d\nu \sum_{k=-\infty}^{\infty} e^{2\pi i n k} e^{-\nu^2\tau} = \frac{1}{2} \sqrt{\frac{\pi}{\tau}} \sum_{k=-\infty}^{\infty} e^{-\pi k^2/\tau},$$  \hspace{1cm} (28.105)

and we obtain

$$C^{(D)} \equiv -\frac{3}{4} \sqrt{\pi} \int_{0}^{\infty} \frac{d\tau}{\tau^{3}} \left[ 1 + 2 \sum_{k=1}^{\infty} e^{-\pi k^2/\tau} \right].$$  \hspace{1cm} (28.106)

The leading term vanishes in analytic regularization by a fundamental rule is due to Veltman:

$$\int_{0}^{\infty} d\tau \tau^\alpha = 0,$$  \hspace{1cm} (28.107)

for all $\alpha$. The sum contributes

$$C^{(4)} \equiv \frac{3}{2} \sqrt{\pi} \sum_{k=1}^{\infty} e^{-\pi k^2/\tau} = \frac{3}{2} \sqrt{\pi} \sum_{k=1}^{\infty} \frac{1}{\pi^2 k^2} = \frac{3}{2} \sqrt{\pi} \frac{1}{\pi^4} \zeta(4),$$  \hspace{1cm} (28.108)

where $\zeta(4) \equiv \sum_{n=1}^{\infty} 1/n^4 = \pi^4/90$ is Riemann’s zeta function. Inserting this into (28.100) we obtain

$$F^{(4)} = \frac{1}{240\alpha}.$$  \hspace{1cm} (28.109)
Note that we could have arrived at the same result by a simple formal calculation

\[ F^{(4)} = \frac{1}{2} \sum_{n=1}^{\infty} n^2 \sqrt{\frac{n^2}{a^2}} = \frac{1}{2a} \zeta(-3) = \frac{1}{2a} \frac{3}{4\pi^4} \zeta(4) = \frac{1}{240a}, \tag{28.110} \]

using the analytic continuation formula for zeta function\(^1\)

\[ \zeta(z) = 2^z \pi^{-z} \sin(\pi z/2) \Gamma(1-z) \zeta(1-z), \tag{28.111} \]

which can also be written as

\[ \zeta(z) = 2^{z-1} \pi^z \zeta(1-z)/\Gamma(z) \cos \frac{z\pi}{2}. \tag{28.112} \]

A third way of doing the calculation is by using Abel-Plana summation formula \([5, 6, 7, 23]\)

\[ \left[ \sum_{n=1}^{\infty} f(n) - \int_0^{\infty} dn \ f(n) \right] = \frac{1}{2} f(0) + i \int_0^{\infty} d\nu f(i\nu) - f(-i\nu) e^{2\pi\nu - 1}, \tag{28.113} \]

which is correct if \(f(\nu)\) is regular for \(\text{Re} \, \nu \geq 0\). Note that this formula goes over into the Poisson formula (28.104) if we expand the denominator in powers of \(e^{2\pi\nu}\) and rotate the integration contour by 90 degrees in the complex \(\nu\)-plane. On the imaginary axis, \(f(\nu)\) may have poles and branch points which are passed during the integration on the right, i.e., with \(\text{Re} \, \nu > 0\). Applying this formula to the sum (28.99) at \(m = 0\), we obtain

\[ F^{(D)} = \frac{1}{4a} d_0^{(D)} \sqrt{-(D - 3) + 6\xi} \tag{28.114} \]

\[ - \frac{1}{2a} \int_0^{\infty} d\nu \frac{d_\nu^{(D)}}{e^{2\pi\nu - 1}} \sqrt{-(i\nu - 1)(i\nu + D - 3) - 6\xi + \text{h.c.}}. \]

For \(D = 4\) and \(\xi \leq 1/6\), this becomes

\[ F^{(4)} = \frac{1}{a} \int_0^{\infty} d\nu \frac{\nu^2 \sqrt{\nu^2 + 1 - 6\xi}}{e^{2\pi\nu - 1}}. \tag{28.115} \]

For \(\xi = 1/6\), the integral can be performed exactly and yields\(^2\) \(\Gamma(4)\zeta(4)/(2\pi)^4 a = 1/240a\), in agreement with (28.109) and (28.110). For \(\xi = 0\), the result is bigger by a factor 1.0933.

A fourth way is most appropriate if we want to know the free energy (28.99) for any mass \(m\) and parameter \(\xi\). Then we proceed as follows. First we write the energy as

\[ \omega_n^{(D)} = \frac{1}{a} \left( n + n_1^{(D)} \right) \left( n + n_2^{(D)} \right) \tag{28.116} \]


\(^2\)I.S. Gradshteyn and I.M. Ryzhik, op. cit., Formula 3.411.1
where
\[ n_{1,2}^{(D)} = \frac{1}{2} \left[ (D - 4) \pm \sqrt{(D - 2)^2 - 4m^2a^2 - 24\xi} \right]. \tag{28.117} \]

In \( D = 4 \) dimensions, we want to sum
\[ F^{(4)} = \frac{1}{2a} \sum_{n=1}^{\infty} n^2 \sqrt{(n + n_{1}^{(4)})(n + n_{2}^{(4)})}, \quad n_{1,2}^{(4)} = \pm \sqrt{1 - m^2a^2 - 6\xi} \equiv \pm \alpha, \tag{28.118} \]

and observe that the sum can equally well begin at \( n = 0 \). Hence we can write
\[ F^{(4)} = \frac{1}{2a} \sum_{k=0}^{\infty} \left[ (k^2 - \alpha^2)^{3/2} + \alpha^2(k^2 - \alpha^2)^{1/2} \right]. \tag{28.119} \]

To sum this we use the generalized zeta function (see Appendix 28A)
\[ \zeta(s, \alpha, \beta) \equiv \sum_{k=0}^{\infty} \frac{1}{[(k + \alpha)(k + \beta)]^s} \tag{28.120} \]

and find
\[ F^{(4)} = \frac{1}{2a} \zeta(-3/2, \alpha, -\alpha) + \alpha^2 \frac{1}{2a} \zeta(-1/2, \alpha, -\alpha). \tag{28.121} \]

Expressing \( \zeta(-3/2, \alpha, -\alpha) \) and \( \zeta(-3/2, \alpha, -\alpha) \) by their integral representations (28A.11), we recover the previous result (28.115).

### 28.0.5 Effective Action due to Zero-Mass Field at \( T = 0 \)

For reasons of general covariance, the free energy (28.28) can be attributed to an effective action due to fluctuations
\[ \mathcal{A}^k = \frac{1}{480\pi^2} \int d^4x \sqrt{-g} \left( \frac{\alpha}{16} R_{\mu\nu} R^{\mu\nu} + \frac{1 - \alpha}{36} R^2 \right). \tag{28.122} \]

The third possible quadratic term \( R_{\mu\nu\lambda\kappa} R^{\mu\nu\lambda\kappa} \) can always be expressed in terms of the other two by the four-dimensional version of the famous Gauss-Bonnet theorem:
\[ \int \sqrt{-g} \left( R_{\mu\nu\lambda\kappa} R^{\mu\nu\lambda\kappa} - 4R_{\mu\nu} R^{\mu\nu} + R^2 \right) = 0. \tag{28.123} \]

From Eqs. (28.52), (28.31), and (28.32) we find for \( k = 1 \):
\[ R^2 = \frac{36}{a^8} \left( a_{\eta\eta} + a^2 \right)^2 = \frac{36}{a^4} \left[ 1 + 2 \left( \frac{a_{\eta\eta}}{a} \right) + \left( \frac{a_{\eta\eta}}{a} \right)^2 \right], \tag{28.124} \]
\[ R_{\mu\nu} R^{\mu\nu} = \frac{9}{a^6} \left( a_{\eta\eta} - a^2 \right)^2 + \frac{4}{a^8} \left( 2a^2 + a^2 + aa_{\eta\eta} \right)^2 \tag{28.125} \]
\[ = \frac{1}{a^8} \left[ 16 + 16 \left( \frac{a_{\eta\eta}}{a} \right) + 16 \left( \frac{a_{\eta\eta}}{a} \right)^2 + 13 \left( \frac{a_{\eta\eta}}{a} \right)^4 + 13 \left( \frac{a_{\eta\eta}}{a} \right)^2 - 10 \frac{a^2 a_{\eta\eta}}{a^3} \right]. \]
Recalling the measure (28.54), we see that for a static universe, any $\alpha$ in the action (28.145) corresponds to a free energy (28.28).

If we want to determine $\alpha$, we have several options. One is to perform a gradient expansion assuming that $\eta$ is a small constant, so that we can approximate

$$A^a = \frac{1}{480\pi^2} \int d^4x \sqrt{-g} \left( \frac{\alpha}{16} R_{\mu\nu} R^{\mu\nu} + \frac{1 - \alpha}{36} R^2 \right) = \frac{1}{120} \left( 1 + \alpha \frac{a^2}{a^2} + \ldots \right).$$

(28.126)

Once $\alpha$ is determined, we are able to calculate how the fluctuation action (28.145) with (28.124) and (28.125) influences the evolution of the universe right after the big bang.

There exists an alternative way of determining $\alpha$, which will be discussed in the next section.

### 28.0.6 Separating $R^2$- and $R_{\mu\nu} R^{\mu\nu}$-Contributions

The action (28.145) can easily be calculated for a space with two euclidean dimensions of area $L^2$, and two dimensions forming the surface of a sphere of radius $a$. Then, according to (12.196) and (12.203),

$$R = \frac{1}{a^2}, \quad R_{ij} R^{ij} = \frac{4}{a^4},$$

(28.127)

so that the action (28.145) becomes

$$A^a = \frac{L^2}{480\pi^2} 4\pi \left( \frac{\alpha}{16} 4 + \frac{1 - \alpha}{36} 1 \right) = \frac{L^2}{120\pi a^2} \left( \frac{1}{36} + \frac{2}{9} \alpha \right).$$

(28.128)

Here we have used the measure

$$\int d^4x \sqrt{-g} = A^{(2)} S^a = L^2 4\pi a^2.$$  

(28.129)

Let us now calculate the effective euclidean action for this universe. This is given by

$$A^{\text{eff}} = \frac{1}{2} \int \frac{d\omega}{2\pi} \text{Tr} \log(\partial_1^2 + \partial_2^2 + \Delta^{(2)})$$

$$= -\frac{1}{2} \int_0^\infty d\tau \int_{-\infty}^{\infty} \frac{dk_1L}{2\pi} \int_{-\infty}^{\infty} \frac{dk_2L}{2\pi} \sum_{l=0}^{\infty} (2l + 1) e^{-(k_1^2 + k_2^2 + \omega_l^{(2)^2})\tau},$$

(28.130)

The frequencies are now

$$\omega_l^{(2)^2} = \frac{l(l + 1)}{a^2} + M^2 = \frac{1}{a^2} [l(l + 1) + 6\xi] + m^2.$$  

(28.131)

We have replaced the quantum number $n$ by $l + 1$, where $l$ is the angular momentum on the spherical surface. The degeneracy is $d_{l+1}^{(2)} = 2l + 1$. 

Integrating over \( k_1 \) and \( k_2 \), the effective action (28.137) becomes

\[
\mathcal{A}^{\text{eff}} = -\frac{L^2}{8\pi} \int_0^\infty \frac{d\tau}{\tau} \sum_{l=0}^\infty (2l + 1) e^{-\omega_l^{(2)2} \tau} 
\tag{28.132}
\]

The \( \tau \)-integral diverges at \( \tau = 0 \) so that we have to use dimensional regularization. Instead of integrating over the two dimensional-vector \( (k_1, k_2) \), we integrate over a \( 2 + \epsilon \) dimensional vector and find, instead of \( \sqrt{\frac{\pi}{\tau}} \) a factor \( \sqrt{\frac{\pi}{\tau}^{2+\epsilon}} \). This replaces (28.139) by

\[
\mathcal{A}^{\text{eff}} = -\frac{L^{2+\epsilon}}{8\pi^{1-\epsilon/2}} \int_0^\infty \frac{d\tau}{\tau^{2+\epsilon/2}} \sum_{l=0}^\infty (2l + 1) e^{-\omega_l^{(2)2} \tau} 
\tag{28.133}
\]

Now the integral can be done and yields

\[
\mathcal{A}^{\text{eff}} = -\frac{L^{2+\epsilon} \Gamma(-1 - \epsilon/2)}{a^{2+\epsilon} 8\pi^{1-\epsilon/2}} \sum_{l=0}^\infty (2l + 1) \left[ \omega_l^{(2)2} \right]^{1+\epsilon/2} . 
\tag{28.134}
\]

For \( m = 0 \) and \( \xi = 0 \), this is simply

\[
\mathcal{A}^{\text{eff}} = -\frac{L^{2+\epsilon} \Gamma(-1 - \epsilon/2)}{a^{2+\epsilon} 8\pi^{1-\epsilon/2}} \sum_{l=0}^\infty (2l + 1) [l(l + 1)]^{2+2\epsilon} 
\tag{28.135}
\]

If we go one step further and calculate action (28.145) in a space with three euclidean dimensions of volume \( L^3 \), and one dimension forming a circle of radius \( a \), Then (12.196) and (12.203) yield

\[
R = 0, \quad R_{ij}R^{ij} = 0, 
\tag{28.136}
\]

so that the action (28.145) vanishes. Let us see whether the same thing happens to the fluctuation energy which yields

\[
\mathcal{A}^{\text{eff}} = \frac{1}{2} \int \frac{d\omega}{2\pi} \text{Tr} \log(\partial_1^2 + \partial_2^2 + \partial_3^2 + \Delta^{(1)}) 
\tag{28.137}
\]

\[
= -\frac{1}{2} \int_0^\infty \frac{d\tau}{\tau} \int_{-\infty}^{\infty} \frac{dk_1 L}{2\pi} \int_{-\infty}^{\infty} \frac{dk_2 L}{2\pi} \int_{-\infty}^{\infty} \frac{dk_3 L}{2\pi} \sum_{l_3=-\infty}^{\infty} e^{-(k_1^2 + k_2^2 + k_3^2 + \omega_l^{(1)2})} 
\tag{28.137}
\]

The frequencies are now

\[
\omega_l^{(2)2} = \frac{l_3^2}{a^2} + M^2 = \frac{1}{a^2} (l_3^2 + 6\xi) + m^2, 
\tag{28.138}
\]

where we have replaced \( n \) by the usual magnetic quantum number\(^3\) \( l_3 = n - 1 \) which is doubly degenerate except for \( l_3 = 0 \). This is accounted for by taking the sum over \( l_3 \) from minus to plus infinity.

\(^3\)A preferable notation for the magnetic quantum number would be \( m \), but this symbol is occupied by the mass.
Integrating over $k_1, k_2,$ and $k_3$, the effective action (28.137) becomes

$$\mathcal{A}_{\text{eff}} = -\frac{L^2}{16\pi^{3/2}} \int_0^\infty \frac{d\tau}{\tau^2} \sum_{l_3=-\infty}^\infty e^{-\omega_{l_3}^2 \tau}$$  \hspace{1cm} (28.139)$$

The $\tau$-integral diverges at $\tau = 0$ so that we have to use dimensional regularization. Instead of integrating over the two dimensional-vector $(k_1, k_2, k_3)$, we integrate over a $3+\epsilon$ dimensional vector and find, instead of $\sqrt{\pi/\tau}$ a factor $\sqrt{\pi/\tau^{3+\epsilon}}$. This replaces (28.139) by

$$\mathcal{A}_{\text{eff}} = -\frac{L^{3+\epsilon}}{16\pi^{3/2-\epsilon/2}} \int_0^\infty \frac{d\tau}{\tau^{5/2+\epsilon/2}} \sum_{l_3=-\infty}^\infty e^{-\omega_{l_3}^2 \tau}$$  \hspace{1cm} (28.140)$$

Now the integral can be done and yields

$$\mathcal{A}_{\text{eff}} = -\left(\frac{L}{a}\right)^{3+\epsilon} \frac{\Gamma(-3/2-\epsilon/2)}{16\pi^{3/2-\epsilon/2}} \sum_{l_3=-\infty}^\infty \left[\omega_{l_3}^2\right]^{3/2+\epsilon/2}.$$  \hspace{1cm} (28.141)$$

For $m = 0$ and $\xi = 0$, this is simply

$$\mathcal{A}_{\text{eff}} = -\left(\frac{L}{a}\right)^{3+\epsilon} \frac{\Gamma(-3/2-\epsilon/2)}{16\pi^{3/2-\epsilon/2}} \sum_{l_3=-\infty}^\infty l_3^{3+2\epsilon} = -\left(\frac{L}{a}\right)^{3+\epsilon} \frac{\Gamma(-3/2-\epsilon/2)}{16\pi^{3/2-\epsilon/2}} 2\zeta(3)\zeta(4).$$  \hspace{1cm} (28.142)$$

Here the limit $\epsilon \rightarrow 0$ can be taken and yields

$$\mathcal{A}_{\text{eff}} = -\frac{1}{720\pi}.$$  \hspace{1cm} (28.143)$$

### 28.0.7 Another Separation of $R^2$- and $R_{\mu\nu} R^{\mu\nu}$-Contributions

The action (28.145) can also be calculated for the four-dimensional surface of a sphere of radius $a$ in $D = 5$ spacetime dimensions, where, according to (12.196) and (12.203),

$$R = \frac{3}{a^2}, \quad R_{ij} R^{ij} = \frac{12}{a^4},$$  \hspace{1cm} (28.144)$$

so that the action (28.145) becomes

$$\mathcal{A}^a = \frac{1}{480\pi^3} \frac{8\pi^2}{3} \left(\frac{\alpha}{16} 12 + \frac{1 - \alpha}{36} 9\right) = \frac{1}{720} (1 + 2\alpha).$$  \hspace{1cm} (28.145)$$

Here we have used the measure

$$\int d^4x \sqrt{-g} = (4) S^a = \frac{8\pi^2 a^4}{3}.$$  \hspace{1cm} (28.146)$$
Let us now calculate the effective euclidean action for this universe. This is given by

\[ \mathcal{A}^{\text{eff}} = \frac{1}{2} \text{Tr} \log \Delta^{(4)} = -\frac{1}{2} \int_0^\infty \frac{d\tau}{\tau} \sum_{n=1}^\infty d_n^{(5)} e^{-\omega_n^{(5)2}\tau}, \quad d_n^{(5)} = \frac{1}{6} (2n+1)(n+1)n. \]  

(28.147)

The eigenvalues are

\[ \omega_n^{(5)2} = \frac{1}{a^2} [(n - 1)(n + 2)] + M^2. \]  

(28.148)

We introduce the generalized zeta function

\[ \zeta(s, \Delta^{(4)}) \equiv \sum_{k=0}^\infty \frac{(2k+3)(k+1)k}{6} \frac{1}{\lambda_k^s}, \quad \lambda_k \equiv a^2 \omega_{k+1}^{(5)2} = k(k+3) + a^2 M^2. \]  

(28.149)

Then the effective action is simply

\[ \mathcal{A}^{\text{eff}} = -\frac{1}{2} \zeta'(0, \Delta^{(4)}). \]  

(28.150)

If we rewrite \( \lambda_k \) as \( (p + 3/2)(p - 3/2) + a^2 M^2 \) with \( p \equiv k - 1/2 \) running from \(-1/2\) to infinity, we can rewrite the sum as

\[ \zeta(s, \Delta^{(4)}) \equiv \sum_{k=0}^\infty \frac{(2k+3)(k+2)(k+1)}{6} \frac{1}{\lambda_k^s}, \quad \lambda_k \equiv a^2 \omega_{k+1}^{(5)2} = k(k+3) + a^2 M^2. \]  

(28.151)

For \( s = 0 \), this can be reexpressed in terms of the Hurwitz zeta function (see Appendix 28A)

\[ \zeta(s, \alpha) \equiv \sum_{k=0}^\infty \frac{1}{(k + \alpha)^s}. \]  

(28.152)

Writing

\[ (2k+3)(k+2)(k+1) = 2(k' + 3/2)^3 - \frac{1}{2}(k' + 3/2), \quad k' \equiv k + 3/2, \]  

(28.153)

we see that

\[ \zeta(0, \Delta^{(4)}) = \frac{1}{3} \zeta(-3, 3/2) - \frac{1}{12} \zeta(-1, 3/2), \]  

(28.154)

where we can insert

\[ \zeta(-3, 3/2) = -\frac{127}{960}, \quad \zeta(-1, 3/2) = \frac{11}{24}. \]  

(28.155)

For the derivative \( \zeta(s, \Delta^{(4)}) \) one finds [30], using the integral representation (28A.3),

\[ \zeta'(0, L) = \frac{2}{3} \zeta'(-3, 3/2) - \frac{1}{6} \zeta'(-3, 1/2) - \frac{1}{72} K + \frac{1}{12} K^2 + \psi(1/2) \left( -\frac{1}{12} K - \frac{1}{6} K^2 \right) \]

\[ + \frac{1}{24} \int_0^\infty \frac{dt}{t^4 \sinh(t/2)} \left[ \sqrt{K} \sin(\sqrt{K}t)(8Kt^3 - 46t) \right. \]

\[ + \cos(\sqrt{K}t)(-46 - 24Kt^2) + (46 - Kt^2 - 2K^2t^4) \left] \right]. \]  

(28.156)

where \( \psi(z) \) denotes the logarithmic derivative of the gamma function.
28.0.8 Finite-Temperature Effects

To perform the sum in the free energy (28.91) we rewrite the expression as

\[ F = T \sum_n n^2 \ln \left[ 2 \sinh \frac{\omega_n}{2T} \right] \]  
(28.1)

and the frequency (2.14) as

\[ \omega_n^2 = m^2 + \frac{n^2 - n_c^2}{a^2}. \]  
(28.2)

with the constant

\[ n_c^2 = (1 - 6\xi), \]  
(28.3)

For \( \xi \) in the typically assumed interval \( 0 \leq \xi \leq \frac{1}{6} \), the parameter \( n_c \) is restricted to \( 0 \leq n_c \leq 1 \). For the conformal coupling \( (\xi = 1/6) \) one has \( n_c = 0 \), and for the minimal coupling \( (\xi = 0) \) one has \( n_c = 1 \).

Due to its simplicity, we first consider a massive scalar field with conformal coupling \( (n_c = 0) \), where

\[ \omega_n^2 = m^2 + \frac{n^2}{a^2}. \]  
(28.4)

Using the dimensionless frequencies

\[ \hat{\omega}_n^2 = m^2 a^2 + n^2, \]  
(28.5)

we rewrite the formula (28.1) as

\[ F = T \sum_{n=1}^{\infty} n^2 \ln \left[ 2 \sinh \frac{\hat{\omega}_n}{2T} \right]. \]  
(28.6)

To isolate the finite-\( a \) effects, we add and subtract the flat \( a \rightarrow \infty \) -limit and write

\[ F = F_\infty + (F - F_\infty) \equiv F_\infty + F_{fa}, \]  
(28.7)

where

\[ F_\infty = T \int_0^\infty dnn^2 \ln \left[ 2 \sinh \frac{\omega_n}{2T} \right] = \int_0^\infty dnn^2 \left[ \frac{\omega_n}{2} + T \log \left( 1 - e^{-\omega_n/T} \right) \right]. \]  
(28.8)

The first integral is divergent, and will again be regularized analytically. The rules for this are explained in the textbook [14]. Using Veltman’s rule (28.107) we see that

\[ \int_0^\infty dnn^2 \omega_n = 0. \]  
(28.9)
Thus we arrive at the finite expression

\[ F_{\infty, \text{ren}} = T \int_0^\infty dn n^2 \ln \left[ 1 - \exp \left( -\sqrt{\frac{m^2}{T}}^2 + \left( \frac{n}{aT} \right)^2 \right) \right], \quad (28.10) \]

which can be integrated by parts to obtain

\[ F_{\infty, \text{ren}} = -\frac{a^3 T^4}{3} \int_0^\infty dx x^4 \frac{1}{\sqrt{\left( \frac{m}{T} \right)^2 + x^2}} \left[ \exp \left( \sqrt{\left( \frac{m}{T} \right)^2 + x^2} \right) - 1 \right]. \]

All finite-size effects are contained in the finite sum-minus-integral expression

\[ F_{\text{fa}} = T \sum_{n=1}^\infty n^2 \ln \left( 2 \sinh \frac{\omega_n}{2T} \right) - \int_0^\infty dn n^2 \ln \left( 2 \sinh \frac{\omega_n}{2T} \right). \quad (28.11) \]

To evaluate this it is convenient to use the Abel-Plana summation formula (28.113). Since \( f(0) = 0 \), we have

\[ F_{\text{fa}} = T i \int_0^\infty d\nu \frac{f(i\nu) - f(-i\nu)}{e^{2\pi\nu} - 1}, \quad (28.12) \]

with

\[ f(z) = z^2 \ln \left( 2 \sinh \frac{\sqrt{m^2 a^2 + z^2}}{2aT} \right), \quad (28.13) \]

\[ = z^2 \ln \left\{ \sqrt{\frac{m^2 a^2 + z^2}{aT}} \prod_{p=1}^\infty \left[ 1 + \left( \frac{1}{2\pi paT} \right)^2 \left( m^2 a^2 + z^2 \right) \right] \right\}. \]

The function \( f(z) \) has logarithmic branch points on the imaginary axis with constant discontinuities starting at \( z = n_p = \left[ m^2 a^2 + (2\pi paT)^2 \right]^{1/2} \) for \( p = 0, 1, 2 \ldots \). The integral (28.12) becomes therefore a sum

\[ F_{\text{fa}} = T 2\pi \int_0^\infty d\nu \nu^2 \sum_{p=0}^\infty \left[ \nu^2 - m^2 a^2 - (2\pi paT)^2 \right] \frac{1}{e^{2\pi\nu} - 1} = T 2\pi \sum_{p=0}^\infty \int_{n_p}^\infty d\nu \nu^2 \frac{1}{e^{2\pi\nu} - 1}, \quad (28.14) \]

where the prime on the sum indicates that the term with \( p = 0 \) should be counted with the weight \( 1/2 \). Note that (28.14) corresponds to a Planck distribution associated with the temperature \( T_{\text{eff}} = 1/a \). The appearance of such an effective temperature is typical for the Casimir effect \([5, 14]\).
The total renormalized free energy is
\[ F_{\text{ren}} = F_{\text{fa}} + F_{\infty,\text{ren}} = \frac{T}{4\pi^2} \int_0^\infty x^2 dx + \frac{a^3 T^4}{3} \int_0^\infty x^4 dx \left[ \exp \left( \sqrt{\frac{m^2}{T^2} + x^2} \right) - 1 \right]. \] (28.15)

The internal energy density of the fluctuations is found from
\[ \rho = \frac{1}{V} \frac{1}{2} \frac{\partial (\beta F_{\text{ren}})}{\partial \beta} = \rho_{\text{fa}} + \rho_{\infty,\text{ren}}, \] (28.16)
where \( V = 2\pi^2 a^3 \) is the volume of the closed Robertson-Walker-Friedmann universe. Explicitly:
\[ \rho_{\text{fa}} = 8\pi^2 T^4 \sum_{p=1}^\infty \frac{p^2 \left[ \left( \frac{m}{2\pi T} \right)^2 + p^2 \right]^{1/2}}{\exp \left( 4\pi^2 aT \sqrt{\left( \frac{m}{2\pi T} \right)^2 + p^2} \right) - 1} \] (28.17)
\[ \rho_{\infty,\text{ren}} = \frac{1}{2\pi^2 T^4} \int_0^\infty dx \frac{x^2 \left[ \sqrt{\frac{m^2}{T^2} + x^2} \right]}{\exp \left( \sqrt{\frac{m^2}{T^2} + x^2} \right) - 1}. \] (28.18)

It is useful to extract from (28.16)–(28.18) the energy density in the high-temperature limit \( T \gg 1/a \), which in (28.17) is obviously dominated by the term \( p = 1 \), whereas (28.18) tends to the Stefan-Boltzmann result for zero-mass bosons: \( \rho_{\infty,\text{ren}} = \pi^2 T^4 / 30 \).

To find the pressure of the fluctuation we use the formula
\[ P = -\frac{\partial F_{\text{ren}}}{\partial V} = -\frac{1}{6\pi^2 a^2 \frac{\partial F_{\text{ren}}}{\partial a}} = P_{\text{fa}} + P_{\infty,\text{ren}}, \] (28.19)
where
\[ P_{\text{fa}} = \frac{1}{3} \rho_{\text{fa}} + \frac{2}{3} T^4 \left( \frac{m}{T} \right)^2 \sum_{p=0}^\infty \frac{\left[ \left( \frac{m}{2\pi T} \right)^2 + p^2 \right]^{1/2}}{\exp \left( 4\pi^2 aT \sqrt{\left( \frac{m}{2\pi T} \right)^2 + p^2} \right) - 1}. \] (28.20)
\[ P_{\infty,\text{ren}} = \frac{1}{6\pi^2 T^4} \int_0^\infty \frac{x^4 dx}{\sqrt{\left( \frac{m}{T} \right)^2 + x^2} \left[ \exp \left( \sqrt{\left( \frac{m}{T} \right)^2 + x^2} \right) - 1 \right]}. \] (28.21)

In high-temperature limit \( T \gg 1/a \), the terms with \( p = 0, 1 \) give the main contribution to \( P_{\text{fa}} \).
28.0.9 Alternative Form

Formula (28.14) for the free energy $F_{\text{fa}}$ is not yet convenient to be applied in the low-temperature limit $T \ll 1/a$. For this purpose we use the Poisson summation formula (28.104) which may also be expressed as

$$\sum_{p=\infty}^{\infty} \tilde{c}(p) = 2\pi \sum_{l=-\infty}^{\infty} c(2\pi l)$$  \hspace{1cm} (28.22)

if $\tilde{c}(p)$ and $c(k)$ are connected by the Fourier transform

$$c(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp \tilde{c}(p)e^{-ikp}.$$  \hspace{1cm} (28.23)

In the sum (28.14) we identify

$$\tilde{c}(p) = \int_{2\pi n_p}^{\infty} dx \frac{x^2}{e^x - 1}$$  \hspace{1cm} (28.24)

so we can write

$$F_{\text{fa}} = \frac{T}{4\pi^2} \sum_{p=0}^{\infty} ' \tilde{c}(p) = \frac{T}{2\pi} \sum_{k=0}^{\infty} ' c(2\pi k).$$  \hspace{1cm} (28.25)

The expression for $c(0)$ is easily obtained

$$c(0) = \frac{1}{4\pi^2 aT} \int_{2\pi n_p}^{\infty} dx \frac{x^2\sqrt{x^2-(2\pi ma)^2}}{e^x - 1},$$  \hspace{1cm} (28.26)

This $p=0$ -part of (28.25) coincides with the finite-$a$ part of the free energy of the vacuum fluctuations $F_{\text{vf}}$:

$$F_{\text{vf}} = \sum_{n=1}^{\infty} \frac{n^2\omega_n}{2} - \int_{0}^{\infty} n^2\omega_n dn = \frac{1}{a} \frac{1}{(2\pi)^3} \int_{2\pi n_p}^{\infty} dx \frac{x^2\sqrt{x^2-(2\pi ma)^2}}{e^x - 1},$$  \hspace{1cm} (28.27)

as can be verified using the Abel-Plana summation formula (28.113). This expression is exponentially small for $ma \ll a$. Thus massive fields influence gravity only at short distances or immediately after the big bang. The more interesting case of massless particle yields once more the simple result (28.109)

$$F_{\text{vf}} = \frac{1}{120a}.$$  \hspace{1cm} (28.28)

Separating the zero-temperature term (28.27) from the sum, Eq. (28.25) takes the form

$$F_{\text{fa}} = F_{\text{vf}} + \frac{T}{2\pi} \sum_{k=1}^{\infty} c(2\pi k).$$  \hspace{1cm} (28.29)
For $k \neq 0$, the Fourier transformation (28.23) yields
\[
c(z) = \frac{1}{\pi z} \left[ -\frac{d^2}{dz^2} + (2\pi ma)^2 \right] \int_{2\pi ma}^{\infty} \frac{dx}{e^x - 1} \sin \left( z\sqrt{x^2 - (2\pi ma)^2} \right) \Bigg|_{z = \frac{2\pi ma}{e^{z/2} - 1}}.
\] (28.30)

This is the most convenient expression for dealing with the massless case below, which corresponds to the high-temperature limit in the massive case. In the opposite low-temperature limit of the massive case, it is preferable to find yet another representation of the same expression.

### 28.0.10 Low-Temperature Limit for Non-Zero Mass

For $m \neq 0$, the integral (28.30) cannot be calculated exactly and it is easier to obtain the low-temperature limit for $F_{fa}$ directly from the formula (28.11), which can be written in the form
\[
F_{fa} = F_{vf} + T \sum_{n=1}^{\infty} n^2 \ln \left[ 1 - \exp \left( -\frac{\omega_n}{T} \right) \right] - F_{\infty,\text{ren}}.
\] (28.31)

Omitting $F_{\infty,\text{ren}}$, we have in the low-temperature limit
\[
F'_{fa} = F_{vf} + T \sum_{n=1}^{\infty} n^2 \ln \left[ 1 - \exp \left( -\frac{\omega_n}{T} \right) \right] \approx F_{vf} - T \exp \left( -\frac{\omega_1}{T} \right).
\] (28.32)

Substituting (28.32) into formula (28.16) we obtain
\[
\rho' = \rho_{vf} + \frac{1}{2\pi^2 a^2} \sum_{n=1}^{\infty} \frac{n^2 \omega_n}{\exp \left( \frac{\omega_n}{aT} \right) - 1} \approx \rho_{vf} + \frac{\omega_1}{2\pi^2 a^2} \exp \left( -\frac{\omega_1}{T} \right),
\] (28.33)

where $\rho_{vf}$ is the well-known expression for the energy density for the vacuum fluctuations of the massive scalar field [5, 6, 7, 8, 14]:
\[
\rho_{vf} = \frac{1}{a^4 \pi (2\pi)^3} \int_{2\pi ma}^{\infty} dx \frac{x^2 \sqrt{x^2 - (2\pi ma)^2}}{e^x - 1}.
\] (28.34)

For the pressure we have from (28.19) and (28.32)
\[
P' = P_{vf} + \frac{1}{6\pi^2 a^5} \sum_{n=1}^{\infty} \frac{n^4 \omega_n}{\exp \left( \frac{\omega_n}{aT} \right) - 1} + P_{vt} \approx P_{vf} + \frac{1}{6\pi^2 a^5} \omega_1 \exp \left( -\frac{\omega_1}{T} \right),
\] (28.35)
where

\[ P_{vf} = \frac{1}{3} \rho_{vf} + \frac{(ma)^2}{24\pi^2 a^4} \int_{2\pi ma}^{\infty} dx \frac{x^2}{\sqrt{x^2 - (2\pi ma)^2}}. \] (28.36)

The integral in (28.36) is convergent.

**28.0.11 Massless Case**

The formulas obtained above are simplified considerably for zero mass, which is the case for radiation. Equation (28.32) for \( F'_{fa} \) reads

\[ F'_{fa} = -a^3 \frac{\pi^4}{45} T^4 + \frac{T}{4\pi} \sum_{p=0}^{\infty} \int_{4\pi^2 aT_p}^{\infty} x^2 dx \exp \left(4\pi^2 paT\right) - 1, \] (28.37)

and the energy density of the fluctuations is

\[ \rho' = \frac{\pi^2}{30} T^4 + 8\pi^2 T^4 \sum_{p=1}^{\infty} \frac{p^3}{\exp \left(4\pi^2 paT\right) - 1}. \] (28.38)

The second term here is usual black-body energy density. The first term contains the finite-\( a \) effects.

If we demand the constancy of the total energy carried by the fluctuations during the evolution of the universe, i.e., \( \rho \cdot a^4 = \text{const} \) [24], then we have from (28.38)

\[ aT = \text{const}. \] (28.39)

This is the usual relation between the temperature of the radiation and the scale factor.

It is convenient to deduce from (28.38) the energy density in the high-temperature limit \( T \gg 1/a \):

\[ \rho \approx 8\pi^2 a^4 \left( aT \right)^4 \exp \left( -4\pi^2 aT \right) + \frac{\pi^2}{30 a^4} \left( aT \right)^4. \] (28.40)

To find the pressure of the fluctuations we use the formula (28.19) and obtain

\[ P = \frac{\rho}{3}, \] (28.41)

where \( \rho \) is defined by formula (28.38). This is the usual equation of state for radiation.

As was stressed above, the formulas of the type (28.38) are not yet useful in the low-temperature limit \( T \ll 1/a \). For this purpose we use the expression (28.29) where in the massless case:

\[ F_{vf} = \frac{1}{240a}. \] (28.42)
617

Expression (28.42) is the well-known free energy of the vacuum fluctuations of the massless scalar field in the closed Friedmann universe [5, 6, 7, 8, 13, 14]. Here we find the alternative formula for the free energy

\[
F'_f = \frac{1}{240a} - T \sum_{n=1}^{\infty} \frac{1}{n} \frac{1 + \exp(-n/aT)}{\left[1 - \exp(-n/aT)\right]^3} \exp(-n/aT). \tag{28.44}
\]

Then the low-temperature limit \( T \ll 1/a \) becomes simply

\[
F'_f \approx \frac{1}{240a} - T \exp(-1/aT). \tag{28.45}
\]

This could have been derived directly from (28.31), (28.32). Substituting (28.44) into (28.16) we obtain the alternative form for the energy density

\[
\rho = \frac{1}{480\pi^2 a^4} \left[ 1 + 4 \exp(-n/aT) + \exp(-2n/aT) \right] \frac{\exp(-n/aT)}{\left[1 - \exp(-n/aT)\right]^4}, \tag{28.46}
\]

which converges fast for low temperatures. It is easy to see from this formula that the requirement of a constant total energy gives again the condition (28.39) \( T \cdot a = \) const. In the low-temperature limit \( T \ll 1/a \) we obtain the approximation

\[
\rho \approx \frac{1}{480\pi^2 a^4} + \frac{1}{2\pi^2 a^4} \exp(-1/aT). \tag{28.47}
\]

For the pressure of the fluctuations we can get with the help of (28.19) the equation of the state (28.41) where \( \rho \) is defined now by formula (28.46).

**28.0.12 Scalar Field with Non-Conformal Coupling**

Let us consider finally a massless scalar field with an arbitrary coupling \( 0 \leq \xi \leq 1/6 \). Then the frequency formula (28.2) can be written as

\[
\omega_n^2 = m^2 - \frac{n^2}{a^2} + \frac{n^2}{a^2} \equiv \tilde{m}(a)^2 + \frac{n^2}{a^2} \tag{28.48}
\]

where

\[
\tilde{m}^2(a) \equiv m^2 - \frac{n^2}{a^2} \tag{28.49}
\]

is an effective mass in the formulas to come. Let us also det \( \tilde{m}^2(a) \equiv m^2 + \nu_c^2/a^2 \).

Then we can use directly the formulas derived for the massive scalar field with conformal coupling, except that we rotate \( \nu_c \) to the imaginary value \( \nu_c \rightarrow -in_c \).
The expression for the non-regularized free energy takes the form
\[
F = T \sum_{n=1}^{\infty} n^2 \ln \left( 2 \sinh \frac{n^2 + \nu^2}{2aT} \right).
\] (28.50)

Performing the regularization similar to (28.6)–(28.14) we have finally
\[
F_{\text{fa}} = T 2\pi \int_0^\infty dn n^2 \sum_{p=0}^\infty \theta \left[ n^2 + n_c^2 - (2\pi paT)^2 \right] \frac{1}{\exp(2\pi n) - 1}
\]
\[
= T \pi (1 + 2p_c) \int_0^\infty dn \frac{n^2}{\exp(2\pi n) - 1}
\]
\[
+ T 2\pi \sum_{p=p_c+1}^{\infty} \int_0^\infty \frac{dn n^2 [\exp(2\pi n) - 1]^{-1}}{\sqrt{(2\pi paT)^2 - n^2}}
\] (28.51)

where
\[
p_c \equiv \left\lfloor \frac{n_c}{2\pi aT} \right\rfloor
\] (28.52)
is the largest integer \( \leq n_c/2\pi aT \). It plays the role of as infrared cut-off in the integrals (28.51).

The renormalized \( a \to \infty \) limit of (28.50) is
\[
F_{\infty,\text{ren}} = -\frac{a^3}{3^3} T^4 \int_0^\infty \frac{x^4 dx}{\sqrt{(\frac{\nu}{aT})^2 + x^2}} \left[ \exp \left( \sqrt{(\frac{\nu}{aT})^2 + x^2} \right) - 1 \right]
\]
\[
\to -\frac{\pi^4}{45} \frac{a^3}{3^4} (T)^4.
\] (28.53)

This is the usual black-body free energy, a result which was predictable since the mass term \( \tilde{m}(a) \) tends to zero for \( a \to \infty \). Thus, \( F_{\infty,\text{ren}} \) has the same form for any coupling \( \xi \).

The renormalized expression for the free energy can be found as the sum
\[
F_{\text{ren}} = F_{\text{fa}} + F_{\infty,\text{ren}}.
\] (28.54)

Then, with the help of the formulas (28.16) and (28.19), we obtain for the energy density \( \rho \)
\[
\rho = \frac{\pi^2}{30} T^4 + 8\pi^2 \frac{T^4}{3} \sum_{p=p_c+1}^{\infty} \frac{p^2 \left( p^2 - \left( \frac{n_c}{2\pi aT} \right)^2 \right)^{1/2}}{\exp \left( 4\pi^2 aT \cdot \sqrt{p^2 - \left( \frac{n_c}{2\pi aT} \right)^2} - 1 \right)}.
\] (28.55)

For the pressure we find the usual equation of state for the radiation: \( P = 1/3\rho \). The requirement of a constant total energy during the evolution of the universe gives again the condition (28.42).
The formulas (28.55), (28.60) are useful to estimate the high-temperature limit $T \gg 1/a$. The main contribution in the sums is given by the term with $p = p_c + 1$. The presence of the “mass” in this model does not permit an exact calculation of the coefficients $c(z)$ in (28.30). To obtain an estimate in the low-temperature limit $T \ll 1/a$, we write, by analogy with (28.35), the expression for the $F_{\text{ren}}$ in the form

$$F_{\text{ren}} = F_{\text{vf}} + T \sum_{n=n^*+1}^{\infty} n^2 \ln \left[ 1 - \exp \left( -\frac{\tilde{\omega}_n}{aT} \right) \right],$$

(28.56)

where we introduced the infrared cut-off similar to (3.53): $n^* = [n]$ and

$$n^* = \begin{cases} 
0, & n_c < 1, \\
1, & n_c = 1.
\end{cases}$$

(28.57)

The energy $F_{\text{vf}}$ is the free energy of the vacuum fluctuations

$$F_{\text{vf}} = \frac{1}{a^3 (2\pi)^4} \int_0^\infty dx \frac{x^2 \sqrt{x^2 + (2\pi n_c)^2}}{e^x - 1}.$$  

(28.58)

The corresponding energy density is

$$\rho_{\text{vf}} = \frac{1}{a^4 \pi (2\pi)^5} \int_0^\infty dx \frac{x^2 \sqrt{x^2 + (2\pi n_c)^2}}{e^x - 1}.$$  

(28.59)

with the pressure satisfying $P_{\text{vf}} = \frac{1}{3} \rho_{\text{vf}}$. In the low-temperature limit $aT/ \ll 1$, we find

$$\rho = \rho_{\text{vf}} + \frac{1}{2 \pi^2 a^3} \frac{\partial}{\partial \beta} (\beta F_{\text{ren}}) \approx \frac{(n^* + 1)^2 \omega_{n^*+1}^*}{2 \pi^2 a^3} \exp \left( -\frac{\omega_{n^*+1}^*}{T} \right),$$

(28.60)

and the pressure satisfies once more the equation of state $P = \rho/3$. The requirement of a constant total energy $\rho \cdot a^4 = \text{constant}$ leads again to the condition (3.40).

It is remarkable that the finite-size effects do not change this formula.

### 28.1 Application to Our Universe

It is interesting to estimate the value of the reduced temperature $aT$ for our universe. If we assume that the standard model of the hot universe [24], which most cosmologists believe describes the evolution of the now observable universe, we take for the present state the value $a \sim 12.48 \cdot 10^{27}$ cm. Using the temperature of the microwave radiation $T \sim 2.7\,\text{K}$, we find $aT \sim 1.5 \cdot 10^{29}$. Approximately the same sign of $aT$ is estimated for the relict neutrinos and gravitons [24]. Thus, at the present state of the evolution, our universe for these radiations is in high-temperature limit and the finite-size quantum effects are certainly unobservable. The static-universe approximation was used in this paper to extract pure finite-size effects. As is shown
in Appendix 28C, the slow-evolution approximation is good at the present time for the standard model of the hot universe. Obviously, the approximation breaks down for $a \to 0$.

Finite temperature effects may only become relevant if the thermalization was achieved during radiation dominated era.

**Appendix 28A Generalized Zeta Functions**

The Hurwitz zeta function is defined by

$$
\zeta(s, \alpha) \equiv \sum_{k=0}^{\infty} \frac{1}{(k+\alpha)^s} \tag{28A.1}
$$

For $\alpha = 1$ and $1/2$, this reduces to Riemann’s zeta function $\zeta(s, 1) = \zeta(s)$, $\zeta(s, 1/2) = (2^s - 1)\zeta(s)$. An important property of the Hurwitz zeta function is

$$
\zeta'(0, \alpha) = \log \Gamma(\alpha) - \frac{1}{2} \log 2 \pi. \tag{28A.2}
$$

For $\text{Re } s > 1$, the sum in (28A.1) converges. A representation applicable for $\text{Re } s < 1$ is obtained by performing the sum with the help of the Abel-Plana formula (28.113), which leads to the integral representation

$$
\zeta(s, \alpha) = \frac{1}{2} \alpha^{-s} + \frac{\alpha^{1-s}}{s-1} + 2 \int_0^\infty \frac{dy}{y^2 + \alpha^2} \frac{\sin \arctan(y/\alpha)}{e^{2\pi y} - 1}. \tag{28A.3}
$$

A generalized zeta function was defined in Eq. (28.120)

$$
\zeta(s, \alpha, \beta) \equiv \sum_{k=0}^{\infty} \frac{1}{[(k+\alpha)(k+\beta)]^s} \tag{28A.4}
$$

For $\alpha = \beta$, it reduces trivially to the Hurwitz zeta function: $\zeta(s, \alpha, \beta) = \zeta(2s, \alpha)$. Note that

$$
\partial_\alpha \zeta(s-1, \alpha, \beta) = -(s-1)\sum_{k=0}^{\infty} \frac{k}{[(k+\alpha)(k+\beta)]^s} - (s-1)\beta \zeta(s, \alpha, \beta) \tag{28A.5}
$$

$$
\sum_{k=0}^{\infty} \frac{k}{[(k+\alpha)(k+\beta)]^s} = -\beta \zeta(s, \alpha, \beta) - \frac{1}{s-1} \partial_\alpha \zeta(s-1, \alpha, \beta), \tag{28A.6}
$$

and

$$
\sum_{k=0}^{\infty} \frac{k^2}{[(k+\alpha)(k+\beta)]^s} = \zeta(s-1, \alpha, \beta) - \alpha \beta \zeta(s, \alpha, \beta)
+ (\alpha + \beta) \left[ \zeta(s, \alpha, \beta) + \frac{1}{s-1} \partial_\alpha \zeta(s-1, \alpha, \beta) \right] \tag{28A.7}
$$

---

5op. cit., Eq. (5).
where \( \alpha \) and \( \beta \) can be interchanged on the right-hand side.

For \( \text{Re } s > 1/2 \), the generalized zeta function (28A.4) can be reduced to an integral over the Hurwitz zeta function with the help of a simple formula used extensively by Feynman (see Appendix 28B)

\[
\frac{1}{[\text{Re } s > 1/2]} \frac{1}{\Gamma^2(s) \int_0^1 dx x^{s-1}(1-x)^{s-1}} \frac{1}{[k + x(1-x)]^{s}}, \tag{28A.8}
\]

so that

\[
\zeta(s, \alpha, \beta) = \frac{1}{\Gamma^2(s)} \int_0^1 dx x^{s-1}(1-x)^{s-1} \zeta(2s, x(1-x)) \tag{28A.9}
\]

For \( \text{Re } s < 1 \), this formula is inapplicable and we have to apply the Abel-Plana formula (28.113) to (28A.4) to derive the integral representation generalizing (28A.3):

\[
\zeta(s, \alpha, \beta) = \frac{1}{2 \alpha \beta^s} - \int_0^\infty dy \frac{1}{e^{2\pi y} - 1} \text{Im} \frac{1}{\sqrt{(iy + \alpha)(iy + \beta)}} \tag{28A.10}
\]

The typical situation in spectral sums of the type ming (28.99) is that \( \alpha + \beta \) and \( \alpha \beta \) are real. Then (28A.10) can be rewritten more explicitly as

\[
\zeta(s, \alpha, \beta) = \frac{1}{2 \alpha \beta^s} + \int_0^\infty dy \frac{\cos(s\phi/2)}{e^{2\pi y} - 1} \frac{[y^2 - \alpha \beta]}{[(y^2 - \alpha \beta)^2 + (\alpha + \beta)^2 y^2]^{s/2}} \tag{28A.11}
\]

where

\[
\tan \phi = \frac{(\alpha + \beta)y}{y^2 - \alpha \beta}, \quad \sin \phi = \frac{(\alpha + \beta)y}{\sqrt{(y^2 - \alpha \beta)^2 + (\alpha + \beta)^2 y^2}}. \tag{28A.12}
\]

For \( m = 0 \) and \( \xi = 0 \), where \( \alpha = 1, \beta = -1, \alpha + \beta = 0 \), Eq. (28.121) requires calculating \( \zeta(-3/2, 1, -1) \) and \( \zeta(-1/2, 1, -1) \). Inserting the integral representation (28A.11) into (28.121), the result agrees with what we derived before in Eq. (28.115).

**Appendix 28B Feynman’s Formula**

The elementary version of Feynman’s formula, used extensively in quantum field theory, reads

\[
\frac{1}{AB} = \int_0^1 dx \frac{1}{[Ax + B(1-x)]^2} = \int_0^1 dx \frac{1}{[A(1-x) + Bx]^2}. \tag{28B.1}
\]

It follows directly from the trivial relation

\[
\frac{1}{AB} = \frac{1}{B - A} \left( \frac{1}{A} - \frac{1}{B} \right) = \frac{1}{B - A} \int_A^B dz \frac{1}{z^2}. \tag{28B.2}
\]
For \( n \) denominators, the formula reads

\[
\frac{1}{A_1 \cdots A_n} = (n - 1)! \int_0^1 \int_0^1 dx_1 \ldots dx_n \frac{\delta \left( \sum_{i=1}^n x_i - 1 \right)}{(A_1 x_1 + \ldots + A_n x_n)^n},
\]

(28B.3)
as can be proved by induction. By differentiating both sides \( \alpha_i \) times with respect to \( A_i \), one finds more generally

\[
\frac{1}{A_1^{\alpha_1} \cdots A_n^{\alpha_n}} = \Gamma \left( \sum_{i=1}^n \alpha_i \right) \prod_{i=1}^n \Gamma \left( \alpha_i \right) \frac{\delta \left( \sum_{i=1}^n x_i - 1 \right)}{\left[ \sum_{i=1}^n x_i A_i \right]^{\sum_{i=1}^n \alpha_i}}.
\]

(28B.4)

By analytic continuation, this formula remains valid for complex values of \( \alpha_i \).

Appendix 28C Stabilization of the Friedmann Universe

In Chapters 2 and 3, the time evolution was assumed to be sufficiently slow to justify the calculation of the Casimir effect in a static metric. The standard model of the universe is not a static one [24]. The scale factor \( a \) is a time-dependent function and the scalar field in the case of conformal coupling is reduced to the harmonic oscillator with a time dependent frequency \( \omega(\eta) \) [7, 8] given by

\[
\omega^2(\eta) = m^2 a^2(\eta) + n^2,
\]

(28C.1)

where \( \eta \) is conformal time which is related to the synchronous time \( t \) by the formula \( a \text{d}\eta = dt \). The slow-evolution approximation is good if the adiabatic parameter \( \delta \) of the oscillators [7] satisfies the condition

\[
\delta \equiv \frac{1}{\omega^2} \frac{d\omega}{d\eta} = \frac{m^2 a}{\omega^3} \frac{da}{d\eta} \ll 1.
\]

(28C.2)

It is clear that in our present universe this parameter is extremely small. As an estimate we obtain, for particles with electron mass \( m_e \) in a typical cosmological model [24] compatible with present-day astronomical observations, the value \( \delta \leq 2 \cdot 10^{-30} \).

Static solutions are important also in multidimensional cosmology where models with compact static inner spaces represent solutions with spontaneous compactification [22].

In this Appendix we show that the quantum fluctuations of scalar field are capable of stabilizing the solutions of Einstein’s equations and yielding a static universe. Thus, in this case, the initial assumption of a slowly evolving universe becomes completely self-consistent.

The Einstein equations with a cosmological constant \( \Lambda \) in the case of the Robertson-Walker-Friedmann metric and in the presence of a perfect-fluid stress-energy tensor are [24]:

\[
\frac{1}{a^2} \left( \frac{da}{dt} \right)^2 = -\frac{k}{a^2} + \frac{\Lambda}{3} + \frac{8\pi}{3} \rho,
\]

(28C.3)

\[\text{\footnote{For a proof of this formula see Vol. 1 of E. Goursat and E.R. Hedrick,} A Course in Mathematical Analysis, Ginn and Co., Boston, 1904.}\]
where \( \rho \) and \( P \) are density of mass-energy and pressure of the perfect fluid. It is easy to see that by taking

\[
\rho = \frac{1}{8\pi} \left( \frac{3k}{a^2} - \Lambda \right),
\]

\[
P = \frac{1}{8\pi} \left( \Lambda - \frac{k}{a^2} \right),
\]

the universe can have a steady state. The parameters of the universe are

\[
\Lambda = 4\pi(\rho + 3P),
\]

\[
a^2 = \frac{k}{4\pi(\rho + P)}.
\]

Inserting \( \rho \) and \( P \) from the earlier results, the equations can be solved self-consistently.

It was shown in Chapter 3 that with the parameters of the standard model of the universe the quantum fluctuations are at present time in the high-temperature limit and the Casimir effect gives an exponential small correction to the energy density and pressure of the fluctuations. Thus, for a quantitative estimate of \( \rho \) and \( P \) it is enough to take for the regularized expression of (28.91) the formula (which is valid in all three cases \( k = 0, \pm 1 \))

\[
F_{\text{ren}} = T \int_0^\infty n^2 dn \ln \left( 1 - e^{-\frac{n^2}{T^2}} \right)
\]

\[
= -\frac{3}{4} \int_0^\infty \frac{n^3 \partial \omega_n}{\partial n} dn \exp \left( \frac{\omega_n}{T} \right) - 1
\]

where \( \omega_n \) is defined by (28.90). For the mass-energy density, we have in a closed universe

\[
\rho = \frac{1}{2\pi^2 a^3} \frac{\partial (\beta F_{\text{ren}})}{\partial \beta} = \frac{1}{2\pi^2 a^3} \int_0^\infty \frac{n^2 \omega_n dn}{\exp \left( \frac{\omega_n}{T} \right) - 1}.
\]

It is not difficult to calculate \( P \) using formulas (3.21), (28C.6) and the connection \( \partial \omega_n / \partial a = -(n/a) \partial \omega_n / \partial n \):

\[
P = -\frac{1}{6\pi^2 a} \int_0^\infty \frac{n^2 \partial \omega_n}{\partial n} dn \exp \left( \frac{\omega_n}{T} \right) - 1 = -\frac{1}{2\pi^2 a^3} F_{\text{ren}}.
\]

Consider two particular cases: high- and low-temperature limits, in which the temperature of the universe \( T \) is much or much smaller than the mass of scalar particles. Note through that the parameter \( T \gg 1/a \) in all cases.
I. High-temperature limit: $T \gg m$.
This is the limit of ultra-relativistic particles: $\omega_k \approx k = n/a$ where the equation of state is $P = \frac{\rho}{3}$. For the energy density we get (in restored dimension) from (28C.7)

$$\rho \left( \frac{\text{g}}{\text{cm}^3} \right) = \frac{\pi^2}{30c^5}$$

in agreement with (3.41) (up to the exponentially small Casimir correction).

The standard model of the universe gives for the present state the estimates [24] $\Lambda \leq 10^{-56}\text{cm}^{-2}$ and $a \geq 10^{28}\text{cm}$. From formulas (B.4) and (B.5) we have $\rho = \frac{\Lambda}{8\pi}$ and $a^2 = \frac{3}{2}\frac{1}{\Lambda}$. If we take now for $\rho$ the expression (B.9) then we obtain the parameters $\Lambda \sim 10^{-56}\text{cm}^{-2}$ and $a \sim 10^{28}\text{cm}$ for $T \sim 40^0\text{K}$. Thus, if ultra-relativistic particles exist at the present time in the thermodynamical equilibrium state with the temperature $T \sim 40^0\text{K}$, they can stabilize the universe. The temperature $T \sim 40^0\text{K}$ gives the upper limit for the mass of the particles to consider them as ultra-relativistic ones: $m < 10^{-8}\text{m}_e$, where $\text{m}_e$ is the electron mass. Such super-light particles are predicted in some types of unified theories, supersymmetry and supergravity [27]. Of course, this consideration is rather rough. We have shown here only the possibility in principle of a stabilization of our universe.

II. Low-temperature limit: $T \ll m$.
In this case $\omega_k \approx k^2/2m + m$ and for the mass-energy density we have

$$\rho = \frac{3\sqrt{2\pi}}{8\pi^2} m^4 e^{-\frac{m}{T}} \left( \frac{T}{m} \right)^{5/2} + \frac{\sqrt{2\pi}}{4\pi^2} m^4 e^{-\frac{m}{T}} \left( \frac{T}{m} \right)^{3/2}$$

$$\equiv \rho_1 + \rho_0,$$ (28C.10)

where $\rho_1$ is the kinetic energy density and $\rho_0$ is the rest mass density and $\rho_1 \ll \rho_0$ in this limit.

For the pressure we can get the relation $P = (2/3)\rho_1$ which coincides with the equation of state for Fermi and Bose gases of elementary particles [28]. As $\rho_0 \gg \rho_1$, $P$ we can omit $\rho_1$ and $P$ in the Einstein equations (28C.4) and for the parameters of the static universe we get the formulas $\Lambda = 4\pi\rho_0$, $a^2 = k/(4\pi\rho_0)$ which for $k = +1$ coincides with the formulas for the Einstein universe in the matter dominated era. It is impossible to get from formulas (28C.10) and (28C.5) the more or less real parameters of the observable universe at present time. The point is that formula (28C.10) relates to particles (and antiparticles) which are in thermodynamical equilibrium state and their energy density and number of particles (antiparticles) for $T \ll m$ is exponentially small. The parameters of real universe under low temperatures are defined by usual neutral matter which is not in equilibrium state already [29].

If the scale parameter $a$ is such that $\delta \ll 1$ we can use also the above formulas for $\rho$ and $P$ change also. If $a \to a_s$, $\rho \to \rho_s$ and $P \to P_s$ where $a_s$, $\rho_s$, and $P_s$ are connected with each other by formulas (28C.5), then the stabilization of the universe near $a_s$ will take place. But this state is a metastable one [24].

H. Kleinert, GRAVITY WITH TORSION
Notes and References


[16] See Section 2.17 in the textbook [14].
