Physics

物理

Summary is the best way to say "Good Bye"

Editor: Yuyang Songsheng
Date: March 24, 2017
Email: songshengyuyang@gmail.com

Version: 1.00

Contents

I Classical Mechanics 11

1 The formulation of Classical Mechanics 12
- 1.1 Lagrangian Formulation . . . 12
- 1.2 Symmetry and Conservation Laws(1) . . . 13
- 1.3 Hamilton formulation . . . 13
  - 1.3.1 Poisson Brackets . . . 14
  - 1.3.2 Canonical transformations . . . 15
  - 1.3.3 Evolution as canonical transformations . . . 16
  - 1.3.4 Liouville's theorem . . . 16
- 1.4 Symmetry and Conservation Laws(2) . . . 17
- 1.5 Hamilton-Jacobi equation . . . 17
- 1.6 Symmetry and Conservation Laws(3) . . . 18

2 Two body problem 19
- 2.1 Reduced mass and central field . . . 19
- 2.2 Kepler Problem . . . 20
- 2.3 Disintegration and collisions of particles . . . 23
- 2.4 Scattering and cross section . . . 26

3 Small Oscillation 29
- 3.1 Small oscillation in one dimensional . . . 29
- 3.2 Forced oscillation . . . 30
- 3.3 Non-linear oscillation and perturbation theory . . . 31
- 3.4 Oscillations of systems with more than one degree of freedom . . . 33

4 Motion of a Rigid Body 35
- 4.1 Angular velocity . . . 35
- 4.2 Dynamics of rigid body . . . 36
- 4.3 Eulerian angle . . . 39

II Classical Field Theory 41

5 Special relativity 42
- 5.1 The principle of relativity . . . 42
- 5.2 Relativistic Mechanics . . . 43

5.3 Relativistic Scattering . . . 45
5.3.1 Distribution function . . . 45
5.3.2 Invariant cross section . . . 46
5.3.3 Elastic scattering between two particles . . . 47

6 Classical field theory **48**
6.1 Lagrangian formulation . . . 48
6.2 Symmetry and conservation law . . . 49
6.3 Functional derivatives . . . 50
6.4 Hamiltonian formulation . . . 51
6.4.1 Poission bracket . . . 51
6.4.2 Momentum . . . 52
6.4.3 Angular momentum . . . 52

7 Classical Electrodynamics **54**
7.1 The formulation of classical electrodynamics . . . 54
7.1.1 Maxwell equations and Lorentz force . . . 54
7.1.2 Lorentz transformation of fields . . . 55
7.1.3 Energy-momentum tensor . . . 56
7.1.4 Charged particles in a given EM field . . . 57
7.2 Constant electromagnetic field . . . 59
7.2.1 Coulomb' law . . . 59
7.2.2 Multipole moments . . . 60
7.2.3 Biot-Savart law . . . 61
7.2.4 Magnetic moment . . . 61
7.3 Electromagnetic waves . . . 63
7.3.1 Electromagnetic waves . . . 63
7.3.2 Monochromatic wave . . . 63
7.3.3 Partially polarized light . . . 65
7.4 The field of moving charges . . . 67
7.4.1 Retarded potential . . . 67
7.4.2 Spectral resolution of the retarded potentials . . . 68
7.5 Radiation . . . 69
7.5.1 Far field approximation . . . 69
7.5.2 Low velocity approximation . . . 70
7.5.3 Radiation from a rapidly moving charge . . . 72
7.6 The interaction between charged particles and EM field . . . 73
7.6.1 Radiation reaction . . . 73
7.6.2 Scattering by free charges . . . 75

III General relativity **77**

8 Elementary Differential Geometry **78**
8.1 Fundamental conception on differential manifolds . . . 78

8.2 Multi linear algebra . . . 79
8.3 Vector Bundle . . . 81
8.4 Tangent vector field . . . 82
8.5 Exterior differential . . . 86
8.6 Connection . . . 92
8.7 Riemann manifold . . . 96

9 A Geometrical Description of Newton Theory 100
9.1 Introduction . . . 100
9.2 Geometry structure of Newtonian Space-time . . . 100
9.3 Geometry formulation of Newtonian gravity . . . 101
9.4 Standard formulation of Newtonian gravity . . . 102
9.5 Galilean coordinate system . . . 102
9.6 Coordinate transformation in space . . . 103

10 Geometry of Space-time 104
10.1 Hodge dual operator . . . 104
10.1.1 Hodge dual . . . 104
10.1.2 Levi-Civita tensor . . . 105
10.2 Metric-induced properties of Riemann curvature tensor . . . 105
10.3 The coordinates of observer . . . 106
10.3.1 Riemann normal coordinates . . . 106
10.3.2 The proper reference frame of an accelerated observer . . . 107
10.4 Hypersurfaces . . . 108
10.4.1 Description of hypersurfaces . . . 108
10.4.2 Integration on hypersurfaces . . . 109
10.4.3 Differentiation of tangent vector fields . . . 110
10.5 Killing vector . . . 111

11 Formulation of General Relativity 113
11.1 Basic assumptions of general relativity . . . 113
11.2 Lagrangian formulation . . . 113
11.2.1 Mechanics . . . 113
11.2.2 Field Theory . . . 114
11.2.3 General relativity . . . 114
11.3 Hamiltonian formulation . . . 115
11.3.1 3+1 decomposition . . . 115
11.3.2 Field theory . . . 116
11.3.3 General relativity . . . 117

12 Weak Field Limit 119
12.1 The linearized theory of gravity . . . 119
12.2 Nearly Newtonian gravitational fields . . . 120
12.3 Gravitational wave . . . 121

13 Mass And Angular Momentum of A Gravitating System **124**
13.1 External field of a gravitating source 124
13.2 Conservation laws for 4-momentum and angular momentum 125

IV Quantum Mechanics **128**

14 Linear Algebra **129**
14.1 Linear Vector Space 129
14.1.1 Definition 129
14.1.2 Linear independence 129
14.1.3 Inner product 130
14.1.4 Dual space 131
14.1.5 Dirac's bra and ket notation 131
14.2 Linear Operators 132
14.3 Self-Adjoint operators 133
14.4 Rigged Hilbert space 136
14.5 Unitary operators 137
14.6 Antiunitary operators 138

15 Formulation of quantum mechanics **139**
15.1 Axioms of quantum mechanics 139
15.2 Transformations of States 139
15.3 Schrödinger equation 140
15.4 Position operators 141
15.5 Momentum operators and canonical quantization 141
15.6 Momentum operators and translation of states 142
15.7 Angular momentum operators and rotation of states 142
15.8 Heisenberg picture 143
15.9 Symmetries and conservation laws 144

16 Coordinate and Momentum Representation **145**
16.1 Coordinate representation 145
16.2 Galilei transformation of Schrödinger equation 146
16.3 Probability flux and conditions on wave functions 146
16.4 Path integrals 147
16.5 Momentum representation 149
16.6 Harmonic oscillator 150
16.6.1 Algebraic solution 150
16.6.2 Solution in coordinate representation 151
16.6.3 Path integral solution 152
16.7 Quantum-mechanical particles in classical electromagnetic field 153
16.7.1 General discussion 153
16.7.2 Motion in a uniform static magnetic field 154
16.7.3 The Aharonov-Bohm effect 156

17 Angular Momentum **158**
17.1 Eigenvalues of angular momentum operator . . . 158
17.2 Orbital Angular Momentum and Spin . . . 159
17.3 Rotation operator . . . 161
17.4 Addition of angular momentum . . . 164
17.5 Tensor operators . . . 166
17.6 Spherical potential well . . . 169

18 Discrete Symmetries **172**
18.1 Space inversion . . . 172
18.2 Time reversal . . . 173

19 Approximation method **176**
19.1 Time independent perturbation theory . . . 176
19.1.1 Brillouin-Wigner perturbation theory . . . 176
19.1.2 Nondegenerate perturbation theory . . . 177
19.1.3 Degenerate perturbation theory . . . 178
19.2 Application of time independent perturbation theory in hydrogen atom . . . 179
19.2.1 Stark effect . . . 179
19.2.2 Fine structure . . . 181
19.2.3 Zeeman effect . . . 182
19.2.4 Hyperfine structure . . . 184
19.3 Time dependent perturbation theory . . . 187
19.4 Atomic Radiation . . . 190
19.5 Variational Method . . . 191

20 Many body problem **193**
20.1 Identical particles . . . 193
20.2 Non-relativistic quantum field theory . . . 195
20.2.1 Motivation and formulation of quantum field theory . . . 195
20.2.2 Particles in quantum field theory . . . 196
20.2.3 Momentum space . . . 199
20.2.4 Fermions . . . 199

21 Scattering theory **201**
21.1 Lippmann–Schwinger equation . . . 201
21.2 Born approximation . . . 203
21.3 Partial wave analysis . . . 205
21.4 Resonance . . . 208
21.4.1 Attractive potential . . . 208
21.4.2 Delta-shell potential . . . 209
21.4.3 General description of resonances . . . 211
21.5 Two-to-two scattering . . . 211
21.6 Time-dependent formulation . . . 212

V Quantum Field Theory 214

22 From classical field to quantum field 215
22.1 Motivation . . . 215
22.2 Lorentz invariance in quantum field theory . . . 216
22.3 Momentum . . . 216
22.4 Angular Momentum . . . 217
22.5 Anticommutation relation . . . 218

23 Spin 0 Field 219
23.1 Klein-Gordon field . . . 219
23.2 Canonical quantization Formulation . . . 219
23.3 Perturbation theory for canonical quantization . . . 221
23.3.1 Perturbation expansion of correlation functions . . . 221
23.3.2 Feynman diagram . . . 222
23.4 Path integral formulation . . . 223
23.4.1 Basic formulation . . . 223
23.4.2 Free field theory . . . 224
23.5 Perturbation theory for path integral quantization . . . 224
23.6 Symmetries in the functional formalism . . . 225
23.7 Cross section and the S-matrix . . . 227
23.8 LSZ reduction formula . . . 229
23.8.1 Field strength renormalization . . . 229
23.8.2 LSZ reduction formula . . . 231
23.9 Renormalization . . . 233
23.9.1 Counting of ultraviolet divergence . . . 233
23.9.2 Renormalized perturbation theory . . . 234
23.9.3 Techniques for renormalization . . . 236
23.9.4 One loop structure of ϕ^4 theory . . . 237
23.9.5 General renormalization theory . . . 239
23.10 Renormalization group . . . 241
23.10.1 Modified minimal-subtraction scheme . . . 241
23.10.2 Beta and gamma function . . . 243
23.10.3 Callen-Symanzik equation . . . 244
23.10.4 Running of coupling constants . . . 246
23.11 Spontaneous symmetry breaking . . . 248
23.11.1 Effective action . . . 248
23.11.2 Computation of the effective action . . . 249
23.11.3 The effective action as a generating functional . . . 250
23.11.4 Renormalization and symmetry . . . 251
23.11.5 Goldstone's theorem . . . 252
23.12 Linear sigma model . . . 253
23.12.1 Introduction . . . 253
23.12.2 Renormalization . . . 254

23.12.3 Effective action . . . 255
23.12.4 Functional determinants . . . 256
23.13 Non-relativistic limit . . . 258
23.13.1 Complex Klein-Gordon field . . . 258
23.13.2 Non-relativistic limit . . . 259

24 Spin 1/2 Field **260**
24.1 Representation of the Lorentz group . . . 260
24.2 Spin–statistics theorem . . . 261
24.3 Spinor field . . . 261
24.4 Lagrangians for spinor fields . . . 264
24.5 Canonical quantization of Dirac field . . . 268
24.6 Parity,time reversal and charge conjugation . . . 273
24.7 Perturbation theory for canonical quantization . . . 275
24.8 Path integral quantization . . . 276
24.8.1 Grassmann numbers . . . 276
24.8.2 Path integral formulation for free Dirac field . . . 278
24.8.3 Perturbation theory for path integral quantization . . . 279
24.9 LSZ reduction formula . . . 280
24.10 Functional determinant . . . 282

25 Vector Field **284**
25.1 Vector field . . . 284
25.2 Electromagnetic field and gauge invariance . . . 285
25.3 Canonical quantization of EM field . . . 286
25.3.1 Canonical quantization in Coulomb gauge . . . 286
25.3.2 Canonical quantization in Lorentz gauge . . . 288
25.4 Perturbation theory for canonical quantization . . . 292
25.4.1 Lagrangian of QED . . . 292
25.4.2 Coulomb gauge . . . 292
25.4.3 Lorentz Gauge . . . 295
25.5 Path integral quantization . . . 295
25.5.1 Path integral formulation for free EM field . . . 295
25.5.2 Perturbation theory for path integral quantization . . . 297
25.5.3 Ward-Takahahsi identity (1) . . . 297
25.6 Exact propagator of photon . . . 298
25.6.1 Photon self-energy . . . 298
25.6.2 Ward identities(2) . . . 299
25.7 LSZ reduction formula . . . 301
25.7.1 LSZ reduction formula and Feynman rules . . . 301
25.7.2 Ward Takahashi identity (3) . . . 303
25.8 Renormalization . . . 304
25.8.1 Renormalized quantum electrodynamics . . . 304
25.8.2 One loop structure of QED . . . 306

25.8.3 Renormalization group . . . 309
25.8.4 The magnetic moment of the electron . . . 311
25.8.5 Lamb shift . . . 313

26 Gauge Field **315**
26.1 Nonabelian gauge theory . . . 315
26.1.1 Nonabelian symmetries . . . 315
26.1.2 Nonabelian gauge theory . . . 316
26.1.3 Group representations . . . 318
26.2 Quantization of nonabelian gauge theory . . . 322
26.2.1 The path integral for nonabelian gauge theory . . . 322
26.2.2 The Feynman rules for nonabelian gauge theory . . . 323
26.3 Renormalization of nonabelian gauge theory . . . 325
26.4 Chiral gauge theories and anomalies . . . 327
26.4.1 Chiral gauge theories . . . 327
26.4.2 Anomalies in global symmetries . . . 330
26.4.3 Anomalies and the path integral for fermions . . . 331
26.5 Spontaneous breaking of gauge symmetries . . . 333
26.6 Quantization of spontaneously broken gauge theory . . . 335
26.6.1 Spontaneously broken abelian gauge theory . . . 335
26.6.2 Spontaneously broken nonabelian gauge theory . . . 337
26.7 The Standard Model . . . 339
26.7.1 Gauge and Higgs sector . . . 339
26.7.2 Lepton sector . . . 341
26.7.3 Quark sector . . . 344

VI Statistical mechanics **348**

27 Thermodynamics **349**
27.1 Introduction to thermodynamics . . . 349
27.2 Entropy formulation of thermodynamics . . . 350
27.2.1 Property of entropy function . . . 350
27.2.2 Simple problems . . . 351
27.2.3 Heat and Work . . . 352
27.3 Thermodynamic potential . . . 354
27.3.1 Energy Scheme . . . 354
27.3.2 Intensive Variables and Thermodynamic Potentials . . . 354
27.3.3 Free Energy and Maxwell Relations . . . 356
27.3.4 Gibbs Free Energy and Enthalpy . . . 356
27.3.5 Other Thermodynamic Potentials . . . 357
27.4 The Euler and Gibbs-Duhem Equations . . . 358
27.5 Thermodynamic systems with multi-components . . . 359
27.5.1 Chemical Reactions . . . 359

27.5.2 Phase Coexistence 360
27.5.3 The Clausius-Clapeyron Equation 361
27.5.4 The Coexistence Curve 361
27.5.5 Coexistence of Several Phases 361
27.5.6 The Critical Point 362

28 Principles of Statistical Mechanics and Ensembles 363
28.1 Density matrix 363
28.2 Statistical ensemble 364
28.2.1 Micro-canonical ensemble 364
28.2.2 Canonical ensemble 364
28.2.3 Grand-canonical ensemble 365
28.3 Fluctuations 366
28.3.1 Canonical Ensemble 366
28.3.2 Grand canonical ensemble 366

29 Interaction-free Systems 368
29.1 General discussion 368
29.1.1 Bose-Einstein Statistics 368
29.1.2 Fermi-Dirac Statistics 369
29.1.3 Boltzmann Statistics 369
29.2 Ideal Boltzmann Gas 370
29.2.1 Molecules without internal motion 370
29.2.2 Molecules without internal motion 372
29.3 Ideal Bose Systems 373

30 Quantum field theory approach 377

31 Phase Transitions and the Renormalization Group 378

Part I

Classical Mechanics

Chapter 1
The formulation of Classical Mechanics

1.1 Lagrangian Formulation

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt, \quad \delta q_i(t_1) = \delta q_i(t_2) = 0$$

$$\delta S = 0 \rightarrow \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = 0$$

1. If we transform the coordinates q to the Q as $q = q(Q, t)$, the new Lagrangian will be

   $$\bar{L}(Q, \dot{Q}, t) \equiv L(q(Q, t), \dot{q}(Q, \dot{Q}, t), t)$$

   We can verify that

   $$\frac{d}{dt}\frac{\partial \bar{L}}{\partial \dot{Q}} - \frac{\partial \bar{L}}{\partial Q} = 0$$

2. If $L_1 = L + \frac{d}{dt}f(q, t)$, then L and L_1 is equivalent and will generate the same dynamical equation.

Example:

1. The form of Lagrangian for an isolated system of particles in inertial frame:

   $$L = \sum_a \frac{1}{2}m_a v_a^2 - U(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots,)$$

   The equation of motion is

   $$m_i \ddot{\boldsymbol{r}}_i = -\nabla_{\boldsymbol{r}_i} U$$

   To get the form of Lagrangian for a system of interacting particles, we must assume:

   - Space and time are homogeneous and isotropic in inertial frame;
   - Galileo's relativity principle and Galilean transformation;
   - Spontaneous interaction between particles;

2. Consider a reference frame K. Suppose the K is moving with the velocity $\boldsymbol{V}(t)$ and rotating with angular velocity $\boldsymbol{\Omega}$ relative to the inertial reference frame. We use the coordinates of the mass point in K as general coordinates, i.e. $\boldsymbol{r} = (x_k, y_k, z_k)$. Then the Lagrangian of the mass point will be

   $$L = \frac{1}{2}m\boldsymbol{v}^2 + m\boldsymbol{v} \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}) + \frac{m}{2}(\boldsymbol{\Omega} \times \boldsymbol{r})^2 - m\dot{\boldsymbol{V}} \cdot \boldsymbol{r} - U$$

The equation of motion will be

$$m\frac{d\boldsymbol{v}}{dt} = -\frac{\partial U}{\partial \boldsymbol{r}} - m\dot{\boldsymbol{V}} + m(\boldsymbol{r}\times\dot{\boldsymbol{\Omega}}) + 2m(\boldsymbol{v}\times\boldsymbol{\Omega}) + m[\boldsymbol{\Omega}\times(\boldsymbol{r}\times\boldsymbol{\Omega})]$$

1.2 Symmetry and Conservation Laws(1)

Theorem 1.1 Nother's theorem

For $q_i \to q_i + \delta q_i$ and $L \to L + \delta L$, if $\delta L = \frac{df(q,\dot{q},t)}{dt}$,then we get

$$\frac{d}{dt}(\sum_i p^i \delta q_i - f) = 0 \quad (p^i = \frac{\partial L}{\partial \dot{q}_i})$$

Example: For an isolated system of particles in inertial frame,
$\delta L = 0$ when $\delta \boldsymbol{r}_i \to \boldsymbol{r}_i + \delta \boldsymbol{a}$, so

$$\frac{d}{dt}(\sum_i \boldsymbol{p}_i) = 0$$

$\delta L = 0$ when $\delta \boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{r}_i \times \delta\boldsymbol{\theta}$, so

$$\frac{d}{dt}(\sum_i \boldsymbol{r}_i \times \boldsymbol{p}_i) = 0$$

Homogeneity of time If $\frac{\partial L}{\partial t} = 0$,then we get

$$\frac{dE}{dt} = 0 \quad (E = \sum_i \dot{q}_i p^i - L)$$

1.3 Hamilton formulation

$$p^i = \frac{\partial L}{\partial \dot{q}_i}$$

$$H(q,p,t) = \sum_i p^i \dot{q}_i - L$$

$$\dot{p^i} = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p^i}$$

Example: For an isolated system of particles in inertial frame,

$$\boldsymbol{p}_i = m_i \boldsymbol{v}_i$$

$$H(q,p,t)=\sum_i \frac{p_i^2}{2m}+U(\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots)$$

$$\dot{\boldsymbol{p}}_i=-\nabla_{\boldsymbol{r}_i}U \quad \dot{\boldsymbol{r}}_i=\frac{\boldsymbol{p}_i}{m_i}$$

1.3.1 Poisson Brackets

First, we assume the bracket operation has the following properties:

$$[f,g]=-[g,f]$$

$$[\alpha_1 f_1+\alpha_2 f_2,\beta_1 g_1+\beta_2 g_2]=\alpha_1\beta_1[f_1,g_1]+\alpha_1\beta_2[f_1,g_2]+\alpha_2\beta_1[f_2,g_1]+\alpha_2\beta_2[f_2,g_2]$$

$$[f_1 f_2,g_1 g_2]=f_1[f_2,g_1]g_2+f_1 g_1[f_2,g_2]+g_1[f_1,g_2]f_2+[f_1,g_1]g_2 f_2$$

$$[f,[g,h]]+[g,[h,f]]+[h,[f,g]]=0$$

Here, f,g,h are functions of p^i,q_i,t. Then, we assume that

$$\left[q_i,p^k\right]=\delta_i^k$$

we can derive that

$$[f,g]=\sum_k\left(\frac{\partial f}{\partial q_k}\frac{\partial g}{\partial p^k}-\frac{\partial f}{\partial p^k}\frac{\partial g}{\partial q_k}\right)$$

So the Hamilton equation can be written as

$$\dot{p^i}=\left[p^i,H\right] \quad \dot{q}_i=[q_i,H]$$

And we can also get

$$\frac{df}{dt}=[f,H]+\frac{\partial f}{\partial t} \quad \frac{d}{dt}[f,g]=[\frac{df}{dt},g]+[f,\frac{dg}{dt}]$$

Example: For an isolated system of particles in inertial frame,

$$[r_{ia},p_{jb}]=\delta_{ab}\delta_{ij}$$

we define $l_a=\epsilon_{abc}r_a p_b$, then

$$[l_a,r_b]=\epsilon_{abc}r_c \quad [l_a,p_b]=\epsilon_{abc}p_c \quad [l_a,l_b]=\epsilon_{abc}l_c$$

1.3.2 Canonical transformations

In Hamiltonian mechanics, a canonical transformation is a change of canonical coordinates that preserves the form of Hamilton's equations (that is, the new Hamilton's equations resulting from the transformed Hamiltonian may be simply obtained by substituting the new coordinates for the old coordinates), although it might not preserve the Hamiltonian itself.

$$Q_i = Q_i(p, q, t) \quad P_i = P_i(p, q, t)$$

$$\dot{Q}_i = \frac{\partial H'}{\partial P_i} \quad \dot{P}_i = -\frac{\partial H'}{\partial Q_i}$$

Proposition 1.1 Canonical condition

If $(q_i, p^i, H) \to (Q_i, P^i, H)$ is a canonical transformation, then there exists a generating function $F(q_i, Q_i, t)$ satisfying that

$$\sum_i p^i \dot{q}_i - H(p^i, q_i) = \sum_i P^i \dot{Q}_i - H'(Q_i, P^i) + \frac{dF}{dt}$$

Applying Legendre transformation, we can get four kinds of generating function.

1. $$\frac{dF}{dt} = \sum_i p^i \dot{q}_i - \sum_i P^i \dot{Q}^i + (H' - H)$$

   Assume $\Phi(q_i, Q_i, t) = F$, so

   $$p^i = \frac{\partial \Phi}{\partial q_i} \quad P^i = -\frac{\partial \Phi}{\partial Q_i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

2. $$\frac{d}{dt}(F + \sum_i P^i Q_i) = \sum_i p^i \dot{q}_i + \sum_i Q_i \dot{P}^i + (H' - H)$$

   Assume $\Phi(q_i, P^i, t) = F + \sum_i P^i Q_i$, so

   $$p^i = \frac{\partial \Phi}{\partial q_i} \quad Q_i = \frac{\partial \Phi}{\partial P^i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

3. $$\frac{d}{dt}(F - \sum_i p^i q_i) = -\sum_i q_i \dot{p}^i - \sum_i P^i \dot{Q}_i + (H' - H)$$

   Assume $\Phi(p^i, Q_i, t) = F - \sum_i p^i q_i$, so

   $$q_i = -\frac{\partial \Phi}{\partial p^i} \quad P^i = -\frac{\partial \Phi}{\partial Q_i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

4.
$$\frac{d}{dt}(F+\sum_i P^iQ_i-\sum_i p^iq_i)=-\sum_i q_i\dot{p}^i+\sum_i Q_i\dot{P}^i+(H'-H)$$

Assume $\Phi(p^i,P^i,t)=F+\sum_i P^iQ_i-\sum_i p^iq_i$, so
$$q_i=-\frac{\partial\Phi}{\partial p^i}\quad Q_i=\frac{\partial\Phi}{\partial P^i}\quad H'=H+\frac{\partial\Phi}{\partial t}$$

Theorem 1.2 The invariance of Poisson Bracket

Suppose that $(q,p,H)\to(Q,P,H')$ is a canonical transformation and $f(q,p,t)=F(Q,P,t)$, $g(q,p,t)=G(Q,P,t)$, then

$$[f,g]_{q,p}=[F,G]_{Q,P}$$

As a result, the condition for canonical transformation can also be stated as

$$[Q_i,Q_j]_{q,p}=0\quad [P^i,P^j]_{p,q}=0\quad [Q_i,P^j]_{q,p}=\delta_i^j$$

1.3.3 Evolution as canonical transformations

Let q_t,p_t be the values of the canonical variables at time t, and $q_{t+\tau}$,$p_{t+\tau}$ their values at another time $t+\tau$. The latter are some functions of the former:

$$q_{t+\tau}=q(q_t,p_t,t,\tau)\quad p_{t+\tau}=p(q_t,p_t,t,\tau)$$

If these formulae are regarded as a transformation from the variables q_t,p_t to $q_{t+\tau}$, $p_{t+\tau}$, then this transformation is canonical. This is evident from the expression

$$dS=p_tdq_t+p_{t+\tau}dq_{t+\tau}-(H_{t+\tau}-H_t)dt$$

for the differential of the action $S(q_t,q_{t+\tau},t,\tau)$, taken along the true path, passing through the points q, and $q_{t+\tau}$ at times t and $t+\tau$ for a given τ. $-S$ is the generating function of the transformation. So we have the following communication relation

$$[q_{i\,t+\tau},q_{j\,t+\tau}]_{q_t,p_t}=0\quad [p^i_{t+\tau},p^j_{t+\tau}]_{q_t,p_t}=0\quad [q_{i\,t+\tau},p^j_{t+\tau}]_{q_t,p_t}=\delta_i^j$$

1.3.4 Liouville's theorem

Lemma 1

Let D be the Jacobian of the canonical transformation

$$\frac{\partial(Q_1,\cdots,Q_s,P^1,\cdots,P^s)}{\partial(q_1,\cdots,q_s,p^1,\cdots,p^s)}$$

Then we have

$$D=1$$

Theorem 1.3 Liouville's theorem

The phase-space distribution function is constant along the trajectories of the system ♣

☞ **Proof:** The phase volume is invariant under canonical transformation.The change in p and q during the motion can be regarded as a canonical transformation. Suppose that each point in the region of phase space moves in the course of time in accordance with the equations of motion of the mechanical system. The region as a whole therefore moves also, but its volume remains unchanged. □

1.4 Symmetry and Conservation Laws(2)

Suppose g is a function of p and q. If the transformation of q and p can be described as

$$q \to q + \epsilon[q, g]$$

$$p \to p + \epsilon[p, g]$$

We can prove that

$$H \to H + \epsilon[H, g]$$

So if H is invariant under the transformation, then $[H, g] = 0$, that means $\frac{dg}{dt} = 0$, i.e. g is a conserved quantity of the motion.

1.5 Hamilton-Jacobi equation

We define

$$S(q, t) = \left(\int_{q_0,t_0}^{q,t} Ldt\right)|_{extremum}$$

We can prove that

$$p = \frac{\partial S}{\partial q}, \quad H = -\frac{\partial S}{\partial t}$$

So, we have

$$-\frac{\partial S}{\partial t} = H(q, \frac{\partial S}{\partial q})$$

This is called Hamiltonian-Jacobi equation.

Suppose the complete integral of the Hamilton-Jacobi equation is

$$S = f(t, q_1, \cdots, q_s; \alpha^1, \cdots, \alpha^s) + A$$

where $\alpha^1, \cdots, \alpha^s$ and A are arbitrary constants. We effect a canonical transformation from the variables q, p to new variables, taking the function $f(t, q, \alpha)$ as the generating function, and the quantities $\alpha^1, \cdots, \alpha^s$ as the new momenta. Let the new co-ordinates be $\beta_1, \cdots, \beta_2$.

$$p^i = \frac{\partial f}{\partial q_i} \quad \beta_s = \frac{\partial f}{\partial \alpha_s} \quad H' = H + \frac{\partial f}{\partial t} = 0$$

So,

$$\alpha^s = \text{ constant }, \beta_s = \text{ constant}$$

By means of the s equations $\beta_s = \frac{\partial f}{\partial \alpha^s}$, the s coordinates q can be expressed in terms of the time and the $2s$ constants. This gives the general integral of the equations of motion.

1.6 Symmetry and Conservation Laws(3)

If S is invariant under transformation $q_i \to q_i + \delta q_i$, then

$$\delta S = (\sum_i p^i \delta q_i)|_{q_0,t_0}^{q,t} = 0$$

So, we have

$$\frac{d}{dt}(p^i \delta q_i) = 0$$

Further more, if

$$\delta S = (\sum_i p^i \delta q_i)|_{q_0,t_0}^{q,t} = f(q_i, \dot{q}_i, t)|_{q_0,t_0}^{q,t}$$

we will have conserved quantity

$$\frac{d}{dt}(p^i \delta q_i - f) = 0$$

Chapter 2
Two body problem

2.1 Reduced mass and central field

The Lagrangian for a two-body system is

$$L = \frac{1}{2}m_1\dot{\boldsymbol{r}}_1^2 + \frac{1}{2}m_2\dot{\boldsymbol{r}}_2^2 + U(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$$

Let $\boldsymbol{r} \equiv \boldsymbol{r}_1 - \boldsymbol{r}_2$ be the relative position vector and let the origin be at the centre of mass, i.e. $m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2 = 0$. These two equations give

$$\boldsymbol{r}_1 = \frac{m_2}{m_1 + m_2}\boldsymbol{r} \quad \boldsymbol{r}_2 = \frac{m_1}{m_1 + m_2}\boldsymbol{r}$$

Then, we have

$$L = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - U(r)$$

where

$$m = \frac{m_1 m_2}{m_1 + m_2}$$

is called reduced mass.The Lagrangian is formally identical with the Lagrangian of a particle of mass m moving in an external field $U(r)$ which is symmetrical about a fixed origin.

L is isotropic, so angular momentum is conserved, i.e. $\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{p} =$ const . Since $\boldsymbol{r}$ is always perpendicular to $\boldsymbol{M}$, the path of the particle lies in one plane. Using polar coordinates, we have

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - U(r)$$

And it is easy to see that

$$M = mr^2\dot{\phi} = \text{ const} \quad E = \frac{1}{2}m\dot{r}^2 + \frac{M^2}{2mr^2} + U(r) = \text{ const}$$

So,

$$\frac{dr}{dt} = \sqrt{\frac{2(E - U(r))}{m} - \frac{M^2}{m^2r^2}}$$

$$\frac{d\phi}{dr} = \frac{M}{r^2\sqrt{2m(E - U(r)) - M^2/r^2}}$$

The radial part of the motion can be regarded as taking place in one dimension in a field where the effective potential energy is

$$U_{eff} = U(r) + \frac{M^2}{2mr^2}$$

The values of r for which

$$U(r) + \frac{M^2}{2mr^2} = E$$

determine the limits of the motion as regards distance from the centre. When equation above is satisfied, the radial velocity $\dot{r}$ is zero. This does not mean that the particle comes to rest as in true one-dimensional motion, since the angular velocity is not zero. The value $\dot{r} = 0$ indicates a turning point of the path, where $r(t)$ begins to decrease instead of increasing, or vice versa. If the range in which r may vary is limited only by the condition $r \geq r_{min}$, the motion is infinite: the particle comes from, and returns to, infinity. If the range of r has two limits r_{min} and r_{max}, the motion is finite and the path lies entirely within the annulus bounded by the circles $r = r_{min}$ and $r = r_{max}$. This does not mean that the path must be a closed curve. During the time in which r varies from r_{min} to r_{max} and back, the radius vector turns through an angle

$$\Delta\phi = 2\int_{r_{min}}^{r_{max}} \frac{M}{r^2\sqrt{2m(E - U(r)) - M^2/r^2}}dr$$

The condition for the path to be closed is that this angle should be a rational fraction of 2π. There are only two types of central field in which all finite motions take place in closed paths. They are those in which the potential energy of the particle varies as $\frac{1}{r}$ or as r^2.
The presence of the centrifugal energy when $M \neq 0$, which becomes infinite as $\frac{1}{r^2}$ when $r \to 0$, generally renders it impossible for the particle to reach the centre of the field, even if the field is an attractive one. A fall of the particle to the centre is possible only if the potential energy tends sufficiently rapidly to $-\infty$ as $r \to 0$. From the inequality

$$\frac{1}{2}m\dot{r}^2 = E - U(r) - \frac{M^2}{2mr^2} > 0$$

it follows that r can take values tending to zero only if

$$[r^2U(r)]_{r\to 0} < -\frac{M^2}{2m}$$

2.2 Kepler Problem

An important class of central fields is formed by those in which the potential energy is inversely proportional to r. They include the fields of N Newtonian gravitational attraction and of Coulomb electrostatic interaction; the latter may be either attractive or repulsive.
Let us first consider an attractive field, where

$$U = -\frac{\alpha}{r}$$

with α a positive constant. The effective potential energy

$$U_{eff} = -\frac{\alpha}{r} + \frac{M^2}{2mr^2}$$

As $r \to 0$, U_{eff} tends to $+\infty$, and as $r \to \infty$ it tends to zero from negative values; for $r = \frac{M^2}{m\alpha}$ it has a minimum value

$$U_{eff,min} = -\frac{m\alpha^2}{2M^2}$$

The motion is finite for $-\frac{m\alpha^2}{2M^2} \leq E < 0$ and infinite for $E \geq 0$.
The shape of path is

$$\frac{p}{r} = 1 + e\cos\phi$$

Here,

$$p = \frac{M^2}{m\alpha} \quad e = \sqrt{1 + \frac{2EM^2}{m\alpha^2}}$$

This is the equation of a conic section with one focus at the origin; $2p$ is called the latus rectum of the orbit and e the eccentricity. Our choice of the origin is such that the point where $\phi = 0$ is the point nearest to the origin (called the perihelion).
If $E < 0$, the orbit is an ellipse and the motion is finite.

Figure 2.1: Attractive Kepler orbit with $e < 1$

The major and minor semi-axes of the ellipse are

$$a = \frac{p}{1-e^2} = \frac{\alpha}{2|E|} \quad b = \frac{p}{\sqrt{1-e^2}} = \frac{M}{\sqrt{2m|E|}}$$

The least and greatest distances from the centre of the field (the focus of the ellipse) are

$$r_{min} = \frac{p}{1+e} = a(1-e) \quad r_{max} = \frac{p}{1-e} = a(1+e)$$

The period of revolution in an elliptical orbit is

$$T = \frac{\pi ab}{\frac{1}{2}r^2\dot{\phi}} = 2\pi a^{3/2}\sqrt{\frac{m}{\alpha}} = \pi\alpha\sqrt{\frac{m}{2|E|^3}}$$

If $E > 0$, the path is a hyperbola with the origin as internal focus.

Figure 2.2: Attractive Kepler orbit with $e > 1$

The distance of the perihelion from the focus is

$$r_{min} = \frac{p}{1+e} = a(1-e)$$

where $a = \frac{p}{(1-e^2)^2} = \frac{\alpha}{2E}$ is the semiaxis of the hyperbola.
If $E = 0$, the eccentricity $e = 1$, and the particle moves in a parabola with perihelion distance $r_{min} = \frac{p}{2}$. This case occurs if the particle starts from rest at infinity.

Let us now consider motion in a repulsive field, where

$$U = \frac{\alpha}{r} \quad (\alpha > 0)$$

Here the effective potential energy is

$$U_{eff} = \frac{\alpha}{r} + \frac{M^2}{2mr^2}$$

and decreases monotonically from $+\infty$ to zero as r varies from zero to infinity. The energy of the particle must be positive, and the motion is always infinite. The calculations are exactly similar to those for the attractive field. The path is a hyperbola:

$$\frac{p}{r} = -1 + e\cos\phi$$

Figure 2.3: Repulsive Kepler orbit

The perihelion distance is

$$r_{min} = \frac{p}{-1+e} = a(1+e)$$

There is an integral of the motion which exists only in fields $U = \frac{\alpha}{r}$. It is easy to verify by direct calculation that the quantity

$$\boldsymbol{v} \times \boldsymbol{M} + \frac{\alpha \boldsymbol{r}}{r}$$

is constant. The direction of the conserved vector is along the major axis from the focus to the perihelion, and its magnitude is αe. This is most simply seen by considering its value at perihelion.

2.3 Disintegration and collisions of particles

Let us consider a spontaneous disintegration of a particle into two constituent parts. This process is most simply described in a frame of reference in which the particle is at rest before the disintegration. The law of conservation of momentum shows that the sum of the momenta of the two particles formed in the disintegration is then zero; that is, the particles move apart with equal and opposite momenta. The magnitude p_0 of either momentum is given by the law of conservation of energy:

$$E_i = E_{1i} + \frac{p_0^2}{2m_1} + E_{2i} + \frac{p_0^2}{2m_2}$$

here m_1 and m_2 are the masses of the particles, E_{1i} and E_{2i}, their internal energies, and E_i the internal energy of the original particle. If ϵ is the disintegration energy, i.e. the difference

$$\epsilon = E_i - E_{1i} - E_{2i}$$

which must obviously be positive, then

$$\epsilon = \frac{p_0^2}{2m}$$

here m is the reduced mass of the two particles.
Let us now change to a frame of reference in which the primary particle moves with velocity $\boldsymbol{V}$ before the break-up. This frame is usually called the laboratory system, or L system, in contradistinction to the centre-of-mass system, or C system, in which the total momentum is zero. Let us consider one of the resulting particles, and let $\boldsymbol{v}$ and $\boldsymbol{v}_0$ be its velocities in the L and the C system-respectively. It can be represented by

Figure 2.4: Disintegration in L and C frame

The relation between the angles θ and θ_0 in the L and C systems is evidently.

$$\tan\theta = \frac{v_0 \sin\theta_0}{V + v_0 \cos\theta_0}$$

In physical applications we are usually concerned with the disintegration of not one but many similar particles, and this raises the problem of the distribution of the resulting particles in direction, energy, etc. We shall assume that the primary particles are randomly oriented in

space, i.e. isotropically on average.
In the C system, every resulting particle has the same energy, and their directions of motion are isotropically distributed. The fraction of particles entering a solid angle element do is $\frac{do}{4\pi}$. So the distribution with respect to the angle θ_0 is

$$\frac{1}{2}\sin\theta_0 d\theta_0$$

The corresponding distributions in the L system are obtained by an appropriate transformation. For example, let us calculate the kinetic energy distribution in the L system. Since

$$v^2 = V^2 + v_0^2 + 2Vv_0\cos\theta_0$$

we have $d(v^2) = d\cos\theta_0$. So the kinetic energy can distributed uniformly over between $T_{min} = \frac{1}{2}(v_0 - V)^2$ and $T_{max} = \frac{1}{2}m(v_0 + V)^2$.

A collision between two particles is said to be elastic if it involves no change in their internal state. The collision is most simply described in the C system.The velocities of the particles before the collision are related to their velocities $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in the L system by $\boldsymbol{v}_{10} = m_2\boldsymbol{v}/(m_1 + m_2)$, $\boldsymbol{v}_{20} = -m_1\boldsymbol{v}/(m_1 + m_2)$, where $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$.
Because of the law of conservation of momentum, the momenta of the two particles remain equal and opposite after the collision, and are also unchanged in magnitude, by the law of conservation of energy. Thus, in the C system the collision simply rotates the velocities, which remain opposite in direction and unchanged in magnitude. The velocities of the two particles after the collision are

$$\boldsymbol{v}'_{10} = \frac{m_2 v\boldsymbol{n}_0}{m_1 + m_2} \quad \boldsymbol{v}'_{20} = -\frac{m_1 v\boldsymbol{n}_0}{m_1 + m_2}$$

The velocities in the L system after the collision are therefore

$$\boldsymbol{v}'_1 = \frac{m_2 v\boldsymbol{n}_0}{m_1 + m_2} + \frac{m_1\boldsymbol{v_1} + m_2\boldsymbol{v_2}}{m_1 + m_2} \quad \boldsymbol{v}'_2 = -\frac{m_1 v\boldsymbol{n}_0}{m_1 + m_2} + \frac{m_1\boldsymbol{v_1} + m_2\boldsymbol{v_2}}{m_1 + m_2}$$

Multiplying equations by m_1 and m_2 respectively, we obtain

$$\boldsymbol{p}'_1 = mv\boldsymbol{n}_0 + \frac{m_1(\boldsymbol{p}_1 + \boldsymbol{p}_2)}{m_1 + m_2} \quad \boldsymbol{p}'_2 = -mv\boldsymbol{n}_0 + \frac{m_2(\boldsymbol{p}_1 + \boldsymbol{p}_2)}{m_1 + m_2}$$

It can be represented by

Figure 2.5: Collision in L and C frame

Let us consider in more detail the case where one of the particles (m_2, say) is at rest before the

Figure 2.6: Collision with 2 at rest

collision. In that case the distance $OB = \frac{m_2 p_1}{m_1+m_2} = mv$ is equal to the radius. The vector $\vec{AB}$ is equal to the momentum $\boldsymbol{p}_1$ of the particle m_1 before the collision.
θ_1 and θ_2 can be expressed in terms of χ by

$$\tan\theta_1 = \frac{m_2 \sin\chi}{m_1 + m_2\cos\chi} \qquad \theta_2 = \frac{1}{2}(\pi - \chi)$$

The magnitudes of the velocities of the two particles after the collision in terms of χ are

$$v_1' = \frac{\sqrt{m_1^2 + m_2^2 + 2m_1 m_2 \cos\chi}}{m_1 + m_2} v \qquad v_2' = \frac{2m_1 v}{m_1 + m_2} \sin\frac{1}{2}\chi$$

If $m_1 < m_2$, the velocity of m_1 after the collision can have any direction. If $m_1 > m_2$, this particle can be deflected only through an angle not exceeding $\theta_m ax$ from its original direction. Evidently

$$\sin\theta_{max} = \frac{m_2}{m_1}$$

The collision of two particles of equal mass, of which one is initially at rest, is especially simple. In this case both B and A lie on the circle, so

Figure 2.7: Collision of equal mass

Then

$$\theta_1 = \frac{1}{2}\chi \qquad \theta_2 = \frac{1}{2}(\pi - \chi)$$

$$v_1' = v\cos\frac{1}{2}\chi \qquad v_2' = v\sin\frac{1}{2}\chi$$

After the collision the particles move at right angles to each other.

2.4 Scattering and cross section

Scattering is a general physical process where some forms of radiation, such as light, sound, or moving particles, are forced to deviate from a straight trajectory by one or more paths due to localized non-uniformities in the medium through which they pass. In classical mechanics, scattering generally refer to particle-particle collisions.
The definition of cross section of a scattering process is

$$\sigma \equiv \frac{\text{Number of Events per target}}{\text{Time} \times \text{Incident Flux}}$$

Here, the incident flux are measured in the frame of target particle. Recall that when we reduce two body problem in to one body problem, $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ is the coordinates of projectile in the frame of target. So the scattering process can be represented by the reduced mass moving in the central field.

Figure 2.8: Scattering in central field

And we have

$$\phi_0 = \int_{r_{min}}^{\infty} \frac{M}{r^2\sqrt{2m(E - U(r)) - \frac{M^2}{r^2}}} \quad \chi = |\pi - 2\phi_0|$$

Since

$$E = \frac{1}{2}mv_\infty^2 \quad M = m\rho v_\infty,$$

we can get the relation between χ and ρ. Suppose the number density of the particles is n, then the incident flux is nv_∞, number of events that particles are scattered into the solid angle $do = \sin\chi d\chi d\phi$ at (χ, ϕ) in time T is

$$d\rho\rho(\chi)d\phi n v_\infty T$$

So, we have

$$d\sigma = \rho(\chi)d\rho d\phi = \frac{\rho(\chi)}{\sin\chi}\left|\frac{d\rho}{d\chi}\right| do$$

In C system, we have $\boldsymbol{r}_1 = \frac{m_2}{m_1+m_2}\boldsymbol{r}$, so the scattering angle of particle 1 is the same as χ. While in L system (particle 2 is at rest before scattering), we must making corresponding transformation to get the right expression for cross section.

Rutherford's formula

One of the most important applications of the formulae derived above is to the scattering of charged particles in a Coulomb field. As $U = \frac{\alpha}{r}$, we have

$$\phi_0 = \arccos \frac{\alpha/mv_\infty^2\rho}{\sqrt{1+(\alpha/mv_\infty^2\rho)^2}}$$

Recall that $\chi = \frac{1}{2}(\pi - \phi_0)$, we can obtain that

$$\rho^2 = \frac{\alpha^2}{m^2v_\infty^4}\cot^2\frac{1}{2}\chi$$

and

$$d\sigma = \left(\frac{\alpha}{2mv_\infty^2}\right)^2 \frac{do}{\sin^4\frac{1}{2}\chi}$$

This is Rutherford's formula. It may be noted that the effective cross-section is independent of the sign of α, so that the result is equally valid for repulsive and attractive Coulomb fields.
Formula above gives the effective cross-section in the frame of reference in which the centre of mass of the colliding particles is at rest. The transformation to the laboratory system is effected by means of

$$\tan\theta_1 = \frac{m_2\sin\chi}{m_1 + m_2\cos\chi} \qquad \theta_2 = \frac{1}{2}(\pi - \chi)$$

For particles initially at rest, we have

$$d\sigma_2 = \left(\frac{\alpha}{mv_\infty^2}\right)^2 \frac{do_2}{\cos^3\theta_2}$$

The same transformation for the incident particles leads, in general, to a very complex formula, and we shall merely note two particular cases.
If the mass m_2 of the scattering particle is large compared with the mass m_1 of the scattered particle, then $\chi = \theta_1$ and $m = m_1$, so that

$$d\sigma_1 = \left(\frac{\alpha}{4E_1}\right)^2 \frac{do_1}{\sin^2(\frac{1}{2}\theta_1)}$$

here $E_1 = \frac{1}{2}m_1v_\infty^2$ is the energy of the incident particle.
If the masses of the two particles are equal, then by $\theta_1 = \frac{\chi}{2}$, we have

$$d\sigma_1 = \left(\frac{\alpha}{E_1}\right)^2 \frac{\cos\theta_1}{\sin^4\theta_1} do_1$$

If the particles are entirely identical, that which was initially at rest cannot be distinguished after the collision. The total effective cross-section for all particles is obtained by adding do_1 and do_2, so

$$d\sigma = \left(\frac{\alpha}{E_1}\right)^2 \left(\frac{1}{\sin^4\theta} + \frac{1}{\cos^4\theta}\right)\cos\theta do$$

Let us return to the general formula and use it to determine the distribution of the scattered particles with respect to the energy lost in the collision. When the masses of the scattered

and scattering particles are arbitrary, the velocity acquired by the latter is given in terms of the angle of scattering in the C system by

$$v_2' = \frac{2m_1}{m_1 + m_2} v_\infty \sin\frac{\chi}{2}$$

The energy acquired by 2 and lost by 1 is therefore

$$\epsilon = \frac{2m^2}{m_2} v_\infty^2 \sin^2\frac{\chi}{2}$$

Expressing $\sin\frac{\chi}{2}$ in terms of ϵ, we obtain

$$d\sigma = 2\pi \frac{\alpha^2}{m_2 v_\infty^2} \frac{d\epsilon}{\epsilon^2}$$

This is the required formula: it gives the effective cross-section as a function, of the energy loss ϵ, which takes values from zero to $\epsilon_{max} = 2m^2 v_\infty^2 / m_2$.

Chapter 3
Small Oscillation

3.1 Small oscillation in one dimensional

Let us consider the motion in one dimension, the potential energy of the particle is $V = V(q)$. If we choose the coordinate of equilibrium point as $q = 0$, then $\frac{\partial V}{\partial q}|_{q=0} = 0$. Expand $V(q)$ around $q = 0$, we have

$$V(q) = V(0) + \frac{1}{2}\frac{\partial^2 V}{\partial q^2}|_{q=0}\, q^2 + \cdots$$

If the equilibrium point is stable,we have

$$\frac{\partial^2 V}{\partial q^2}|_{q=0} \equiv V''(0) > 0$$

For small oscillation, we can neglect the higher orders of q and the Lagrangian can be written as

$$L = \frac{1}{2}m\dot{q} - \frac{1}{2}V''(0)q^2$$

The Euler-Lagrangian equation is

$$\ddot{q} + \omega_0^2 q = 0 \quad \omega_0^2 = \frac{V''(0)}{m}$$

The general solution is

$$q = A\cos(\omega_0 t + \phi)$$

A and ϕ depends on the initial condition.

If there is a damped force which is proportional to the velocity of the particle, then we have

$$\ddot{q} + \frac{1}{Q}\dot{q} + \omega_0^2 q = 0$$

If $Q > \frac{1}{2\omega_0}$, we have

$$q = Ae^{-\frac{t}{2Q}}\cos(\omega t + \phi) \quad \omega = \sqrt{\omega_0^2 - \frac{1}{4Q^2}}$$

It is called under damped oscillation.
If $Q < \frac{1}{2\omega_0}$, we have

$$q = Ae^{\lambda_+ t} + BAe^{\lambda_- t}$$

Here,

$$\lambda_{\pm} = -\frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2} - \omega_0^2}$$

It is called over damped oscillation.
If $Q = \frac{1}{2\omega_0}$, we have

$$q = C(1 + Dt)e^{-\omega_0 t}$$

It is called critical damped oscillation.

3.2 Forced oscillation

The equation of motion for forced oscillation is

$$\ddot{q} + \frac{1}{Q}\dot{q} + \omega_0^2 q = F(t)$$

The form of solution is

$$q = q_s + q_g$$

and q_g is the general solution of the homogeneous equation and q_s is an arbitrary special solution of the equation. In order to get q_s, we consider the following equation

$$\ddot{G} + \frac{1}{Q}\dot{G} + \omega_0^2 G = \delta(t - t')$$

and its solution is $G(t, t')$. Then we have

$$q_s = \int_{-\infty}^{\infty} F(t')G(t - t')dt'$$

For under damped oscillation, we have

$$G(t, t') = \begin{cases} \frac{e^{-\frac{t-t'}{2Q}}}{\omega} \sin\omega(t - t') & t > t' \\ 0 \quad t < t' \end{cases}$$

And we can verify that

$$q_s = \int_0^{\infty} F(t - t')\frac{e^{-\frac{t'}{2Q}}}{\omega} \sin\omega t' dt'$$

A special case is that

$$F(t) = F_0 \cos\Omega t$$

and we have

$$q(t) = \frac{F_0}{\sqrt{(\Omega^2 - \omega_0^2 + \frac{1}{2Q^2})^2 + \frac{\omega^2}{Q^2}}} \cos(\Omega t + \phi)$$

Here

$$\tan\phi = \frac{\frac{\Omega}{Q}}{\omega_0^2 - \Omega^2}$$

When

$$\Omega = \sqrt{\omega_0^2 - \frac{1}{2Q^2}}$$

we have

$$q_{max} = \frac{QF_0}{\omega}$$

It is called resonance.

3.3 Non-linear oscillation and perturbation theory

Consider the equation of motion

$$\ddot{q} + \omega_0^2 q + \epsilon q^3 = 0$$

Suppose ϵ is very small, then we can expand

$$q = q_0 + \epsilon q_1 + \epsilon^2 q_2 + \cdots$$

and

$$\omega = \omega_0 + \epsilon\omega_1 + \epsilon^2\omega_2 + \cdots$$

The equation of motion can be written as

$$\ddot{q} + \omega^2 q = (\omega^2 - \omega_0^2)q - \epsilon q^3$$

Let $\tau \equiv \omega t$ and $q' \equiv \frac{dq}{d\tau} = \frac{\dot{q}}{\omega}$, we have

$$q'' + q = (1 - \frac{\omega_0^2}{\omega^2})q - \frac{\epsilon}{\omega^2}q^3 = 0$$

We can solve the equation above power by power and get

$$q_0'' + q_0 = 0$$

$$q_1'' + q_1 = -\frac{q_0^3}{\omega_0^2} + \frac{2\omega_1}{\omega_0}q_0$$

and so on. When doing the perturbation, we must adjust the ω_i to avoid the resonance solution. The details will be neglect here.

Now let us consider the non-linear oscillation with drive force. The equation of motion is

$$\ddot{q} + \frac{1}{Q}\dot{q} + \omega_0^2 q + \epsilon q^3 = F_0 \cos\omega t$$

It can be rewritten as

$$\ddot{q} + \omega^2 q = -\frac{1}{Q}\dot{q} + (\omega^2 - \omega_0^2)q - \epsilon q^3 + F_0 \cos\omega t$$

We treat the right hand of the equation as perturbation, so we multiply it by a parameter μ and let it be 1 later,

$$\ddot{q} + \omega^2 q = \mu\left(-\frac{1}{Q}\dot{q} + (\omega^2 - \omega_0^2)q - \epsilon q^3 + F_0 \cos\omega t\right)$$

Concerning on the phase lagging effect, we redefine the "time" as

$$\tau \equiv \omega t - \delta$$

so,

$$q'' + q = \mu\left[-\frac{1}{Q\omega}q' + \left(1 - \frac{\omega_0^2}{\omega^2}\right)q - \frac{\epsilon}{\omega^2}q^3 + \frac{F_0}{\omega^2}\cos(\tau + \delta)\right]$$

The expansion series of q and δ are

$$q = q_0 + \mu q_1 + \mu^2 q_2 + \cdots$$

$$\delta = \delta_0 + \mu\delta_1 + \mu^2\delta_2 + \cdots$$

We can solve the equation above power by power and get

$$q_0'' + q_0 = 0$$

$$q_1'' + q_1 = -\frac{1}{Q\omega}q_0' + \left(1 - \frac{\omega_0^2}{\omega^2}\right)q_0 - \frac{\epsilon}{\omega^2}q_0^3 + \frac{F_0}{\omega^2}\cos(\tau + \delta_0)$$

and so on. The solution of the zeroth order perturbation is

$$q_0 = A_0 \cos\tau$$

Substitute it into the first order equation, we have

$$q_1''+q_1 = \left(\frac{A_0}{Q\omega} - \frac{F_0}{\omega^2}\sin\delta_0\right)\sin\tau + \left[\left(1 - \frac{\omega_0^2}{\omega^2}\right)A_0 - \frac{3\epsilon A_0^3}{4\omega^2} + \frac{F_0}{\omega^2}\cos\delta_0\right]\cos\tau - \frac{\epsilon A_0^3}{4\omega^2}\cos 3\tau$$

To avoid non-physical solution, we have

$$\sin\delta_0 = \frac{A_0\omega}{F_0 Q}$$

and

$$\left(1 - \frac{\omega_0^2}{\omega^2}\right)A_0 - \frac{3\epsilon A_0^3}{4\omega^2} + \frac{F_0}{\omega^2}\cos\delta_0 = 0$$

At last, we can get

$$A_0 = \frac{F_0}{\sqrt{\left[(\omega^2 - \omega_0^2) - \frac{3}{4}\epsilon A_0^2\right]^2 + \frac{\omega^2}{Q^2}}}$$

If we define

$$x \equiv \frac{\omega}{\omega_0} \quad y \equiv \frac{A_0\omega_0^2}{F_0}$$

the equation above can be written as

$$y^2 = \frac{1}{(x^2 - 1 - ay^2)^2 + bx^2}$$

Here

$$a = \frac{3\epsilon F_0^2}{4\omega_0^6} \quad b = \frac{1}{\omega_0^2 Q^2}$$

We than can solve for x in terms of y.,

$$x^2 = \frac{2 + 2ay - b \pm \sqrt{(b - 2 - 2ay)^2 - 4(a^2y^2 + 2ay + 1 - \frac{1}{y})}}{2}$$

So the resonance curve has two branches, corresponding to the two roots of the equation above.

Figure 3.1: Resonance curve

When the frequency of the drive force increase from left, the amplitude of oscillation will become larger and larger. But when it comes to the point of inflection, the amplitude will drop to the low-right part of the curve. When the frequency of the drive force decrease from right, the amplitude of oscillation will also become larger and larger. When it comes to the point of inflection, the amplitude will jump to the hight-left part of the curve. This effect is called hysteresis.

3.4 Oscillations of systems with more than one degree of freedom

Let's look at a system with many degrees of freedom; we have

$$L = \frac{1}{2}\sum_{i,j} T_{ij}\dot{q}_i\dot{q}_j - V\left(q_1, \ldots q_n\right)$$

Let $q_{0,i}$ be an equilibrium position and expand about this point $q_i = q_{0,i} + \eta_i$, so $\dot{q}_i = \dot{\eta}_i$. We can expand the potential energy to give

$$V\left(q_1, \ldots q_n\right) = V\left(q_{0,1}, \ldots q_{0,n}\right) + \sum_i \left(\frac{\partial V}{\partial q_i}\right)_{q_{0,i}} \eta_i + \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_{q_{0,i}} \eta_i\eta_j + \cdots$$

The first term is a constant with respect to η_i and constant terms do not affect the motion. The second term is zero, because $q_{0,i}$ is a point of equilibrium so we are left with

$$L = \frac{1}{2}\sum_{i,j}\left(T_{ij}\dot{\eta}_i\dot{\eta}_j - V_{ij}\eta_i\eta_j\right)$$

where

$$T_{ij} = T_{ij}\left(q_{0,1},\ldots q_{0,n}\right) \quad V_{ij} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_{q_{0,i}}$$

yielding the equations of motion

$$\sum_j \left(T_{ij}\ddot{\eta}_j - V_{ij}\eta_j\right) = 0$$

This is a linear differential equation with constant coefficients. We can try the solution

$$\eta_i = Ca_i e^{-i\omega t}$$

so we have

$$\sum_j \left(V_{ij}a_j - \omega^2 T_{ij}a_j\right) = 0$$

This is a matrix equation such that

$$\vec{\vec{A}} \cdot \vec{a} = 0$$

with

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_j \end{bmatrix}$$

and

$$\vec{\vec{A}} = \begin{bmatrix} V_{11} - \omega^2 T_{11} & V_{12} - \omega^2 T_{12} & \cdots \\ V_{21} - \omega^2 T_{21} & V_{22} - \omega^2 T_{22} & \cdots \\ \vdots & & \end{bmatrix}$$

This equation only has a solution is $\det \vec{\vec{A}} = 0$. This gives a nth-degree polynomial to solve for ω^2. We will get n solutions for ω^2 that we can substitute into the matrix equation and solve for a_j.

Chapter 4
Motion of a Rigid Body

4.1 Angular velocity

Suppose there are two coordinate frames. The frame 2 is rotating relatively to frame 1. If the coordinates of the particles in frame 2 is $r_2 = (x_2, y_2, z_2)$. The the coordinates of the particle in frame 1 is

$$r_1 = O(t)r_2$$

and we have

$$O^T O = I$$

So, we can derive that

$$\Omega + \Omega^T = 0 \quad \Omega \equiv O^T \frac{dO}{dt}$$

It is easy to verify that

$$\frac{dr_1}{dt} = \frac{dO}{dt} r_2 + O\frac{dr_2}{dt} = OO^T \frac{dO}{dt} r_2 + Ov_2 = O(\Omega r_2 + v_2)$$

Suppose

$$\Omega = \begin{bmatrix} 0 & -\omega_{2z} & \omega_{2y} \\ \omega_{2z} & 0 & -\omega_{2x} \\ -\omega_{2y} & \omega_{2x} & 0 \end{bmatrix}$$

So, we have

$$v_1 = O(\omega_2 \times r_2 + v_2)$$

Now, we define

$$\omega_1 \equiv O\omega_2$$

We can derive that

$$O\Omega O^T = \begin{bmatrix} 0 & -\omega_{1z} & \omega_{1y} \\ \omega_{1z} & 0 & -\omega_{1x} \\ -\omega_{1y} & \omega_{1x} & 0 \end{bmatrix}$$

So, we have

$$v_1 = \omega_1 \times r_1 + Ov_2$$

ω is the so-called angular velocity.

Note: *ω_1 is independent of the base vector we choose for frame 2.*
If we choose frame 1 differently, ω_1 will transform like an vector.

4.2 Dynamics of rigid body

Inertial tensor of rigid body

Suppose there is frame attached to the rigid body, then the coordinate of all the mass point of the rigid body is constant in this frame, i.e.

$$r_1 = r_0(t) + O(t)r_2 \quad r_2 \text{ is a constant}$$

So we have

$$v_1 = V + \omega_1 \times r_1 = V + O(\omega_2 \times r_2)$$

The kinetic energy of the rigid body is

$$T = \sum \frac{m}{2}(\boldsymbol{V} + \boldsymbol{\omega} \times \boldsymbol{r})^2 = \sum \frac{m}{2}V^2 + \sum m\boldsymbol{V} \cdot (\boldsymbol{\omega} \times \boldsymbol{r}) + \sum \frac{m}{2}(\boldsymbol{\omega} \times \boldsymbol{r})^2$$

If we choose the origin of the frame 2 to be the center of mass of the rigid body, we have

$$T = \frac{\mu V^2}{2} + \frac{1}{2}\sum m[\omega^2 r^2 - (\boldsymbol{\omega} \cdot \boldsymbol{r})^2]$$

If we define the inertial tensor as

$$I_{ik} = \sum m(x_l^2 \delta_{ik} - x_i x_k)$$

the kinetic energy of the rigid body can be rewritten as

$$T = \frac{\mu V^2}{2} + \frac{1}{2}I_{ik}\omega_i\omega_k$$

and the Lagrangian of the rigid body is

$$L = \frac{\mu V^2}{2} + \frac{1}{2}I_{ik}\omega_i\omega_k - U$$

If the body is regarded as continuous, the sum becomes an integral over the volume of the body:

$$I_{ik} = \int \rho(x_l^2 \delta_{ik} - x_i x_k)dV$$

Like any symmetrical tensor of rank two, the inertia tensor can be reduced to diagonal form by an appropriate choice of the directions of the axes x_1, x_2 x_3. These directions are called the principal axes of inertia, and the corresponding values of the diagonal components of the tensor are called the principal moments of inertia; we shall denote them by I_1, I_2, I_3. When the axes x_1, x_2 x_3 are so chosen, the kinetic energy of rotation takes the very simple form

$$T_{rot} = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$$

A body whose three principal moments of inertia are all different is called an asymmetrical top.

If two are equal ($I_1 = I_2 \neq I_3$), we have a symmetrical top. In this case the direction of one of the principal axes in the x_1x_2-plane may be chosen arbitrarily.
If all three principal moments of inertia are equal, the body is called a spherical top, and the three axes of inertia may be chosen arbitrarily as any three mutually perpendicular axes.
The determination of the principal axes of inertia is much simplified if the body is symmetrical, for it is clear that the position of the centre of mass and the directions of the principal axes must have the same symmetry as the body.
For example, if the body has a plane of symmetry, the centre of mass must lie in that plane, which also contains two of the principal axes of inertia, while the third is perpendicular to the plane.
If a body has an axis of symmetry of any order, the centre of mass must lie on that axis, which is also one of the principal axes of inertia, while the other two are perpendicular to it.
If the axis is of order higher than the second, the body is a symmetrical top. For any principal axis perpendicular to the axis of symmetry can be turned through an angle different from π about the latter, i.e. the choice of the perpendicular axes is not unique, and this can happen only if the body is a symmetrical top.

Finally, we may note one further result concerning the calculation of the inertia tensor. Although this tensor has been defined with respect to a system of co-ordinates whose origin is at the centre of mass , it may sometimes be more conveniently found by first calculating a similar tensor,

$$I'_{ik} = \sum m(x_l'^2\delta_{ik} - x'_i x'_k)$$

defined with respect to some other origin O'. If the distance OO' is represented by a vector $\boldsymbol{a}$, then $\boldsymbol{r} = \boldsymbol{r'} + \boldsymbol{a}$; since, by the definition of O, $\sum m\boldsymbol{r} = 0$, we have

$$I'_{ik} = I_{ik} + \mu(a^2\delta_{ik} - a_i a_k)$$

Using this formula, we can easily calculate I_{ik} if I'_{ik} is known.

Angular momentum

The value of the angular momentum of systems depends on the point with respect to which it is defined. In the mechanics of a rigid body, the most appropriate point to choose for this purpose is the origin of the moving system of co-ordinates, i.e. the centre of mass of the body. Then we have

$$\boldsymbol{M} = \sum m\boldsymbol{r} \times (\boldsymbol{\omega} \times \boldsymbol{r} + \boldsymbol{V}) = \sum m\left[r^2\boldsymbol{\omega} - (\omega \cdot \boldsymbol{r})\boldsymbol{r}\right]$$

or, in tensor notation,

$$M_i = I_{ik}\omega_k$$

If the axes x_1, x_2 x_3 are the same as the principal axes of inertia, we have

$$M_1 = I_1\omega_1 \quad M_2 = I_2\omega_2 \quad M_3 = I_3\omega_3$$

Equation of motion

Since a rigid body has, in general, six degrees of freedom, the general equations of motion must be six in number. They can be put in a form which gives the time derivatives of two vectors, the momentum and the angular momentum of the body. The first equation is obtained by simply summing the equations $\dot{\boldsymbol{p}} = \boldsymbol{f}$ for each particle in the body. In terms of the total momentum of the body

$$\boldsymbol{P} = \sum \boldsymbol{p} = \mu \boldsymbol{V}$$

and total force acting on it $\boldsymbol{F} = \sum \boldsymbol{f}$, we have

$$\frac{d\boldsymbol{P}}{dt} = \boldsymbol{F}$$

Although $\boldsymbol{F}$ has been defined as the sum of all the forces $\boldsymbol{f}$ acting on the various particles, including the forces due to other particles, $\boldsymbol{F}$ actually includes only external forces: the forces of interaction between the particles composing the body must cancel out, since if there are no external forces the momentum of the body, like that of any closed system, must be conserved, i.e. we must have $\boldsymbol{F} = 0$.

Let us now derive the second equation of motion, which gives the time derivative of the angular momentum $\boldsymbol{M}$. To simplify the derivation, it is convenient to choose the fixed (inertial) frame of reference in such a way that the centre of mass is at rest in that frame at the instant considered. We have

$$\dot{\boldsymbol{M}} = \frac{d}{dt} \sum \boldsymbol{r} \times \boldsymbol{p} = \sum \dot{\boldsymbol{r}} \times \boldsymbol{p} + \sum \boldsymbol{r} \times \dot{\boldsymbol{p}}$$

Our choice of the frame of reference (with $\boldsymbol{V} = 0$) means that the vectors $\dot{r}$ and $\boldsymbol{p} = m\boldsymbol{v}$ are parallel, so $\dot{\boldsymbol{r}} \times \boldsymbol{p} = 0$. We have finally

$$\frac{d\boldsymbol{M}}{dt} = \boldsymbol{K}$$

where

$$\boldsymbol{K} = \sum \boldsymbol{r} \times \boldsymbol{f}$$

Since $\boldsymbol{M}$ has been defined as the angular momentum about the centre of mass, it is unchanged when we go from one inertial frame to another. We can therefore deduce that the equation of motion, though derived for a particular frame of reference, is valid in any other inertial frame, by Galileo's relativity principle.
The vector $\boldsymbol{r} \times \boldsymbol{f}$ is called the moment of the force $\boldsymbol{f}$, and so $\boldsymbol{K}$ is the total torque, i.e. the sum of the moments of all the forces acting on the body. Like the total force, $\sum \boldsymbol{r} \times \boldsymbol{f}$ need include only the external forces: by the law of conservation of angular momentum, the sum of the moments of the internal forces in a closed system must be zero.

Euler's equation

Let $\frac{d\boldsymbol{A}}{dt}$ be the rate of change of any vector $\boldsymbol{A}$ with respect to the fixed system of co-ordinates. We have

$$\frac{d\boldsymbol{A}}{dt} = \frac{d'\boldsymbol{A}}{dt} + \boldsymbol{\omega} \times \boldsymbol{A}$$

where $\frac{d'\boldsymbol{A}}{dt}$ is the rate of change of the $\boldsymbol{A}$'s components in the body system of coordinates. So, we have

$$\frac{d'\boldsymbol{M}}{dt} + \boldsymbol{\omega} \times \boldsymbol{M} = \boldsymbol{K}$$

Suppose the principal axes of inertia are x_1, x_2 x_3, so we have

$$\begin{aligned} I_1\frac{d\omega_1}{dt} + (I_3 - I_2)\omega_2\omega_3 &= K_1 \\ I_2\frac{d\omega_2}{dt} + (I_1 - I_3)\omega_1\omega_3 &= K_2 \\ I_3\frac{d\omega_1}{dt} + (I_2 - I_1)\omega_1\omega_2 &= K_3 \end{aligned}$$

This is called Euler's equation.

4.3 Eulerian angle

The motion of a rigid body can be described by means of the three co-ordinates of its centre of mass and any three angles which determine the orientation of the axes x_1, x_2 x_3 in the moving system of coordinates relative to the fixed system X, Y, Z. These angles may often be conveniently taken as what are called Eulerian angles. The moving x_1x_2-plane intersects the

Figure 4.1: Eulerian angle

fixed XY-plane in some line ON, called the line of nodes. This line is evidently perpendicular to both the Z-axis and the x_3-axis; we take its positive direction as that of the vector product $\boldsymbol{z} \times \boldsymbol{x}_3$. We take, as the quantities defining the position of the axes x_1, x_2 x_3 relative to the axes X, Y, Z the angle θ between the Z and x_3 axes, the angle ϕ between the X-axis and ON, and the angle ψ between the x_1 and ON.

Let us now express the components of the angular velocity vector ω along the moving axes x_1, x_2 x_3 in terms of the Eulerian angles and their derivatives. To do this, we must find the components along those axes of the angular velocities $\dot{\theta}$, $\dot{\phi}$, $\dot{\psi}$. The angular velocity $\dot{\theta}$ is along the line of nodes ON. The angular velocity $\dot{\phi}$ is along the Z-axis. The angular velocity ψ is

along the x_3-axis. Collecting the components along each axis, we have

$$\begin{aligned}\omega_1 &= \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\ \omega_2 &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \\ \omega_3 &= \dot{\phi}\cos\theta + \dot{\psi}\end{aligned}$$

For a symmetrical top, by using the fact that the choice of directions of the principal axes x_1, x_2 is arbitrary for a symmetrical top. If the x_1 axis is taken along the line of nodes ON, i.e. $\psi = 0$, the components of the angular velocity are simply

$$\omega_1 = \dot{\theta} \quad \omega_2 = \dot{\phi}\sin\theta \quad \omega_3 = \dot{\phi}\cos\theta + \dot{\psi}$$

For the free motion of a symmetrical top, we take the Z-axis of the fixed system of coordinates in the direction of the constant angular momentum M of the top. The x_3-axis of the moving system is along the axis of the top; let the x_1-axis coincide with the line of nodes at the instant considered. Then the components of the vector M are

$$M_1 = I_1\dot{\theta} \quad M_2 = I_2\dot{\phi}\sin\theta \quad M_3 = I_3(\dot{\phi}\cos\theta + \dot{\psi})$$

Since the x_1-axis is perpendicular to the Z-axis, we have

$$M_1 = 0 \quad M_2 = M\sin\theta \quad M_3 = M\cos\theta$$

Comparison gives

$$\dot{\theta} = 0 \quad \dot{\phi} = \frac{M}{I_1} \quad \dot{\phi}\cos\theta + \dot{\psi} = \frac{M\cos\theta}{I_3}$$

The first of these equations gives $\theta =$ constant, i.e. the angle between the axis of the top and the direction of $\boldsymbol{M}$ is constant. The second equation gives the angular velocity of precession $\dot{\phi} = \frac{M}{I_1}$. Finally, the third equation gives the angular velocity with which the top rotates about its own axis $\omega_3 = \frac{M\cos\theta}{I_3}$.

Part II

Classical Field Theory

Chapter 5
Special relativity

5.1 The principle of relativity

First, we assume there is an upper limit of velocity of propagation of interaction c. Second, we assume that inertial reference frame are all the same in describing the law of physics. Then, we can find the invariant intervals when transforming from one inertial reference frame to another, $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$. (In the following, we assume that $c = 1$.)This transformation is called Lorentz transformation, which can be written as

$$\overline{x}^\mu = \Lambda^\mu{}_\nu x^\nu$$

The invariant symbol of the vector representation of Lorentz transformation is $\eta^{\mu\nu}$

$$\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta^{\rho\sigma} = \eta^{\mu\nu},$$

where,

$$\eta^{\mu\nu} \equiv \begin{bmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{bmatrix}$$

The inverse matrix of $\eta^{\mu\nu}$ is

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{bmatrix}$$

We can use $\eta^{\mu\nu}$ and and its inverse $\eta_{\mu\nu}$ to raise and lower vector indices,

$$x_\mu \equiv \eta_{\mu\nu} x^\nu$$

And we can verify the following equations,

$$\Lambda^\mu{}_\rho \Lambda_\nu{}^\rho = \delta^\nu_\mu$$

$$x^\mu = \eta^{\mu\nu} x_\nu$$

$$\overline{x}_\mu = \Lambda_\mu{}^\nu x_\nu$$

$$\Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma \eta_{\rho\sigma} = \eta_{\mu\nu},$$

In a special case when the new reference frame move along $\hat{1}$ direction with velocity β, we have

$$\overline{x}^0 = \gamma x^0 - \gamma\beta x^1$$

$$\overline{x}^1 = -\gamma\beta x^0 + \gamma x^1$$

Some physical quantity will behave like a tensor (vector,scalar) when transforming form one inertial frame to another. For example,

scalar proper time: $d\tau$, mass: m, electrical charge e

vector four velocity: $u^\mu = \frac{dx^\mu}{d\tau}$, four momentum: $p^\mu = mu^\mu$, four acceleration: $a^\mu = \frac{du^\mu}{d\tau}$, four force: $f^\mu = ma^\mu$.

We can also define the corresponding three vector.

three velocity : $\hat{v}^i = \frac{dx^i}{dt}$

$$u^0 = \gamma_v, u^i = \gamma\hat{v}^i$$

When the new reference frame move along $\hat{1}$ direction with velocity β, we have

$$\overline{\hat{v}}^1 = \frac{\hat{v}^1 - \beta}{1 - \hat{v}^1\beta}$$

$$\overline{\hat{v}}^2 = \frac{\hat{v}^2}{\gamma(1 - \hat{v}^2\beta)}$$

$$\overline{\hat{v}}^3 = \frac{\hat{v}^3}{\gamma(1 - \hat{v}^3\beta)}$$

three momentum : $\hat{p}^i = p^i$

three acceleration : $\hat{a}^i = \frac{dv^i}{dt}$

three force : $\hat{f}^i = \frac{d\hat{p}^i}{dt}$

$$f^i = \gamma_v \hat{f}^i$$

5.2 Relativistic Mechanics

For a free particle, we have

$$\frac{dp^\mu}{d\tau} = 0$$

It can be formulated in several ways

Lagrangian formulation

$$S = -m\int_a^b d\tau, \quad \delta x^\mu(a) = \delta x^\mu(b) = 0$$

$$\delta S = 0 \Rightarrow m\frac{du^\mu}{d\tau} = 0$$

Hamiltonian formulation

$$S = -m\int_{t_1}^{t_2} \sqrt{1 - \dot{x}_i\dot{x}^i}dt$$

$$L = -m\sqrt{1 - \dot{x}_i\dot{x}^i}$$

$$\pi^i = \frac{\partial L}{\partial \dot{x}_i} = \gamma m\eta^{ij}\dot{x}_j$$

$$H = \pi^i\dot{x}_i - L = \gamma m = \sqrt{m^2 + \pi^i\pi_i}$$

So, the Hamiltonian equations are

$$\dot{\pi}^i = 0, \quad \dot{x}_i = \eta_{ij}\frac{\pi^j}{\sqrt{m^2 + \pi^k\pi_k}}$$

Hamiltonian-Jacobi equation

$$H = -\frac{\partial S}{\partial t}, \quad \pi^i = \frac{\partial S}{\partial x_i}$$

If we define $p^0 = H$, $p^i = \pi^i$, then we can verify that $p^\mu = \frac{\partial S}{\partial x_\mu}$. So, p^μ is a vector under Lorentz transformation. The Hamiltonian-Jacobi equation can be written as

$$(\frac{\partial S}{\partial t})^2 = m^2 + (\frac{\partial S}{\partial x})^2 + (\frac{\partial S}{\partial y})^2 + (\frac{\partial S}{\partial z})^2$$

For a non-free particle, we have the revised newton's second law:

$$f^\mu = \frac{dp^\mu}{d\tau}$$

The formula is the definition of the four force. It can also be written in three vector as

$$\hat{f}^i = \gamma_v m\hat{a}^i + \gamma_v^3(\hat{a}^j\hat{v}_j)m\hat{v}^i$$

If the system consists of more than one particles interacting with each other. We can derive the conservation laws from the symmetry.

Translational symmetry and conservation of momentum

$$\overline{x}^\mu = x^\mu + \delta x^\mu$$

$$\delta S = \sum m u_\mu \delta x^\mu|_a^b = 0$$

$\sum p^\mu$ is conserved.

Rotational symmetry and conservation of angular momentum

$$\overline{x}^\mu = x^\mu + x_\nu \delta\Omega^{\mu\nu}$$

$$\delta S = \sum m u^\mu x^\nu \delta\Omega_{\mu\nu}|_a^b = 0$$

$\sum M^{\mu\nu}$ is conserved, where $M^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu$.

5.3 Relativistic Scattering

5.3.1 Distribution function

The number of particles in the region $\boldsymbol{r} + d\boldsymbol{r}$ and $\boldsymbol{p} + d\boldsymbol{p}$ is $f(\boldsymbol{p}, \boldsymbol{r}) dp_x dp_y dp_z dx dy dz$. Then $f(\boldsymbol{p}, \boldsymbol{r})$ is called distribution function.
We first determine the properties of the "volume element" $dp_x dp_y dp_z$, with respect to Lorentz transformations. If we introduce a four-dimensional coordinate system, on whose axes are marked the components of the four-momentum of a particle, then $dp_x dp_y dp_z$, can be considered as the zeroth component of an element of the hypersurface defined by the equation $p^\mu p_\mu + m^2 = 0$. The element of hypersurface is a four-vector directed along the normal to the hypersurface; in our case the direction of the normal obviously coincides with the direction of the four-vector p^μ. From this it follows that the ratio

$$\frac{dp_x dp_y dp_z}{E}$$

is an invariant quantity, since it is the ratio of corresponding components of two parallel four-vectors.
Then, we notice that $dV dt$ is invariant under Lorentz transformation and $dt = \frac{E}{m} d\tau$. So we can infer that

$$dx dv dz E$$

is an invariant quantity quantity. Putting all together, we know the phase volume

$$dp_x dp_y dp_z dx dy dz$$

is an invariant volume. So, we have

$$f(\boldsymbol{r}, \boldsymbol{p}) = f'(\boldsymbol{r}', \boldsymbol{p}')$$

in coordinate transformation.

5.3.2 Invariant cross section

Recall the definition of cross section

$$\sigma \equiv \frac{\text{Number of Events per target}}{\text{Time} \times \text{Incident Flux}}$$

Here, the incident flux and time are measured in the frame of target particle.
Suppose that we have two colliding beams; we denote by n_1 and n_2 the particle densities in them and by $\boldsymbol{v}_1$ and $\boldsymbol{v_2}$ the velocities of the particles. In the reference system in which particle 2 is at rest, we are dealing with the collision of the beam of particles 1 with a stationary target. Then according to the usual definition of the cross-section σ, the number of collisions occurring in volume dV in time dt is

$$dN = \sigma v_{rel} n_1 n_2 dV dt$$

,where v_{rel} is the velocity of particle 1 in the rest system of particle 2 (which is just the definition of the relative velocity of two particles in relativistic mechanics).
The number dN is by its very nature an invariant quantity. We would like to express it in a form which is applicable in any reference system:

$$dN = A n_1 n_2 dV dt$$

where A is a number to be determined, for which we know that its value in the rest frame of one of the particles is $v_{rel}\sigma$. We shall always mean by σ precisely the cross-section in the rest frame of one of the particles, i.e. by definition, an invariant quantity. From its definition, the relative velocity v_{rel} is also invariant. The product dVdt is an invariant. Therefore the product An_1n_2 must also be an invariant. The law of transformation of the particle density n is

$$n = \frac{n_0}{\sqrt{1-v^2}} = n_0 E/m$$

so we can construct A in an arbitrary frame as

$$A = -\sigma v_{rel} \frac{p_1^\mu p_{2\mu}}{E_1 E_2}$$

Note that

$$-p_1^\mu p_{2\mu} = \frac{m_1}{\sqrt{1-v_{rel}^2}} m_2 = m_1 m_2 \frac{1 - \boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{\sqrt{(1-\boldsymbol{v}_1^2)\cdot(1-\boldsymbol{v}_2^2)}}$$

we can get the following expression for v_{rel}:

$$v_{rel} = \frac{\sqrt{(\boldsymbol{v}_1 - \boldsymbol{v}_2)^2 - (\boldsymbol{v}_1 \times \boldsymbol{v}_2)^2}}{1 - \boldsymbol{v}_1 \cdot \boldsymbol{v}_2}$$

Finally, we have

$$dN = \sigma \sqrt{(\boldsymbol{v}_1 - \boldsymbol{v}_2)^2 - (\boldsymbol{v}_1 \times \boldsymbol{v}_2)^2} n_1 n_2 dV dt$$

If the velocities v_1 and v_2 are collinear, then we have

$$dN = \sigma |\boldsymbol{v}_1 - \boldsymbol{v}_2| n_1 n_2 dV dt$$

If we have only one target, then

$$dN = \sigma |\boldsymbol{v}_1 - \boldsymbol{v}_2| n_1 dt$$

5.3.3 Elastic scattering between two particles

$$(E_1, \boldsymbol{p}_1, E_2, \boldsymbol{p}_2) \to (E_1', \boldsymbol{p}_1', E_2', \boldsymbol{p}_2')$$

In lab frame (L frame),

$$E_2 = m_2 \quad \boldsymbol{p}_2 = 0$$

By the conservation of momentum, we can derive that

$$\cos\theta_1 = \frac{E_1'(E_1 + m_2) - E_1 m_2 - m_1^2}{p_1 p_1'}$$

$$\cos\theta_2 = \frac{(E_1 + m_2)(E_2' - m_2)}{p_1 p_2'}$$

Here $\theta_1(\theta_2)$ is the angle between $\boldsymbol{p}_1'(\boldsymbol{p}_2')$ with $\boldsymbol{p}_1$. In a special case where $m_1 = 0$, we have

$$E_1' = \frac{m_2}{1 - \cos\theta_1 + \frac{m_2}{E_1}}$$

Suppose in the center of mass frame (C frame), the scattering angle is χ,then we can derive that

$$E_1' = E_1 - \Delta E \quad E_2' = m_2 + \Delta E$$

Here,

$$\Delta E = \frac{m_2(E_1^2 - m_1^2)}{m_1^2 + m_2^2 + 2m_2 E_1}(1 - \cos\chi)$$

Next, suppose in L frame, let $x = p_1' \cos\theta_1$,$y = p_1' \sin\theta_1$,we can get that

$$\frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} = 1$$

Here,

$$a = \frac{p_1(E_1 m_2 + m_2^2)}{m_1^2 + m_2^2 + 2m_2 E_1} = \frac{m_2 V}{\sqrt{1 - V^2}} \quad b = \frac{m_2 p_1}{\sqrt{m_1^2 + m_2^2 + 2m_2 E_1}} = \frac{a}{\sqrt{1 - V^2}} \quad c = \frac{p_1(E_1 m_2 + m_1^2)}{m_1^2 + m_2^2 + 2m_2 E_1}$$

Here $V = \frac{p_1}{E_1 + m_2}$ is the velocity of particle 2 before scattering in C frame. And it is easy to see that $a + c = p_1$. The result above can be represented by a picture.

Figure 5.1: Relativistic scattering

Chapter 6
Classical field theory

6.1 Lagrangian formulation

$$S = \int \mathcal{L}(\phi_a, \dot{\phi}_a, \nabla\phi_a) d^4x, \quad \delta\phi_a|_\Sigma = 0$$

$$\delta S = 0 \Rightarrow \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0$$

Locality of the theory

There are no terms in the Lagrangian coupling $\phi(\boldsymbol{x}, t)$ directly to $\phi(\boldsymbol{y}, t)$ with $\boldsymbol{x} \neq \boldsymbol{y}$. The closet we get for the $\boldsymbol{x}$ label is coupling between $\phi(\boldsymbol{x}, t)$ and $\phi(\boldsymbol{x} + \delta\boldsymbol{x}, t)$ through the gradient term $\boldsymbol{\nabla}\phi$.

Lorentz invariance

Scalar fields:

$$\overline{\phi}(x) = \phi(\Lambda^{-1}x)$$

Vector fields:

$$\overline{A}^\mu(x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$$

$$\overline{A}_\mu(x) = (\Lambda^{-1})^\nu{}_\mu A_\nu(\Lambda^{-1}x) = \Lambda_\mu{}^\nu A_\nu(\Lambda^{-1}x)$$

$$\overline{\partial_\mu \phi}(x) = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \phi(\Lambda^{-1}x) = \Lambda_\mu{}^\nu \partial_\nu \phi(\Lambda^{-1}x)$$

Lagrangian is a scalar, or more loosely, action is invariant under Lorentz transformation.

6.2 Symmetry and conservation law

Theorem 6.1 Noether's theorem

Every continuous symmetry of the Lagrangian gives rise to a conserved current $j^\mu(x)$ such that the equation of motion imply $\partial_\mu j^\mu = 0$. Suppose that the infinitesimal transformation is

$$\phi_a \to \phi_a + \delta\phi_a$$
$$\mathcal{L} \to \mathcal{L} + \delta\mathcal{L}$$

and if $\delta\mathcal{L} = \partial_\mu K^\mu$, we can get

$$j^\mu = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a + K^\mu$$

space-time translation

$\overline{x} = x - a$

$$j^\mu = a_\nu T^{\mu\nu}$$
$$T^{\mu\nu} \equiv -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial^\nu\phi_a + \eta^{\mu\nu}\mathcal{L}$$

If we define $P^\mu \equiv \int T^{0\mu}d^3x$,then we have the law of momentum conservation:

$$\frac{dP^\mu}{dt} = 0$$

Lorentz Transformation

$\overline{x}^\mu = x^\mu + \delta\omega^\mu{}_\nu x^\nu$

The infinitesimal Lorentz transformation can be written as $I + \delta\omega^\mu{}_\nu$

$$\delta\omega^\mu{}_\nu = \begin{bmatrix} 0 & \beta_1 & \beta_2 & \beta_3 \\ \beta_1 & 0 & -\theta_3 & \theta_2 \\ \beta_2 & \theta_3 & 0 & -\theta_1 \\ \beta_3 & -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

This time, we assume that

$$\overline{\phi}_a(x) = \mathcal{S}_a{}^b\phi_b(\Lambda^{-1}x)$$

In the limit of infinitesimal Lorentz transformation, we have

$$\mathcal{S}_a{}^b = \delta_a{}^b + \frac{1}{2}\delta\omega_{\alpha\beta}(\Sigma^{\alpha\beta})_a{}^b$$

$$j^\mu = \frac{1}{2}M^{\mu\nu\rho}\delta\omega_{\nu\rho}$$

$$M^{\mu\nu\rho} \equiv x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} (\Sigma^{\nu\rho})_a{}^b \phi_b$$

If we define $M^{\nu\rho} \equiv \int M^{0\nu\rho} d^3x$, then we have the law of angular momentum conservation:

$$\frac{dM^{\nu\rho}}{dt} = 0$$

6.3 Functional derivatives

Definition 6.1 Functional derivatives

Given a manifold M representing (continuous/smooth) functions ρ (with certain boundary conditions etc.), and a functional F defined as

$$F\colon M \to \mathbb{R} \quad \text{or} \quad F\colon M \to \mathbb{C}\,,$$

the functional derivative of $F[\rho]$, denoted $\frac{\delta F}{\delta \rho}$,is defined by

$$\begin{aligned}\int \frac{\delta F}{\delta \rho}(x)\phi(x)\, dx &= \lim_{\varepsilon \to 0} \frac{F[\rho + \varepsilon\phi] - F[\rho]}{\varepsilon} \\ &= \left[\frac{d}{d\epsilon} F[\rho + \epsilon\phi]\right]_{\epsilon=0},\end{aligned}$$

where ϕ is an arbitrary function. The quantity $\epsilon\phi$ is called the variation of ρ.

Like the derivative of a function, the functional derivative satisfies the following properties, where $F[\rho]$ and $G[\rho]$ are functionals:
Linearity:

$$\frac{\delta(\lambda F + \mu G)[\rho]}{\delta\rho(x)} = \lambda \frac{\delta F[\rho]}{\delta\rho(x)} + \mu \frac{\delta G[\rho]}{\delta\rho(x)},$$

where λ, μ are constants.
Product rule:

$$\frac{\delta(FG)[\rho]}{\delta\rho(x)} = \frac{\delta F[\rho]}{\delta\rho(x)} G[\rho] + F[\rho] \frac{\delta G[\rho]}{\delta\rho(x)}\,,$$

Chain rules:
If F is a functional and G an operator, then

$$\frac{\delta F[G[\rho]]}{\delta\rho}(y) = \int dx \frac{\delta F[G[\rho]]}{\delta G[\rho]}(x) \cdot \frac{\delta G[\rho]}{\delta\rho}(x, y)$$

If G is an ordinary differentiable function g, then this reduces to

$$\frac{\delta F[g(\rho)]}{\delta\rho}(y) = \frac{\delta F[g(\rho)]}{\delta g[\rho]}(y)\, \frac{dg(\rho)}{d\rho}(y)$$

Proposition 6.1 Properties of functional derivatives

$$\frac{\delta F}{\delta \rho}(y) = \lim_{\epsilon \to \infty} \frac{1}{\epsilon}\{F[\rho(x) + \epsilon\delta(x-y)] - F[\rho(x)]\}$$

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x-y)$$

$$\frac{\delta}{\delta f(y)} \int g\left(f(x)\right) dx = g'(f(y))$$

$$\frac{\delta f'(x)}{\delta f(y)} = \frac{d}{dx}\delta(x-y)$$

$$\frac{\delta}{\delta f(y)} \int g\left(f'(x)\right) dx = -\frac{d}{dy} g'(f'(y))$$

6.4 Hamiltonian formulation

$$\pi^a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a}$$

$$\mathcal{H}(\phi_a, \nabla\phi_a, \pi^a) = \pi^a \dot{\phi}_a - \mathcal{L}$$

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

Now, we can get the Hamilton equation form $\delta S = 0$,

$$\dot{\phi}_a(\boldsymbol{x}, t) = \frac{\delta}{\delta \pi^a(\boldsymbol{x}, t)} H = \frac{\partial \mathcal{H}}{\partial \pi^a}$$

$$\dot{\pi^a}(\boldsymbol{x}, t) = -\frac{\delta}{\delta \phi_a(\boldsymbol{x}, t)} H = -\frac{\partial \mathcal{H}}{\partial \phi_a} + \left(\frac{\partial \mathcal{H}}{\partial \phi_{a,b}}\right)_{,b}$$

6.4.1 Poission bracket

First, we define that

$$[\phi_a(\boldsymbol{x}), \phi_b(\boldsymbol{y})] \equiv [\pi^a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] \equiv 0$$

$$[\phi_a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] \equiv \delta_a^b \delta(\boldsymbol{x} - \boldsymbol{y})$$

then, we demand that the bracket operation has the same properties as the Poission bracket in classical mechanics. And we also demand that

$$[\partial_x A(\boldsymbol{x}), B(\boldsymbol{y})] = \partial_x [A(\boldsymbol{x}), B(\boldsymbol{y})]$$

and

$$\left[\int d^3x A(\boldsymbol{x}), B(\boldsymbol{y})\right] = \int d^3x [A(\boldsymbol{x}), B(\boldsymbol{y})]$$

As a result, we can verify that

$$[W[\phi(\boldsymbol{x}),\pi(\boldsymbol{x})],Z[\phi(\boldsymbol{x}),\pi(\boldsymbol{x})]]=\int d^3x\left\{\frac{\delta W}{\delta\phi(\boldsymbol{x})}\frac{\delta Z}{\delta\pi(\boldsymbol{x})}-\frac{\delta W}{\delta\pi(\boldsymbol{x})}\frac{\delta Z}{\delta\phi(\boldsymbol{x})}\right\}$$

Especially,

$$[\phi_a(\boldsymbol{x}),H]=\frac{\delta}{\delta\pi^a(\boldsymbol{x})}H,\quad [\pi^a(\boldsymbol{x}),H]=-\frac{\delta}{\delta\phi_a(\boldsymbol{x})}H$$

So, the Hamilton equation can be written as

$$\dot{\phi}_a=[\phi_a,H],\quad \dot{\pi^a}=[\pi^a,H]$$

Further more, we can prove that

$$\frac{dO(\phi,\pi,t)}{dt}=[O,H]+\frac{\partial O}{\partial t}$$

and

$$\frac{d[A,B]}{dt}=[A,\frac{dB}{dt}]+[\frac{dA}{dt},B]$$

6.4.2 Momentum

It is easy to verify that

$$P^0=H,\quad P^i=\int -\pi^a\partial^i\phi_a d^3x$$

And we can get the Poisson bracket

$$\begin{aligned}[\phi_a,P^\mu] &= -\partial^\mu\phi_a\\ [\pi^a,P^\mu] &= -\partial^\mu\pi^a\\ [P^\mu,P^\nu] &= 0\end{aligned}$$

6.4.3 Angular momentum

It is easy to verify that

$$M^{\mu\nu}=\int(x^\mu T^{0\nu}-x^\nu T^{0\mu}-\pi^a(\Sigma^{\mu\nu})_a{}^b\phi_b)d^3x$$

We define that

$$M_L^{\mu\nu}\equiv\int(x^\mu T^{0\nu}-x^\nu T^{0\mu})d^3x\quad M_S^{\mu\nu}\equiv\int(-\pi^a(\Sigma^{\mu\nu})_a{}^b\phi_b)d^3x$$

$$(L^{\mu\nu})_a{}^b\equiv-(x^\mu\partial^\nu-x^\nu\partial^\mu)\delta_a{}^b\quad (S^{\mu\nu})_a{}^b\equiv-(\Sigma^{\mu\nu})_a{}^b$$

So, we can get the Poisson bracket

$$[\phi_a,M_L^{\mu\nu}]=(L^{\mu\nu})_a{}^b\phi_b\quad [\phi_a,M_S^{\mu\nu}]=(S^{\mu\nu})_a{}^b\phi_b$$

$$[\pi^a,M_L^{\mu\nu}]=(L^{\mu\nu})_b{}^a\pi^b\quad [\pi^a,M_S^{\mu\nu}]=-(S^{\mu\nu})_b{}^a\pi^b$$

Because $\frac{dM^{\mu\nu}}{dt} = 0$, $M^{\mu\nu}$ can commutate with $\frac{d}{dt}$, so

$$[[\phi(x), M^{\mu\nu}], M^{\rho\sigma}] = (L^{\mu\nu} + S^{\mu\nu})(L^{\rho\sigma} + S^{\rho\sigma})\phi(x)$$

At last, we can get the Poisson bracket

$$[\phi(x), [M^{\mu\nu}, M^{\rho\sigma}]] = (L^{\mu\nu}L^{\rho\sigma} - L^{\rho\sigma}L^{\mu\nu} + S^{\mu\nu}S^{\rho\sigma} - S^{\rho\sigma}S^{\mu\nu})\phi(x)$$

Since we can prove that

$$L^{\mu\nu}L^{\rho\sigma} - L^{\rho\sigma}L^{\mu\nu} = -\eta^{\nu\rho}L^{\mu\sigma} + \eta^{\sigma\mu}L^{\rho\nu} + \eta^{\mu\rho}L^{\nu\sigma} - \eta^{\sigma\nu}L^{\rho\mu}$$

if we demand that

$$S^{\mu\nu}S^{\rho\sigma} - S^{\rho\sigma}S^{\mu\nu} = -\eta^{\nu\rho}S^{\mu\sigma} + \eta^{\sigma\mu}S^{\rho\nu} + \eta^{\mu\rho}S^{\nu\sigma} - \eta^{\sigma\nu}S^{\rho\mu}$$

We will get get the Poisson bracket of the $M^{\mu\nu}$,

$$[M^{\mu\nu}, M^{\rho\sigma}] = -\eta^{\nu\rho}M^{\mu\sigma} + \eta^{\sigma\mu}M^{\rho\nu} + \eta^{\mu\rho}M^{\nu\sigma} - \eta^{\sigma\nu}M^{\rho\mu}$$

up to the possibility of a term on the right-hand side that commutes with $\phi(x)$ and its derivatives.

We now define $J_i \equiv \frac{1}{2}\epsilon_{ijk}M^{jk}$ and $K_i \equiv M^{i0}$, so we have

$$M^{\mu\nu} = \begin{bmatrix} 0 & -K_1 & -K_2 & -K_3 \\ K_1 & 0 & J_3 & -J_2 \\ K_2 & -J_3 & 0 & J_1 \\ K_3 & J_2 & -J_1 & 0 \end{bmatrix} \quad \left(\delta\omega_{\mu\nu} = \begin{bmatrix} 0 & -\beta_1 & -\beta_2 & -\beta_3 \\ \beta_1 & 0 & -\theta_3 & \theta_2 \\ \beta_2 & \theta_3 & 0 & -\theta_1 \\ \beta_3 & -\theta_2 & \theta_1 & 0 \end{bmatrix} \right)$$

the Poisson bracket can be written as

$$\begin{aligned} [J_i, J_j] &= \epsilon_{ijk}J_k \\ [J_i, K_j] &= \epsilon_{ijk}K_k \\ [K_i, K_j] &= -\epsilon_{ijk}J_k \end{aligned}$$

We can use the similar method to derive that

$$[P^\mu, M^{\rho\sigma}] = \eta^{\mu\sigma}P^\mu - \eta^{\mu\rho}P^\sigma$$

It can also be written as

$$\begin{aligned} [J_i, H] &= 0 \\ [J_i, P_j] &= \epsilon_{ijk}P_k \\ [K_i, H] &= P_i \\ [K_i, P_j] &= \delta_{ij}H \end{aligned}$$

At last, we define $L_i \equiv \frac{1}{2}\epsilon_{ijk}M_L^{jk}$ and $S_i \equiv \frac{1}{2}\epsilon_{ijk}M_S^{jk}$ we can demonstrate that

$$\begin{aligned} [L_i, S_j] &= 0 \\ [S_i, P_j] &= 0 \\ [L_i, P_j] &= \epsilon_{ijk}P_k \end{aligned}$$

Chapter 7
Classical Electrodynamics

7.1 The formulation of classical electrodynamics

7.1.1 Maxwell equations and Lorentz force

The Lagrangian of the EM field A^μ when coupling with current is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^\mu A_\mu$$

Here

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad j^\mu = \rho_{e0}u^\mu$$

ρ_{e0} is the charge density measured in the frame of charge. Then we can derive the Euler-Lagrange equation of EM field:

$$\partial_\nu F^{\mu\nu} = j^\mu$$

And we can derive the charge conservation equation

$$\partial_\mu j^\mu = 0$$

For a charged particle moving in EM field with orbit $x^\mu(\tau)$, we have

$$j^\mu = e_a\delta(\boldsymbol{r} - \boldsymbol{r}_a(t))\sqrt{1-\boldsymbol{v}_a^2}u_a^\mu$$

Then we can derive that

$$\int dV dt j^\mu A_\mu = \int dx^\mu e A_\mu$$

So, the action for a charged particle when coupling with EM field is

$$S = -m\int d\tau + e\int dx^\mu A_\mu$$

We can derive the Euler-Lagrangian equation of the particle:

$$ma_\mu = eF_{\mu\nu}u^\nu$$

The Hamiltonian formulation of electrodynamics will be discussed in detail in the Hamiltonian formulation of general relativity and canonical quantization formulation of quantum electrodynamics.
We now define the electric field and magnetic field as follows,

$$E^i \equiv F^{0i} = -\dot{A}^i - \nabla^i A^0 \quad B^i \equiv \epsilon_{ijk}\nabla^j A^k$$

We also define $\rho_e \equiv j^0$ and $J^i \equiv j^i$,so the field equation can then be written

$$\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{B} &= \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{J} \\
\boldsymbol{\nabla} \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} \\
\boldsymbol{\nabla} \cdot \boldsymbol{E} &= \rho_e \\
\boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0
\end{aligned}$$

That is the so-called Maxwell equation.
The equation of motion of the charged particle can be written as

$$\frac{d\boldsymbol{p}}{dt} = e(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \quad \frac{d\mathcal{E}}{dt} = e\boldsymbol{E} \cdot \boldsymbol{v}$$

That is the so-called Lorentz force.
We note that the Field A cannot be completely determined by Maxwell and Lorentz equation. If we make the transformation $A_\mu \to A_\mu + \partial_\mu \xi(x)$, $\mathcal{L}$ and $F^{\mu\nu}$ would be invariant, and the Maxwell and Lorentz equation is still valid. This arbitrariness of ξ is called gauge invariance. The topic will be discussed in detail in QED.

7.1.2 Lorentz transformation of fields

In the new coordinate system $\overline{x}^\mu = \Lambda^\mu{}_\nu x^\nu$, we have

$$\overline{A}^\mu(\overline{x}) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}\overline{x})$$

In a special case when the new reference frame move along $\hat{1}$ direction with velocity β, we have

$$\overline{x}^0 = \gamma x^0 - \gamma\beta x^1 \quad \overline{x}^1 = -\gamma\beta x^0 + \gamma x^1$$

So, we have

$$\overline{A}^0 = \gamma A^0 - \gamma\beta A^1 \quad \overline{A}^1 = -\gamma\beta A^0 + \gamma A^1$$

We can further derive that

$$\overline{E}_1 = E_1 \quad \overline{E}_2 = \gamma E_2 - \gamma\beta B_3 \quad \overline{E}_3 = \gamma E_3 + \gamma\beta B_2$$
$$\overline{B}_1 = B_1 \quad \overline{B}_2 = \gamma B_2 + \gamma\beta E_3 \quad \overline{B}_3 = \gamma B_3 - \gamma\beta E_2$$

It can be written in the three vector form as

$$\overline{\boldsymbol{E}} = \gamma(\boldsymbol{E}_\perp + \boldsymbol{\beta} \times \boldsymbol{B}) + \boldsymbol{E}_\parallel$$
$$\overline{\boldsymbol{B}} = \gamma(\boldsymbol{B}_\perp - \boldsymbol{\beta} \times \boldsymbol{E}) + \boldsymbol{B}_\parallel$$

If $\beta \ll 1$, neglect the higher order of β^2, we have

$$\overline{\boldsymbol{E}} = \boldsymbol{E} + \boldsymbol{\beta} \times \boldsymbol{B}$$
$$\overline{\boldsymbol{B}} = \boldsymbol{B} - \boldsymbol{\beta} \times \boldsymbol{E}$$

We also note that $F_{\mu\nu}F^{\mu\nu}$ and $\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$ is invariant under Lorentz transformation. It can be represented by the electric and magnetic field as

$$E^2 - B^2 = \text{ inv} \quad \boldsymbol{E} \cdot \boldsymbol{B} = \text{ inv}$$

7.1.3 Energy-momentum tensor

For a free EM field, the energy-momentum tensor is

$$T_f^{\mu\nu} \equiv -\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\rho)}\partial^\nu A_\rho + \eta^{\mu\nu}\mathcal{L} = \partial^\nu A^\rho F^\mu{}_\rho - \frac{1}{4}\eta^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}$$

We note that the energy-momentum tensor defined above is not symmetric, so we define a modified energy-momentum tensor by adding a term $-\partial^\rho A^\nu F^\mu{}_\rho$,so

$$T_e'^{\mu\nu} = F^{\nu\rho}F^\mu{}_\rho - \frac{1}{4}\eta^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}$$

Note that for free EM field, $\partial^\rho A^\nu F^\mu{}_\rho = \partial^\rho\left(A^\nu F^\mu{}_\rho\right)$.So, we can get that

$$\partial_\mu T_f'^{\mu\nu} = 0 \quad P_f'^\mu = P_f^\mu$$

So, from now on, we will use T_f' as the energy-momentum tensor of EM field and omit the prime of T_f' for simplicity. The momentum of the free EM field is

$$P_f^0 = \int dV \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) \equiv \int dV w \quad P_f^i = \int dV \boldsymbol{E}\times\boldsymbol{B} \equiv \int dV \boldsymbol{S}$$

If there also exists charged particles in the system, i.e. the source of EM field, we must also include the energy-momentum tensor of the particles to get the right conservation equation. The energy-momentum tensor of particles is defined as

$$T_p^{\mu\nu} \equiv \sum_a m_a \delta(\boldsymbol{r}-\boldsymbol{r}_a)\sqrt{1-\boldsymbol{v}_a^2}u_a^\mu u_a^\nu$$

In this definition, we have

$$P_p^0 = \sum_a \frac{m_a}{\sqrt{1-\boldsymbol{v}_a^2}} \quad \boldsymbol{P}_p = \sum_a \frac{m_a}{\sqrt{1-\boldsymbol{v}_a^2}}\boldsymbol{v}_a$$

So, our definition is consistent with the definition of four momentum in previous chapter. Recall that

$$j_e^\mu = \sum_a e_a \delta(\boldsymbol{r}-\boldsymbol{r}_a)\sqrt{1-\boldsymbol{v}_a^2}u_a^\mu$$

We can define the mass current as

$$j_m^\mu \equiv \sum_a m_a \delta(\boldsymbol{r}-\boldsymbol{r}_a)\sqrt{1-\boldsymbol{v}_a^2}u_a^\mu$$

If there is no particle creation and annihilation, and the mass of the particle is constant during the motion, we would have

$$\partial_\mu j_{ma}^\mu = 0$$

Recall the Lorentz force equation $ma_\mu = eF_{\mu\nu}u^\nu$, it can be rewritten as

$$\rho_{m0a}\frac{du_\mu}{d\tau} = F_{\mu\nu}j_{ea}^\nu$$

Here, $\rho_{m0a} \equiv m_a \delta(\boldsymbol{r} - \boldsymbol{r}_a)\sqrt{1-\boldsymbol{v}_a^2}$ is the mass density measured in the rest frame of the particle. Then, on the one hand, we have

$$\partial_\mu T_p^{\mu\nu} = \sum_a \rho_{m0a} u^\mu \frac{\partial u_a^\nu}{\partial x^\mu} = \sum_a \rho_{m0a} \frac{du_a^\nu}{d\tau} = F^\nu{}_\mu j_e^\mu$$

On the other hand, we can derive that

$$\partial_\mu T_f^{\mu\nu} = -F^\nu{}_\mu j_e^\mu$$

by implementing Maxwell equation. Define

$$T^{\mu\nu} \equiv T_f^{\mu\nu} + T_p^{\mu\nu}$$

We have

$$\partial_\mu T^{\mu\nu} = 0$$

We define the Maxwell stress tensor as

$$f^{ij} \equiv -T_f^{ij} = E^i E^j + B^i B^j - w\delta^{ij}$$

Then, we can write down the conservation law of energy and momentum as

$$\frac{d}{dt}\left(P_p^0 + \int w dV\right) = -\oint \boldsymbol{S} \cdot d\boldsymbol{\sigma}$$

$$\frac{d}{dt}\left(\boldsymbol{P}_p + \int \boldsymbol{S} dV\right) = \oint \boldsymbol{f} \cdot d\boldsymbol{\sigma}$$

We assume there is no particle cross the boundary.
At last, because $T^{\mu\nu} = T^{\nu\mu}$, we have

$$\partial_\mu (x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu}) = 0$$

It is easy to write down the conservation law of angular momentum as

$$\frac{d}{dt}\left(\boldsymbol{L}_p + \int \boldsymbol{r} \times \boldsymbol{S} dV\right) = \oint \boldsymbol{r} \times \boldsymbol{f} \cdot d\boldsymbol{\sigma}$$

Here,

$$\boldsymbol{L}_p = \sum_a \frac{m_a}{\sqrt{1-\boldsymbol{v}_a^2}} \boldsymbol{r}_a \times \boldsymbol{v}_a$$

7.1.4 Charged particles in a given EM field

Now we suppose the EM field is given, i.e. we will neglect the EM field generated by the test charged particles. Then we have

$$S = \int_{t_1}^{t_2} (-\sqrt{1-v^2} m + e\boldsymbol{A} \cdot \boldsymbol{v} - e\phi) dt$$

So,

$$L = -\sqrt{1-v^2} m + e\boldsymbol{A} \cdot \boldsymbol{v} - e\phi$$

The canonical momentum is

$$\boldsymbol{\pi} = \frac{\partial L}{\partial \boldsymbol{v}} = \gamma m \boldsymbol{v} + e\boldsymbol{A}$$

The Hamiltonian is therefore

$$H = \boldsymbol{\pi} \cdot \boldsymbol{v} - L = \gamma m + e\phi = \sqrt{m^2 + (\boldsymbol{\pi} - e\boldsymbol{A})^2} + e\phi$$

If $v \ll 1$, we would have

$$L = \frac{mv^2}{2} + e\boldsymbol{A} \cdot \boldsymbol{v} - e\phi \quad \boldsymbol{\pi} = m\boldsymbol{v} + e\boldsymbol{A} \quad H = \frac{(\boldsymbol{\pi} - e\boldsymbol{A})^2}{2m} + e\phi$$

If the EM field is constant in time, we have

$$\boldsymbol{\nabla} \times \boldsymbol{E} = 0$$

we would choose the gauge that

$$\dot{\boldsymbol{A}} = 0 \quad \boldsymbol{E} = -\boldsymbol{\nabla}\phi$$

Because $\frac{\partial L}{\partial t} = 0$, we can know that $\gamma m + e\phi$ is a constant.
Now we would list some special case.

Motion in a uniform and constant electric field

Suppose the the direction of electric field is $\hat{x}$, the orbit is in the $x - y$ plane. So, the equation of motion is

$$\dot{p}_x = eE \quad \dot{p}_y = 0$$

The final solution is

$$x = \frac{1}{eE}\sqrt{\mathcal{E}_0^2 + (eEt)^2} \quad y = \frac{p_0}{eE}\operatorname{arcsinh}\frac{eEt}{\mathcal{E}_0}$$

Here, we assume when $t = 0$, $p_x = 0$,$p_y = p_0$. The orbit function is

$$x = \frac{\mathcal{E}_0}{eE}\cosh\frac{eEy}{p_0}$$

Motion in a uniform and constant magnetic field

Suppose the direction of magnetic field is $\hat{z}$. Note that particle's kinetic energy $\mathcal{E} = \gamma m$ is constant if there are no electric field, we can derive the equation of motion

$$\dot{v}_x = \omega v_y \quad \dot{v}_y = -\omega v_x \quad \dot{v}_z = 0$$

Here, $\omega = \frac{eB}{\gamma m}$. The final solution is

$$x = x_0 + r\sin(\omega t + \alpha) \quad y = y_0 + r\cos(\omega t + \alpha) \quad z = z_0 + v_{0z}t$$

Here, $x_0, y_0, z_0, r, \alpha, v_{0z}$ should be determined by initial condition.

Motion in a uniform and constant EM field

We only focus on the case that the velocity of particle is much smaller the velocity of light. Suppose the direction of magnetic field is $\hat{z}$, the direction of electric field is within $y-z$ plane. The equation of motion is

$$m\ddot{x} = eB\dot{y} \quad m\ddot{y} = eE_y - eB\dot{x} \quad m\ddot{z} = eE_z$$

The solution is

$$\dot{x} = a\cos\omega t + \frac{E_y}{B} \quad \dot{y} = -a\sin\omega t \quad \dot{z} = v_{0z} + \frac{eE}{m}t$$

Here, $\omega = \frac{eB}{m}$, a, v_{z0} is determined by initial condition. As we suppose that $v \ll 1$ is satisfied,we must have that

$$a \ll 1 \quad v_{0z} \ll 1 \quad \frac{eE_z t}{m} \ll 1 \quad E_y \ll B$$

7.2 Constant electromagnetic field

7.2.1 Coulomb' law

For constant electric field, the Maxwell equation take the form

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho_e \quad \boldsymbol{\nabla} \times \boldsymbol{E} = 0$$

So, we have

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi \quad \nabla^2\phi = -\rho_e$$

The solution is

$$\phi(\boldsymbol{r}) = \int \frac{\rho_e(\boldsymbol{r}')}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|}dV'$$

If $\rho_e(\boldsymbol{r}') = Q\delta(\boldsymbol{r}')$, we have

$$\phi(\boldsymbol{r}) = \frac{Q}{4\pi|\boldsymbol{r}|} \quad E(\boldsymbol{r}) = \frac{Q\boldsymbol{r}}{4\pi|\boldsymbol{r}|^3}$$

For a system of static charged particles, the total energy is

$$U = \frac{1}{2}\int E^2 dV = \frac{1}{2}\int \rho\phi dV = \frac{1}{2}\sum e_a\phi_a + \frac{1}{2}\sum e_a\Phi_a$$

Here, ϕ_a is the electric potential at the point where e_a is located, produced by e_a itself, while Φ_a is the potential produced by other charges. It is obvious that $U_{self} = \frac{1}{2}e_a\phi_a$ is infinite, indicating that classical electrodynamics is no more valid in small distance. This problem will be solved in quantum electrodynamics: the mass of charged particle we measured is already renormalized to include the electromagnetic self energy. So, actually, we have

$$U = \frac{1}{2}\int E^2 dV - U_{self} = \frac{1}{2}\sum_{a\neq b}\frac{e_a e_b}{4\pi R_{ab}}$$

If the charged particle is moving with a constant velocity, we can derive the electric field it produced by Lorentz transformation, the final result is that

$$\boldsymbol{E} = \frac{e\boldsymbol{R}}{4\pi R^3}\frac{1-V^2}{(1-V^2\sin^2\theta)^{3/2}} \quad \boldsymbol{B} = \boldsymbol{V}\times\boldsymbol{E}$$

Here, $\boldsymbol{R}$ is the vector point from the particle to the point we measure the electric field, and θ is the angle between $\boldsymbol{V}$ and $\boldsymbol{R}$.If $V \sim 1$, the electric field will be concentrated in the direction perpendicular to the $\boldsymbol{V}$. If $\boldsymbol{V} \ll 1$, we have

$$\boldsymbol{E} = \frac{e\boldsymbol{R}}{4\pi R^3} \quad \boldsymbol{B} = \frac{e\boldsymbol{V}\times\boldsymbol{R}}{4\pi R^3}$$

7.2.2 Multipole moments

For a system of charged particles, the potential it produced at $\boldsymbol{R}$ is

$$\phi = \sum_a \frac{e_a}{4\pi|\boldsymbol{R}-\boldsymbol{r}_a|}$$

If $R \gg r_a$, we can expand the equation around $r_a = 0$. Generally, we have

$$\frac{1}{|\boldsymbol{R}-\boldsymbol{r}|} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{r^l}{R^{l+1}}\frac{4\pi}{2l+1}Y^*_{lm}(\Theta,\Phi)Y_{lm}(\theta,\phi)$$

So,we have

$$\phi = \sum \phi^{(l)}$$

Here,

$$\phi^{(l)} = \frac{1}{4\pi R^{l+1}}\sum_{m=-l}^{l}\sqrt{\frac{4\pi}{2l+1}}Q^{(l)}_m Y^*_{lm}(\Theta,\Phi)$$

$$Q^{(l)}_m = \sum_a e_a r_a^l\sqrt{\frac{4\pi}{2l+1}}Y_{lm}(\theta_a,\phi_a)$$

Special case:

$$\phi^{(0)} = \frac{Q}{4\pi R} \quad E^{(0)} = \frac{Q\boldsymbol{n}}{4\pi R^2} \quad Q = \sum_a e_a$$

$$\phi^{(1)} = \frac{\boldsymbol{d}\cdot\boldsymbol{n}}{4\pi R^2} \quad E^{(1)} = \frac{3(\boldsymbol{d}\cdot\boldsymbol{n})\boldsymbol{n}-\boldsymbol{d}}{4\pi R^3} \quad \boldsymbol{d} = \sum_a e_a\boldsymbol{r}_a$$

$$\phi^{(2)} = \frac{\boldsymbol{n}\cdot\boldsymbol{D}\cdot\boldsymbol{n}}{8\pi R^3} \quad E^{(2)} = \frac{5(\boldsymbol{n}\cdot\boldsymbol{D}\cdot\boldsymbol{n})\boldsymbol{n}-(\boldsymbol{n}\cdot\boldsymbol{D}+\boldsymbol{D}\cdot\boldsymbol{n})}{8\pi R^4} \quad D_{ij} = \sum e(3x_ix_j - r^2\delta_{ij})$$

For a system of charged particles in the electric field $\phi(\boldsymbol{r})$, if all the particles is near the $r = 0$, we can make the expansion

$$\phi(\boldsymbol{r}) = \sum_{l=0}^{\infty} r^l \sum_{m=-l}^{m=l} a_{lm}\sqrt{\frac{4\pi}{2l+1}}Y_{lm}(\theta,\phi)$$

So

$$U = \sum_{l=0}^{\infty} U^{(l)} \quad U^{(l)} = \sum_{m=-l}^{l} a_{lm} Q_m^{(l)}$$

Special case:

$$U^{(0)} = Q\phi_0 \quad F^{(0)} = Q\boldsymbol{E}_0$$

$$U^{(1)} = -\boldsymbol{d}\cdot\boldsymbol{E}_0 \quad F^{(1)} = \boldsymbol{d}\cdot\boldsymbol{\nabla}\boldsymbol{E}_0 \quad M^{(1)} = \boldsymbol{d}\times\boldsymbol{E}_0$$

$$U^{(2)} = -\frac{1}{6}\boldsymbol{D}\cdot\boldsymbol{\nabla}\boldsymbol{E}_0 \quad F^{(2)} = \frac{1}{6}\boldsymbol{D}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\boldsymbol{E}_0 \quad M^{(2)} = \frac{1}{3}\boldsymbol{\nabla}\cdot(\boldsymbol{D}\times\boldsymbol{E}_0)$$

7.2.3 Biot-Savart law

Let us consider the magnetic field produced by charges which perform a finite motion, in which the particles are always within a finite region of space and the momenta also always remain finite. Consider the time average magnetic field $\overline{\boldsymbol{B}}$, produced by the charges; this field will now be a function only of the coordinates and not of the time. We take the time average of the Maxwell equations

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \quad \boldsymbol{\nabla}\times\boldsymbol{B} = \frac{\partial\boldsymbol{E}}{\partial t} + \boldsymbol{j}$$

Note that the average value of the derivative $\partial\boldsymbol{E}/\partial t$, like the derivative of any quantity which varies over a finite range, is zero. We can get

$$\boldsymbol{\nabla}\cdot\overline{\boldsymbol{B}} = 0 \quad \boldsymbol{\nabla}\times\overline{\boldsymbol{B}} = \overline{\boldsymbol{j}}$$

Recall that $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$. We impose the gauge condition $\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0$, we have

$$\nabla^2\overline{\boldsymbol{A}} = -\overline{\boldsymbol{j}}$$

The solution is

$$\overline{\boldsymbol{A}}(\boldsymbol{r}) = \frac{1}{4\pi}\int\frac{\overline{\boldsymbol{j}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}dV' = \frac{1}{4\pi}\sum\overline{\frac{e_a\boldsymbol{v}_a}{|\boldsymbol{r}-\boldsymbol{r}_a|}}$$

And the magnetic field is

$$\overline{\boldsymbol{B}}(\boldsymbol{r}) = \frac{1}{4\pi}\int\frac{\overline{\boldsymbol{j}}(\boldsymbol{r}')\times\boldsymbol{R}}{R^3}dV' \quad \boldsymbol{R} = \boldsymbol{r}-\boldsymbol{r}'$$

7.2.4 Magnetic moment

For a system of charged particles, the potential it produced at $\boldsymbol{R}$ is

$$\overline{\boldsymbol{A}} = \frac{1}{4\pi}\sum\overline{\frac{e_a\boldsymbol{v}_a}{|\boldsymbol{R}-\boldsymbol{r}_a|}}$$

If $R \gg r_a$, we can expand the equation around $r_a = 0$ the first order,

$$4\pi\overline{\boldsymbol{A}} = \frac{1}{R}\sum e\overline{\boldsymbol{v}} - \sum\overline{e\boldsymbol{v}\left(\boldsymbol{r}\cdot\boldsymbol{\nabla}\frac{1}{R}\right)}$$

Firstly,

$$\sum e\overline{\boldsymbol{v}} = \overline{\frac{d}{dt}\sum e\boldsymbol{r}} = 0$$

Secondly,

$$-\sum \overline{e\boldsymbol{v}\left(\boldsymbol{r}\cdot\boldsymbol{\nabla}\frac{1}{R}\right)} = \frac{1}{R^3}\sum \overline{e\boldsymbol{v}(\boldsymbol{r}\cdot\boldsymbol{R})}$$

Note that

$$\sum e\boldsymbol{v}(\boldsymbol{r}\cdot\boldsymbol{R}) = \frac{1}{2}\frac{d}{dt}e\boldsymbol{r}(\boldsymbol{r}\cdot\boldsymbol{R}) + \frac{1}{2}\left(\sum e\boldsymbol{r}\times\boldsymbol{v}\right)\times\boldsymbol{R}$$

Define the magnetic moment as

$$\boldsymbol{m} \equiv \frac{1}{2}\left(\sum e\boldsymbol{r}\times\boldsymbol{v}\right)$$

We can get

$$\overline{\boldsymbol{A}} = \frac{\overline{\boldsymbol{m}}\times\boldsymbol{R}}{4\pi R^3} \quad \overline{\boldsymbol{B}} = \frac{3\boldsymbol{n}(\overline{\boldsymbol{m}}\cdot\boldsymbol{n}) - \overline{\boldsymbol{m}}}{4\pi R^3}$$

If all the particles have the same charge-mass ratio, and the velocity of all the particles is much smaller than that of light, we have

$$\boldsymbol{m} = \frac{e}{2m}\sum m\boldsymbol{r}\times\boldsymbol{v} = \frac{e}{2m}\boldsymbol{M}$$

Let us consider a system of charges in an external constant uniform magnetic field. The time average of the force acting on the system,

$$\boldsymbol{F} = \sum e\overline{\boldsymbol{v}\times\boldsymbol{B}} = \overline{\frac{d}{dt}\sum e\boldsymbol{r}\times\boldsymbol{B}} = 0$$

The average value of the moment of the forces is

$$\overline{\boldsymbol{K}} = \sum e\overline{\boldsymbol{r}\times(\boldsymbol{v}\times\boldsymbol{B})}$$

We can derive that

$$\overline{\boldsymbol{K}} = \overline{\boldsymbol{m}}\times\boldsymbol{B}$$

Let us consider the change in the average angular momentum $\overline{\boldsymbol{M}}$ of the system. According to a well-known equation of mechanics, the derivative of $\boldsymbol{M}$ is equal to the moment $\boldsymbol{K}$ of the forces acting on the system. We therefore have

$$\frac{d\overline{\boldsymbol{M}}}{dt} = \overline{\boldsymbol{m}}\times\boldsymbol{B}$$

If the charge-mass ratio is the same for all particles of the system, the angular momentum and magnetic moment are proportional to one another, and we find:

$$\frac{d\overline{\boldsymbol{M}}}{dt} = -\boldsymbol{\Omega}\times\overline{\boldsymbol{M}} \quad \boldsymbol{\Omega} = \frac{e}{2m}\boldsymbol{B}$$

This equation states that the vector $\overline{\boldsymbol{M}}$ rotates with angular velocity $-\boldsymbol{\Omega}$ around the direction of the field, while its absolute magnitude and the angle which it makes with this direction remain fixed. This motion is called the Larmor precession. Let us consider the Lagrangian for a system of charges in an external constant uniform magnetic field, it contains the additional term

$$L_H = \sum e\boldsymbol{A} = \sum e(\boldsymbol{B}\times\boldsymbol{r})\cdot\boldsymbol{v} = \sum e(\boldsymbol{r}\times\boldsymbol{v})\cdot\boldsymbol{B} = \boldsymbol{m}\cdot\boldsymbol{B}$$

7.3 Electromagnetic waves

7.3.1 Electromagnetic waves

electromagnetic fields occurring in vacuum in the absence of charges are called electromagnetic waves. We choose the Coulomb's Gauge, i.e.

$$\phi = 0 \quad \boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$$

So,

$$\boldsymbol{E} = -\frac{\partial A}{\partial t} \quad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$$

Then,from Maxwell equation we can derive that

$$\nabla^2 \boldsymbol{A} - \frac{\partial^2 \boldsymbol{A}}{\partial t^2} = 0$$

This is the equation which determines the potentials of electromagnetic waves. It is called the d'Alembert equation, or the wave equation. We can verify that the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy the same wave equation.

We consider the special case of electromagnetic waves in which the fields depends only on one coordinates,say x,Such waves are said to be plane. In this case the equation for the field becomes

$$\frac{\partial^2 f}{\partial t^2} - \frac{\partial^2 f}{\partial x^2} = 0$$

The solution is

$$f(t, x) = f_1(t - x) + f_2(t + x)$$

$f_1(t - x)$ represents a plane wave moving in the positive direction along the x axis. $f_2(t - x)$ represents a plane wave moving in the negative direction along the x axis. The Coulomb's gauge would imply that $A_x = 0$, and we can obtain

$$\boldsymbol{E} = -\boldsymbol{A}' \quad \boldsymbol{B} = -\boldsymbol{n} \times \boldsymbol{A}' = \boldsymbol{n} \times \boldsymbol{E}$$

where the prime denotes differentiation with respect to $t - x$ and $\boldsymbol{n}$ is a unit vector along the direction of propagation of the wave. We see that the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ of a plane wave are directed perpendicular to the direction of propagation of the wave. For this reason, electromagnetic waves are said to be transverse.
The energy density and flux of the plane waves are

$$W = \boldsymbol{E}^2 \quad \boldsymbol{S} = W\boldsymbol{n}$$

7.3.2 Monochromatic wave

A very important special case of electromagnetic waves is a wave in which the field is a simply periodic function of the time. Such a wave is said to be monochromatic. All quantities (potentials, field components) in a monochromatic wave depend on the time through a factor

of the form $\cos(\omega t + a)$. The quantity ω is called the cyclic frequency of the wave (we shall simply call it the frequency). For the monochromatic wave, the wave equation becomes

$$\frac{\partial^2 f}{\partial x^2} + \omega^2 f = 0$$

The vector potential of such a wave is most conveniently written as the real part of a complex expression

$$\boldsymbol{A} = \mathrm{Re}\left\{\boldsymbol{A}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right\} \quad \boldsymbol{k} = \omega \boldsymbol{n}$$

The time average of the product of field intensity can be calculated as

$$\overline{\boldsymbol{AB}} = \frac{1}{2}\mathrm{Re}\left\{\boldsymbol{A}_0 \boldsymbol{B}_0^*\right\}$$

The electric and magnetic field are

$$\boldsymbol{E} = ik\boldsymbol{A} \quad \boldsymbol{B} = i\boldsymbol{k} \times \boldsymbol{A}$$

And we can verify that $(\omega, \boldsymbol{k})$ transform like a four-vector.

Generally, the electric field can be written as

$$E_y = A\cos(\phi) \quad E_z = B\cos(\phi + \delta) \quad \phi = kx - \omega t \quad -\pi < \delta \leq \pi$$

The end of the vector $\boldsymbol{E}$ in $y - z$ plane will form an ellipse. The magnitudes of the semiaxes of the polarized ellipse are

$$|\sqrt{A^2 + B^2 + 2AB\sin\delta} \pm \sqrt{A^2 + B^2 - 2AB\sin\delta}|$$

The angle θ between the major axis and y axis satisfies the equation

$$\tan 2\theta = \frac{2AB\cos\delta}{A^2 - B^2}$$

If $-\frac{\pi}{2} < \delta < \frac{\pi}{2}$, major axis is in the second and forth quadrant. If $\delta > \frac{\pi}{2}$ or $\delta < -\frac{\pi}{2}$, major axis is in the first and third quadrant. If $\delta = \pm\frac{\pi}{2}$ and $A > B$, the major axis is y axis. If $\delta = \pm\frac{\pi}{2}$ and $A < B$, the major axis is z axis. If $\delta = \pm\frac{\pi}{2}$ and $A = B$, the ellipse becomes a circle.
If $0 < \delta < \pi$, the rotation is positive in the direction of x axis (left handed). If $-\pi < \delta < 0$, the rotation is negative in the direction of x axis(right handed). If $\delta = 0, \pi$, the ellipse becomes a line.

Any field is expandable in a Fourier integral containing a continuous or discrete distribution of different frequencies. Such an expansion has the form

$$f(t) = \int_{-\infty}^{\infty} f_\omega e^{-i\omega t}\frac{d\omega}{2\pi}$$

where the Fourier components are given in terms of the function $f(t)$ by the integrals

$$f_\omega = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$$

Because $f(t)$ must be real, so

$$f_{-\omega} = f_{\omega}^*$$

The total intensity of the wave is

$$\int_{-\infty}^{\infty} f^2 dt = \int_{-\infty}^{\infty} |f_\omega|^2 \frac{d\omega}{2\pi} = 2\int_0^{\infty} |f_\omega|^2 \frac{d\omega}{2\pi}$$

There is a special case that $f(t)$ is a periotic function with angular frequency ω_0, then $f(t)$ can be expanded as

$$f(t) = \sum_{-\infty}^{\infty} f_n e^{-in\omega_0}$$

here,

$$f_n = \frac{1}{T}\int_0^T f(t) e^{in\omega_0} dt$$

The average intensity of the wave is

$$\frac{1}{T}\int_0^T f^2 dt = \sum_{-\infty}^{\infty} |f_n|^2$$

And we can verify that

$$f_\omega = \sum_{-\infty}^{\infty} 2\pi f_n \delta(\omega - n\omega_0)$$

7.3.3 Partially polarized light

Every monochromatic wave is necessarily polarized. However we usually have to deal with waves which are only approximately monochromatic, and which contain frequencies in a small interval $\Delta\omega$. We consider such a wave, and let ω be some average frequency for it. Then its field at a fixed point in space can be written in the form

$$\boldsymbol{E_0}(t) e^{-i\omega t}$$

where the complex amplitude $\boldsymbol{E}_0$ is some slowly varying function of the time. Since $\boldsymbol{E}_0$ determines the polarization of the wave, this means that at each point of the wave, its polarization changes with time, such a wave is said to be partially polarized.

The polarization properties of electromagnetic waves are observed experimentally by passing the light to be investigated through various bodies and then observing the intensity of the transmitted light. From the mathematical point of view this means that we draw conclusions concerning the polarization properties of the light from the values of certain quadratic functions of its field. Here of course we are considering the time averages of such functions.

Quadratic functions of the field are made up of terms proportional to the products $E_\alpha E_\beta$, $E_\alpha^* E_\beta^*$ or $E_\alpha^* E_\beta$. Products of the form $E_\alpha E_\beta$ and $E_\alpha^* E_\beta^*$ contain the rapidly oscillating factors $e^{-i2\omega t}$ and will give zero when the time average is taken. Thus we see that the polarization properties of the light are completely characterized by the tensor

$$J_{\alpha\beta} = E_{0\alpha} E_{0\beta}^*$$

The trace of the tensor

$$J \equiv J_{\alpha\alpha} = \boldsymbol{E}_0 \boldsymbol{E}_0^*$$

determines the intensity of the wave, as measured by the energy flux density.To eliminate this quantity which is not directly related to the polarization properties, we introduce the tensor

$$\rho_{\alpha\beta} = \frac{J_{\alpha\beta}}{J}$$

we call it the polarization tensor.

Generally, the polarization tensor is expressed as

$$\rho = \frac{1}{2}\begin{bmatrix} 1+p_3 & p_1 - ip_2 \\ p_1 + ip_2 & 1-p_3 \end{bmatrix}$$

If we introduce the Pauli matrix, i.e.

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The ρ can be expressed as

$$\frac{1}{2}(1-P)I + \frac{1}{2}P(I + \boldsymbol{n}\cdot\boldsymbol{\sigma})$$

Here,

$$P = \sqrt{p_1^2 + p_2^2 + p_3^2} \quad \boldsymbol{n} = (\frac{p_1}{P}, \frac{p_2}{P}, \frac{p_3}{P})$$

For a totally polarized light with polarization state$|E\rangle = (\cos\frac{\theta}{2}e^{-i\frac{\phi}{2}}, \sin\frac{\theta}{2}e^{i\frac{\phi}{2}})$, the polarization tensor $\rho = |E\rangle\langle E|$. We can verify that

$$P = 1 \quad \boldsymbol{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

So, for an arbitrary light,

$$\rho = (1-P)\rho_n + P\rho_p \quad \rho_n = \frac{1}{2}I \quad \rho_p = \frac{1}{2}(I + \boldsymbol{n}\cdot\boldsymbol{\sigma})$$

Thus, we call P the Polarization degree.

Suppose there is a polarizing filter, which allow the light with polarization state

$$|D\rangle = (\cos\frac{\theta}{2}e^{-i\frac{\phi}{2}}, \sin\frac{\theta}{2}e^{i\frac{\phi}{2}})$$

to pass totally. If a light with polarization tensor ρ pass through the device, the relative intensity will become

$$\langle D|\rho|D\rangle = \frac{1}{2} + \frac{1}{2}\boldsymbol{p}\cdot\boldsymbol{m}$$

Here, $\boldsymbol{p} = (p_1, p_2, p_3)$, $\boldsymbol{m} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$

In optics, the Stokes vectors are defined as The Stokes parameters are defined by

$$\begin{aligned} I &\equiv \langle E_x^2\rangle + \langle E_y^2\rangle \\ &= \langle E_a^2\rangle + \langle E_b^2\rangle \\ &= \langle E_l^2\rangle + \langle E_r^2\rangle \\ Q &\equiv \langle E_x^2\rangle - \langle E_y^2\rangle \\ U &\equiv \langle E_a^2\rangle - \langle E_b^2\rangle \\ V &\equiv \langle E_l^2\rangle - \langle E_r^2\rangle \end{aligned}$$

where the subscripts refer to three different bases of the space of Jones vectors: the standard Cartesian basis $\hat{x}, \hat{y}$, a Cartesian basis rotated by 45° $\hat{a}, \hat{b}$, and a circular basis $\hat{l}, \hat{r}$. The symbols $\langle\cdot\rangle$ represent expectation values. The light can be viewed as a random variable taking values in the space C^2 of Jones vectors (E_1, E_2). It is easy to verify that

$$Q = Ip_3 \quad U = Ip_2 \quad V = Ip_1$$

7.4 The field of moving charges

7.4.1 Retarded potential

This time we impose the Lorenz Gauge $\partial_\mu A^\mu = 0$. Now, Maxwell equations would imply that

$$\partial^2 A^\mu = j^\mu$$

We can rewrite it in three dimension form,

$$\nabla^2 \boldsymbol{A} - \frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\boldsymbol{J} \quad \nabla^2\phi - \frac{\partial^2\phi}{\partial t^2} = -\rho$$

To find the particular solution, we divide the whole space into infinitely small regions and determine the field produced by the charges located in one of these volume elements. Because of the linearity of the field equations, the actual field will be the sum of the fields produced by all such elements.

The charge de in a given volume element is a function of the time. If we choose the origin of coordinates in the volume element under consideration, then the charge density is $de(t)\delta(\boldsymbol{R})$, where R is the distance from the origin. Thus we must solve the equation

$$\nabla^2\phi - \frac{\partial^2\phi}{\partial t^2} = -de(t)\delta(\boldsymbol{R})$$

The particular solution is

$$\phi = \frac{de(t-R)}{4\pi R}$$

For an arbitrary distribution of charges $\rho(\boldsymbol{r}', t)$, we have

$$\phi(\boldsymbol{r}, t) = \int \frac{\rho(\boldsymbol{r}', t - |\boldsymbol{r} - \boldsymbol{r}'|)}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|} dV'$$

Similarly we have for the vector potential

$$\boldsymbol{A}(\boldsymbol{r},t)=\int\frac{\boldsymbol{J}(\boldsymbol{r}',t-|\boldsymbol{r}-\boldsymbol{r}'|)}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|}dV'$$

The particular solution above is called retarded potential.

Now we consider an arbitrary charged particle with kinetic equation $\boldsymbol{r}=\boldsymbol{r}_0(t)$, so we have

$$\rho(\boldsymbol{r},t)=e\delta(\boldsymbol{r}-\boldsymbol{r}_0(t))$$

$$\phi(\boldsymbol{r},t)=\int\frac{e\delta(\boldsymbol{r}'-\boldsymbol{r}_0(t-|\boldsymbol{r}-\boldsymbol{r}'|))}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|}dV'=\frac{e}{4\pi R^*(1-\frac{\boldsymbol{R}^*\cdot\boldsymbol{v}^*}{R^*})}$$

Here,

$$\boldsymbol{R}^*=\boldsymbol{r}-\boldsymbol{r}_0(t^*)\quad \boldsymbol{v}^*=\boldsymbol{v}_0(t^*)\quad t^*=t-R^*$$

Similarly, we have

$$\boldsymbol{A}(\boldsymbol{r},t)=\frac{e\boldsymbol{v}^*}{4\pi R^*(1-\frac{\boldsymbol{R}^*\cdot\boldsymbol{v}^*}{R^*})}$$

The potential is called Lienard-Wiechert potentials.
We can derive that

$$\frac{\partial t^*}{\partial t}=\frac{1}{1-\frac{\boldsymbol{R}^*\cdot\boldsymbol{v}^*}{R^*}}\quad \boldsymbol{\nabla}t^*=\frac{\boldsymbol{R}^*}{R^*-\boldsymbol{R}^*\cdot\boldsymbol{v}^*}$$

With the help of the equation above, we can calculate the electric and magnetic field intensity.

$$\boldsymbol{E}=\frac{e}{4\pi(R^*-\boldsymbol{R}^*\cdot\boldsymbol{v}^*)^3}\left\{(1-v^{*2})(\boldsymbol{R}^*-R^*\boldsymbol{v}^*)+\boldsymbol{R}^*\times[(\boldsymbol{R}^*-R^*\boldsymbol{v}^*)\times\boldsymbol{a}^*]\right\}$$

$$\boldsymbol{B}=\frac{\boldsymbol{R}^*}{R^*}\times\boldsymbol{E}$$

The electric field consists of two parts of different type. The first term depends only on the velocity of the particle (and not on its acceleration) and varies at large distances like $\frac{1}{R^2}$. The second term depends on the acceleration, and for large R it varies like $\frac{1}{R}$. This latter term is related to the electromagnetic waves radiated by the particle.

7.4.2 Spectral resolution of the retarded potentials

Suppose

$$\rho(\boldsymbol{r},t)=\int_{-\infty}^{\infty}\rho_\omega(\boldsymbol{r})e^{-i\omega t}\frac{d\omega}{2\pi}\quad \phi(\boldsymbol{r},t)=\int_{-\infty}^{\infty}\phi_\omega(\boldsymbol{r})e^{-i\omega t}\frac{d\omega}{2\pi}$$

We can derive that

$$\phi_\omega(\boldsymbol{r})=\int dV'\int_{-\infty}^{\infty}dt\frac{\rho(\boldsymbol{r}',t)}{4\pi R}e^{i\omega(R+t)}$$

If there is just one point charge, we set

$$\rho=e\delta(\boldsymbol{r}-\boldsymbol{r}_0(t))$$

we can get

$$\phi_\omega(\boldsymbol{r}) = \int_{-\infty}^{\infty} dt \frac{e}{4\pi R(t)} e^{i\omega[R(t)+t]} \quad \boldsymbol{R}(t) = \boldsymbol{r} - \boldsymbol{r}_0(t)$$

Similarly, for vector potential, we have

$$\boldsymbol{A}_\omega(\boldsymbol{r}) = \int_{-\infty}^{\infty} dt \frac{e\boldsymbol{v}_0(t)}{4\pi R(t)} e^{i\omega[R(t)+t]} \quad \boldsymbol{R}(t) = \boldsymbol{r} - \boldsymbol{r}_0(t)$$

For electric and magnetic field, we have

$$\boldsymbol{E}_\omega = -\boldsymbol{\nabla}\phi_\omega + i\omega\boldsymbol{A}_\omega$$

$$\boldsymbol{B}_\omega = \boldsymbol{\nabla} \times \boldsymbol{A}_\omega$$

7.5 Radiation

7.5.1 Far field approximation

We consider the field produced by a system of moving charges at distances large compared with the dimensions of the system. We choose the origin of coordinates O anywhere in the interior of the system of charges. The radius vector from O to the point P, where we determine the field, we denote by $\boldsymbol{R_0}$, and the unit vector in this direction by $\boldsymbol{n}$. Let the radius vector of the charge element be $\boldsymbol{r}$, and the radius vector from de to the point P be $\boldsymbol{R}$. At large distances from the system of charges,$R_0 \gg r$, and we have approximately,

$$R \approx R_0 - \boldsymbol{n} \cdot \boldsymbol{r}$$

We substitute this for the retarded potentials. In the denominator of the integrands we can neglect $\boldsymbol{n} \cdot \boldsymbol{r}$ compared with R_0. In $t - R$, whether it is possible to neglect these terms is determined by how much the quantities e and $\boldsymbol{j}$ change during the time $\boldsymbol{n} \cdot \boldsymbol{r}$. The potentials of the field at large distances from the system of charges are

$$\phi(\boldsymbol{R}_0, t) = \frac{1}{4\pi R_0} \int \rho(\boldsymbol{r}, t - R_0 + \boldsymbol{n} \cdot \boldsymbol{r}) dV$$

$$\boldsymbol{A}(\boldsymbol{R}_0, t) = \frac{1}{4\pi R_0} \int \boldsymbol{J}(\boldsymbol{r}, t - R_0 + \boldsymbol{n} \cdot \boldsymbol{r}) dV$$

At sufficiently large distances from the system of charges, the field over small regions of space can be considered to be a plane wave. For this it is necessary that the distance be large compared not only with the dimensions of the system, but also with the wavelength of the electromagnetic waves radiated by the system. We refer to this region of space as the wave zone of the radiation. In wave zone, we have

$$\boldsymbol{B} = \frac{\partial \boldsymbol{A}}{\partial t} \times \boldsymbol{n} \quad \boldsymbol{E} = \left(\frac{\partial \boldsymbol{A}}{\partial t} \times \boldsymbol{n}\right) \times \boldsymbol{n}$$

The energy flux is given by the Poynting vector which, for a plane wave, is

$$\boldsymbol{S} = B^2 \boldsymbol{n}$$

The intensity of radiation into the element of solid angle do

$$dI = B^2 R_0^2 do$$

Since the field is inversely proportional to R_0, we see that the amount of energy radiated by the system in unit time into the element of solid angle do is the same for all distances. For the radiation produced by a single arbitrarily moving point charge, it turns out to be convenient to use the Lienard-Wiechert potentials. At large distances, we have

$$\boldsymbol{A}(\boldsymbol{R}_0, t) = \frac{e\boldsymbol{v}(t')}{4\pi R_0(1 - \boldsymbol{n}\cdot\boldsymbol{v}_0(t'))}$$

Here,

$$t' - \boldsymbol{n}\cdot\boldsymbol{r}_0(t') = t - R_0$$

For the spectral resolution of the field of the waves radiated by the system, we have

$$\boldsymbol{A}_\omega(\boldsymbol{R_0}) = \frac{e^{ikR_0}}{R_0}\int \boldsymbol{J}_\omega e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} dV$$

In wave zone, we have

$$\boldsymbol{B}_\omega = i\boldsymbol{k}\times\boldsymbol{A}_\omega \quad \boldsymbol{E} = \frac{1}{\omega}(\boldsymbol{k}\times\boldsymbol{A}_\omega)\times\boldsymbol{k}$$

Suppose $d\mathcal{E}_{\omega\boldsymbol{n}}$ is the energy radiated into the element of solid angle do in the form of waves with frequencies in the interval $d\omega$. We have

$$d\mathcal{E}_{\omega\boldsymbol{n}} = 2B_\omega^2 R_0^2 do\frac{d\omega}{2\pi}$$

For the radiation produced by a single arbitrarily moving point charge,we have

$$\boldsymbol{A}_\omega(\boldsymbol{R}_0) = \frac{ee^{i\omega R_0}}{4\pi R_0}\int_{-\infty}^{\infty} e^{i\omega[t-\boldsymbol{n}\cdot\boldsymbol{v}_0]}\boldsymbol{v}_0 dt$$

$$\boldsymbol{B}_\omega(\boldsymbol{R}_0) = \frac{e\omega e^{i\omega R_0}}{4\pi R_0}\int_{-\infty}^{\infty} e^{i\omega[t-\boldsymbol{n}\cdot\boldsymbol{v}_0]}\boldsymbol{n}\times d\boldsymbol{r}_0$$

7.5.2 Low velocity approximation

If the charge density (current density) distribution changes a little during time $\boldsymbol{r}\cdot\boldsymbol{n}$, then we can expand $f(r, t - R_0 + \boldsymbol{r}\cdot\boldsymbol{n})$ in series of $(\boldsymbol{r}\cdot\boldsymbol{n})^2$. In this case, if the typical angular frequency of the motion is ω, the typical scale of the motion is a, then we have

$$a \ll T \sim \frac{2\pi}{\omega} \sim \lambda$$

or

$$v \sim \frac{a}{T} \ll 1$$

which means that the velocity of the particles is much smaller then that of light, or the scale of the system is much smaller then the wavelength.
As for the zeroth order approximation, we just drop the $\boldsymbol{r}\cdot\boldsymbol{n}$ in the equation, so

$$\boldsymbol{A}(\boldsymbol{R}_0,t)=\frac{1}{4\pi R_0}\int \boldsymbol{J}(\boldsymbol{r},t-R_0)dt$$

For the radiation produced by arbitrarily moving point charges, we have

$$\boldsymbol{A}(\boldsymbol{R}_0,t)=\frac{1}{4\pi R_0}\sum e_a\boldsymbol{v}_a(\boldsymbol{r},t-R_0)=\frac{1}{4\pi R_0}\dot{\boldsymbol{d}}$$

So,

$$\boldsymbol{B}=\frac{1}{4\pi R_0}\ddot{\boldsymbol{d}}\times\boldsymbol{n}\quad \boldsymbol{E}=\frac{1}{4\pi R_0}(\ddot{\boldsymbol{d}}\times\boldsymbol{n})\times\boldsymbol{n}$$

Radiation of this kind is called dipole radiation. We note that a closed system of particles, for all of which the ratio of charge to mass is the same, cannot radiate (by dipole radiation).The intensity of the dipole radiation is

$$dI=\frac{\ddot{\boldsymbol{d}}^2}{16\pi^2}\sin^2\theta do$$

where θ is the angle between $\ddot{\boldsymbol{d}}$ and n. Integrate over all over the direction, we have

$$I=\frac{\ddot{\boldsymbol{d}}^2}{6\pi}$$

If we have just one charge moving in the external field, we have

$$I=\frac{e^2w^2}{6\pi}$$

$\boldsymbol{w}$ is the acceleration of the charge. For the spectral resolution of the intensity of dipole radiation, we have

$$d\mathcal{E}_\omega=\frac{|(\ddot{\boldsymbol{d}})_\omega|^2}{3\pi}=\frac{\omega^4}{3\pi}|\boldsymbol{d}_\omega|^2\frac{d\omega}{2\pi}$$

More details on dipole radiation during collisions and Coulomb interaction can be found in section 68,69 and 70 of *The classical theory of fields (L.D.Landau & E.M.Lifshitz)*.

If we keep the first order of $\boldsymbol{n}\cdot\boldsymbol{r}$, we have

$$\boldsymbol{A}=\frac{\dot{\boldsymbol{d}}}{4\pi R_0}+\frac{\ddot{\boldsymbol{D}}}{24\pi R_0}+\frac{\dot{\boldsymbol{m}}\times\boldsymbol{n}}{4\pi R_0}$$

Here,

$$\boldsymbol{D}=\sum e[3\boldsymbol{r}(\boldsymbol{n}\cdot\boldsymbol{r}-\boldsymbol{n}r^2)]\quad \boldsymbol{m}=\frac{1}{2}\sum e\boldsymbol{r}\times\boldsymbol{v}$$

So

$$\boldsymbol{B}=\frac{1}{4\pi R_0}\left[\ddot{\boldsymbol{d}}\times\boldsymbol{n}+\frac{1}{6}\dddot{\boldsymbol{D}}\times\boldsymbol{n}+(\ddot{\boldsymbol{m}}\times\boldsymbol{n})\times\boldsymbol{n}\right]$$

$$\boldsymbol{E}=\frac{1}{4\pi R_0}\left[(\ddot{\boldsymbol{d}}\times\boldsymbol{n})\times\boldsymbol{n}+\frac{1}{6}(\dddot{\boldsymbol{D}}\times\boldsymbol{n})\times\boldsymbol{n}+\boldsymbol{n}\times\ddot{\boldsymbol{m}}\right]$$

$$I = \frac{1}{6\pi}\ddot{\boldsymbol{d}}^2 + \frac{1}{720\pi}\dddot{D}_{ij}^2 + \frac{1}{6\pi}\ddot{\boldsymbol{m}}^2 \quad D_{ij} = \sum e(3r_i r_j - r^2\delta_{ij})$$

The total radiation consists of three independent parts: dipole, quadrupole, and magnetic dipole radiation. The details of the derivation can be found in section 71 of *The classical theory of fields (L.D.Landau & E.M.Lifshitz)*.

7.5.3 Radiation from a rapidly moving charge

Consider the particle in the system of reference in which the particle is at rest at a given moment; in this system of reference we can apply low velocity approximation. Here, we have

$$d\mathcal{E} = \frac{e^2 w^2}{6\pi}dt \quad d\boldsymbol{P} = 0$$

So, in an arbitrary reference, we have

$$dP^\mu = -\frac{e^2}{6\pi}\frac{du^\nu}{d\tau}\frac{du_\nu}{d\tau}u^\mu d\tau$$

Recall the Lorentz equation, the total four-momentum radiated during the time of passage of the particle through a given electromagnetic field is equal to

$$\Delta P^\mu = -\frac{e^4}{6\pi m^2}\int F_{\nu\rho}u^\rho F^{\nu\sigma}u_\sigma dx^\mu$$

Particularly, we have

$$\Delta\mathcal{E} = \frac{e^2}{6\pi}\int\frac{w^2 - (\boldsymbol{v}\times\boldsymbol{w})^2}{(1-v^2)^3}dt = \frac{e^4}{6\pi m^2}\int\frac{(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})^2 - (\boldsymbol{E}\cdot\boldsymbol{v})^2}{1-v^2}dt$$

It is clear from formula above that for velocities close to the velocity of light, the total energy radiated per unit time is proportionally to the square of the energy of the moving particle. The only exception is motion in an electric field, along the direction of the field. In this case the factor $(1-v^2)$ standing in the denominator is cancelled by an identical factor in the numerator, and the radiation does not depend on the energy of the particle.

Now we discuss the angular distribution of the radiation from a rapidly moving charge. The radiation electric field is

$$\boldsymbol{E} = \frac{e}{4\pi R}\frac{\boldsymbol{n}\times[(\boldsymbol{n}-\boldsymbol{v})\times\boldsymbol{w}]}{(1-\boldsymbol{n}\cdot\boldsymbol{v})^3} \quad \boldsymbol{B} = \boldsymbol{n}\times\boldsymbol{E}$$

where all the quantities on the right sides of the equations refer to the retarded time $t' = t - \frac{R}{c}$. The intensity radiated into the solid angle do is

$$dI = \frac{e^2}{16\pi^2}\left\{\frac{2(\boldsymbol{n}\cdot\boldsymbol{w})(\boldsymbol{v}\cdot\boldsymbol{w})}{(1-\boldsymbol{v}\cdot\boldsymbol{n})^5} + \frac{\boldsymbol{w}^2}{(1-\boldsymbol{v}\cdot\boldsymbol{n})^4} - \frac{(1-v^2)(\boldsymbol{n}\cdot\boldsymbol{w})^2}{(1-\boldsymbol{v}\cdot\boldsymbol{n})^6}\right\}do$$

If we want to determine the angular distribution of the total radiation throughout the whole motion of the particle, we must integrate the intensity over the time. In doing this, it is important to remember that the integrand is a function of t'; therefore we must write

$$dt = (1-\boldsymbol{n}\cdot\boldsymbol{v})dt'$$

after which the integration over t' is immediately done.
In the ultrarelativistic case, the intensity is large within the narrow range of angles in which $1-\boldsymbol{v}\cdot\boldsymbol{n}$ is small. Thus an ultrarelativistic particle radiates mainly along the direction of its own motion, within the small range of angles around the direction of its velocity. We also point out that, for arbitrary velocity and acceleration of the particle, there are always two directions for which the radiated intensity is zero. These are the directions for which the vector $\boldsymbol{n}-\boldsymbol{v}$ is parallel to the vector $\boldsymbol{w}$.

If the velocity and acceleration of the particle are parallel,

$$\boldsymbol{B} = \frac{e}{4\pi R}\frac{\boldsymbol{w}\times\boldsymbol{n}}{(1-\boldsymbol{v}\cdot\boldsymbol{n})^3}$$

and the intensity is

$$dI = \frac{e^2}{16\pi^2}\frac{w^2\sin^2\theta}{(1-v\cos\theta)^6}$$

It is naturally, symmetric around the common direction of $\boldsymbol{v}$ and $\boldsymbol{w}$, and vanishes along ($\theta = 0$) and opposite to ($\theta = \pi$) the direction of the velocity. In the ultrarelativistic case, the intensity as a function of θ has a sharp double maximum near $\boldsymbol{v}$, with a steep drop to zero for $\theta = 0$.

If the velocity and acceleration are perpendicular to one another,

$$dI = \frac{e^2w^2}{16\pi^2}\left[\frac{1}{(1-v\cos\theta)^4} - \frac{(1-v^2)\sin^2\theta\cos^2\phi}{(1-v\cos\theta)^6}\right]$$

where θ is again the angle between $\boldsymbol{v}$ and $\boldsymbol{n}$, and ϕ is the azimuthal angle of the vector $\boldsymbol{n}$ relative to the plane passing through $\boldsymbol{v}$ and $\boldsymbol{w}$.

The discussion of synchrotron radiation (magnetic bremsstrahlung) can be found in section 74 of *The classical theory of fields (L.D.Landau & E.M.Lifshitz)*.

7.6 The interaction between charged particles and EM field

7.6.1 Radiation reaction

If a charged particle accelerates, it radiates away energy. This means that if an external force is applied to a charge, not all of the energy transferred to the charge by the force is converted to the kinetic energy of the charge; some of the energy is radiated away in the form of electromagnetic waves. From Newton's law $\boldsymbol{F} = m\boldsymbol{a}$, the net force on the charge must be less than the applied external force. In effect, the fields surrounding the charge exert a recoil or reaction force on the charge.
The fields of a moving point charge are

$$\boldsymbol{E} = \frac{e}{4\pi(\boldsymbol{R}^*\cdot\boldsymbol{u}^*)^3}\left\{(1-v^{*2})R^*\boldsymbol{u} + \boldsymbol{R}^*\times[R^*\boldsymbol{u}^*\times\boldsymbol{a}^*]\right\} \qquad \boldsymbol{B} = \frac{\boldsymbol{R}^*}{R^*}\times E$$

Here,$\boldsymbol{u}^* = \frac{\boldsymbol{R}^*}{R^*} - v^*$. It is only those terms that go as $\frac{1}{R}$ that contribute to radiation energy. The other term falls off as $\frac{1}{R^2}$ so contributes nothing to the integral of the Poynting vector over a large sphere. This term is called the velocity field and, although it doesn't contribute to the energy radiated away by the EM field, it does store energy, so some of the energy imparted by the force that gets the charge moving must be siphoned off to create these velocity fields. These velocity fields are curious beasts, however, for they contain energy that is never actually lost to the charge. If a charge is accelerated to some velocity, the velocity fields are constructed around the moving charge, but if the charge is then decelerated to rest again, the velocity fields disappear without having radiated away any energy. It would seem to be reabsorbed by the charge as it slows down.

If we look at a charged particle that starts off in some state, then goes through an acceleration followed by a deceleration, and finally ends up in the same state that it started from, what we can say is that the velocity fields are the same at the end as they were at the start, so over this period, the only energy that is truly lost from the particle is the energy that is radiated away. In non-relativistic case,

$$\int_{t_1}^{t_2} \boldsymbol{F}_{rad} \cdot \boldsymbol{v} dt = -\int_{t_1}^{t_2} P dt = -\frac{e^2}{6\pi} \int_{t_1}^{t_2} a^2 dt$$

Since

$$\int_{t_1}^{t_2} a^2 dt = \int_{t_1}^{t_2} \dot{\boldsymbol{v}} \cdot \dot{\boldsymbol{v}} dt = \boldsymbol{v} \cdot \dot{\boldsymbol{v}}|_{t_1}^{t_2} - \int_{t_1}^{t_2} \boldsymbol{v} \cdot \ddot{\boldsymbol{v}} dt = -\int_{t_1}^{t_2} \boldsymbol{v} \cdot \ddot{\boldsymbol{v}} dt$$

So,we have

$$\boldsymbol{F}_{rad} = \frac{e^2}{6\pi} \cdot \boldsymbol{a}$$

This is known as the Abraham-Lorentz formula for radiation reaction. This equation can only be applied when the frequency and intensity of the EM field is not very big, i.e.

$$\lambda \gg \frac{e^2}{m} \quad H \ll \frac{m^2}{e^3}$$

The details can be found in section 75 of *The classical theory of fields (L.D.Landau & E.M.Lifshitz)*.

We derive the relativistic expression for the radiation damping for a single charge, which is applicable also to motion with velocity comparable to that of light. This force is now a four-vector g^μ, which must be included in the equation of motion of the charge, written in four-dimensional form:

$$m\frac{du^\mu}{d\tau} = eF^{\mu\nu}u_\nu + g^\mu$$

To determine g^μ we note that for $v \ll 1$, its three space components must go over into the components of the vector $\frac{e^2}{6\pi} \cdot \boldsymbol{a}$. It is easy to see that the vector $\frac{e^2}{6\pi}\frac{d^2u^\mu}{d\tau^2}$ has this property. However, it does not satisfy the identity $g^\mu u_\mu = 0$, which is valid for any force four-vector. In order to satisfy this condition, we must add to the expression given a certain auxiliary four-vector, made up from the four-velocity u^μ and its derivatives. The three space components of this vector must become zero in the limiting case $v = 0$. As a result we find

$$g^\mu = \frac{e^2}{6\pi}\left(\frac{d^2u^\mu}{d\tau^2} - u^\mu u^\nu \frac{d^2u^\nu}{d\tau^2}\right)$$

It is called Abraham–Lorentz–Dirac force.
The integral of the four-force g^μ over the world line of the motion of a charge, passing through a given field, must coincide (except for opposite sign) with the total four-momentum ΔP^μ of the radiation from the charge. The first term in equation above goes to zero on performing the integration, since at infinity the particle has no acceleration. We integrate the second term by parts and get:

$$-\int g^\mu d\tau = -\frac{e^2}{6\pi}\int \frac{du^\nu}{d\tau}\frac{du_\nu}{d\tau}u^\mu d\tau = \Delta P^\mu$$

7.6.2 Scattering by free charges

If an electromagnetic wave falls on a system of charges, then under its action the charges are set in motion. This motion in turn produces radiation in all directions; there occurs, we say, a scattering of the original wave. The scattering is most conveniently characterized by the ratio of the amount of energy emitted by the scattering system in a given direction per unit time, to the energy flux density of the incident radiation. This ratio clearly has dimensions of area, and is called the reflective scattering cross section. Let dI be the energy radiated by the system into solid angle do per second for an incident wave with Poynting vector $\boldsymbol{S}$. Then the effective cross-section for scattering (into the solid angle do) is

$$d\sigma = \frac{d\bar{I}}{\bar{S}}$$

(the dash over a symbol means a time average). The integral σ of $d\sigma$ over all directions is the total scattering cross-section.
Let us consider the scattering produced by a free charge at rest. Suppose there is incident on this charge an approximately plane monochromatic wave (partially polarized light). We shall assume that the velocity acquired by the charge under the influence of the incident wave is small compared with the velocity of light. Then we can neglect the force exerted by magnetic field. We also assume the wavelength of the EM field is much larger than the displacement of the charge during its vibrations. So, we have

$$m\ddot{\boldsymbol{r}} = e\boldsymbol{E}_0(t)e^{-i\omega t}$$

and

$$\ddot{\boldsymbol{d}} = \frac{e^2}{m}\boldsymbol{E}_0(t)e^{-i\omega t}$$

Now we assume the incident direction of the EM wave is $\hat{x}$, the scattered direction of the EM wave is $\boldsymbol{n}' = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. The dipole radiation is

$$\overline{dI} = \frac{e^4}{16\pi^2 m^2}\overline{(\boldsymbol{E}\times\boldsymbol{n}')^2}do$$

Since

$$(\boldsymbol{E}\times\boldsymbol{n}')^2 = -2\,E_yE_z\cos(\theta)\sin(\phi)\sin(\theta) - \left({E_y}^2\sin(\phi)^2 - {E_z}^2\right)\sin(\theta)^2 + {E_y}^2$$

we have

$$\overline{dI} = \frac{e^4\bar{S}}{16\pi^2 m^2}[-(\rho_{12}+\rho_{21})\cos(\theta)\sin(\phi)\sin(\theta) - \left(\rho_{11}\sin(\phi)^2 - \rho_{22}\right)\sin(\theta)^2 + \rho_{11}]$$

Here

$$\bar{S} = \overline{\mathrm{Re}(\boldsymbol{E}) \cdot \mathrm{Re}(\boldsymbol{E})} = \frac{1}{2}\mathrm{Re}\overline{\boldsymbol{E}_0 \cdot \boldsymbol{E}_0^*}$$

The effective cross-section for scattering is

$$d\sigma = \frac{e^4}{16\pi^2 m^2}[-(\rho_{12} + \rho_{21})\cos(\theta)\sin(\phi)\sin(\theta) - \left(\rho_{11}\sin(\phi)^2 - \rho_{22}\right)\sin(\theta)^2 + \rho_{11}]do$$

The total cross section is

$$\sigma = \frac{8\pi}{3}\left(\frac{e^2}{4\pi m}\right)^2$$

If the incident light is totally linear polarized in $\hat{z}$ direction, then we have

$$d\sigma = \frac{e^4}{16\pi^2 m^2}\sin^2\theta do$$

If the incident light is unpolarized, we have

$$d\sigma = \frac{e^4}{32\pi^2 m^2}\left(1 + \cos^2\Theta\right)do$$

Here, $\cos\Theta = \cos\phi\sin\theta$, Θ is the angle between the direction of incident light and scatted light.

Scattering by bound charges

The dynamic equation of the bound charges are

$$\ddot{\boldsymbol{\xi}} = \frac{e}{m}\boldsymbol{E}_0 e^{-i\omega t} - \omega_0^2\boldsymbol{\xi} + \frac{e^2}{6\pi m}\dddot{\boldsymbol{\xi}}$$

Suppose $\boldsymbol{\xi} = \boldsymbol{\xi}_0 e^{-i\omega t}$, we can get

$$\boldsymbol{\xi} = \frac{e\boldsymbol{E}_0}{m(\omega_0^2 - \omega^2 - i\omega\gamma)}e^{-i\omega t}$$

Here,

$$\gamma = \frac{e^2\omega^2}{6\pi m}$$

We can show that

$$\sigma = \sigma_0\frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}$$

Here, σ_0 is the total cross section when EM wave is scattered by free charges. When $\omega \gg \omega_0$, σ is independent of ω,i.e. $\sigma \approx \sigma_0$. When $\omega \ll \omega_0$, σ_0 is proportional to ω^4, i.e. $\sigma = \sigma_0\left(\frac{\omega}{\omega_0}\right)^4$, and it is called Rayleigh scattering.

Part III

General relativity

Chapter 8

Elementary Differential Geometry

8.1 Fundamental conception on differential manifolds

Definition 8.1 Manifold

Manifold *Formally, a topological manifold is a second countable Hausdorff space that is locally homeomorphic to Euclidean space.*
Differentiable manifold *In formal terms, a differentiable manifold is a topological manifold with a globally defined differential structure.*
Tangent space *In mathematics, the tangent space of a manifold facilitates the generalization of vectors from affine spaces to general manifolds, since in the latter case one cannot simply subtract two points to obtain a vector pointing from one to the other.*
Cotangent space *Typically, the cotangent space is defined as the dual space of the tangent space at x.*

Definition 8.2 Submanifold

Submanifold
Immersed submanifolds *An immersed submanifold of a manifold M is the image S of an immersion map $f : N \to M$; in general this image will not be a submanifold as a subset, and an immersion map need not even be injective (one-to-one) – it can have self-intersections.*
Injective immersion submanifolds *More narrowly, one can require that the map $f : N \to M$ be an inclusion (one-to-one), in which we call it an injective immersion, and define an immersed submanifold to be the image subset S together with a topology and differential structure such that S is a manifold and the inclusion f is a diffeomorphism: this is just the topology on N, which in general will not agree with the subset topology: in general the subset S is not a submanifold of M, in the subset topology.*
Open submanifolds
Closed submanifolds

Definition 8.3 Embedded Submanifold

An embedded submanifold (also called a regular submanifold), is an immersed submanifold for which the inclusion map is a topological embedding. That is, the submanifold topology on S is the same as the subspace topology. Given any embedding $f : N \to M$ of a manifold N in M the image $f(N)$ naturally has the structure of an embedded submanifold. That is, embedded submanifolds are precisely the images of embeddings.

Proposition 8.1

If an n dimensional injective immersed submanifold N of a m dimensional manifold M is a closed submanifold of an open submanifold of M, then for every point $p \in f(N)$ there exists a chart ($U \subset M$,$\phi : U \to R_n$) containing p such that $\phi(f(N) \cap U)$ is the intersection of a n-dimensional plane with $\phi(U)$.
Closed submanifolds of an open submanifold are equal to embedded submanifolds.

8.2 Multi linear algebra

Definition 8.4 Tensor

Vector space

Dual space

In mathematics, any vector space V has a corresponding dual vector space (or just dual space for short) consisting of all linear functionals on V together with a naturally induced linear structure.

Tensor product

$$V \otimes W = \mathit{Span}\{v \otimes w\} = \mathcal{L}(V^*, W^*; F)$$

$$V^* \otimes W^* = \mathit{Span}\{v^* \otimes w^*\} = \mathcal{L}(V, W; F)$$

$$\mathcal{L}(V, W; Z) = \mathcal{L}(V \otimes W; Z)$$

$$(\phi \otimes \psi) \otimes \xi = \phi \otimes (\psi \otimes \xi)$$

Tensor

$$V^r_s = V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

$$x = x^{i_1 \cdots i_r}{}_{k_1 \cdots k_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{*k_1} \otimes \cdots \otimes e^{*k_s}$$

$$(x \otimes y)^{i_1 \cdots i_{r_1+r_2}}{}_{k_1 \cdots k_{s_1+s_2}} = x^{i_1 \cdots i_{r_1}}{}_{k_1 \cdots k_{s_1}} \cdot y^{i_{r_1+1} \cdots i_{r_1+r_2}}{}_{k_{s_1+1} \cdots k_{s_1+s_2}}$$

Definition 8.5 Symmtric and antisymmetric tensor

Permutation$(\sigma \in \mathcal{P}(r))$

$$\sigma x(v^{*1}, \cdots, v^{*r}) = x(v^{*\sigma(1)}, \cdots, v^{*\sigma(r)})$$

Symmetric contra-variant tensor

$$\sigma x = x$$

Antisymmetric contra-variant tensor

$$\sigma x = \ sgn\ \sigma \cdot x$$

Symmetrization operator

$$S_r(x) = \frac{1}{r!} \sum_{\sigma \in \mathcal{P}(x)} \sigma x$$

Antisymmetrization operator

$$A_r(x) = \frac{1}{r!} \sum_{\sigma \in \mathcal{P}(x)} sgn\ \cdot \sigma x$$

Definition 8.6 Exterior vector space

Exterior vector space

$$\Lambda^r(V) = A_r(T^r(V))$$

$$\Lambda^0(V) = F \quad \Lambda^1(V) = V$$

Wedge product

$$\xi \wedge \eta \equiv \frac{(k+l)!}{k!l!} A_{k+l}(\xi \otimes \eta)$$

Pull-back mapping $f : V \to W$ *is a linear mapping, we define* $f^* : \Lambda^r(W^*) \to \Lambda^r(V^*)$ *as*

$$f^*\phi(v_1, \cdots, v_r) = \phi(f(v_1), \cdots, f(v_r)).$$

Proposition 8.2 Properties of Wedge product

$$(\xi_1 + \xi_2) \wedge \eta = \xi_1 \wedge \eta + \xi_2 \wedge \eta$$

$$\xi \wedge (\eta_1 + \eta_2) = \xi \wedge \eta_1 + \xi \wedge \eta_2$$

$$\xi \wedge \eta = (-1)^{kl} \eta \wedge \xi$$

$$(\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta) = \frac{(k+l+h)!}{k!l!h!} A_{k+l+h}(\xi \otimes \eta \otimes \zeta)$$

$$f^*(\phi \wedge \psi) = f^*\phi \wedge f^*\psi$$

Proposition 8.3 Properties of exterior space

$$e_{i_1} \wedge \cdots \wedge e_{i_r}(v^{*1}, \cdots, v^{*r}) = \det\langle e_{i_\alpha}, v^{*\beta}\rangle$$
$$e_{i_1} \wedge \cdots \wedge e_{i_r}(e^{*j_1}, \cdots, e^{*j_r}) = \det\langle e_{i_\alpha}, e^{*j_\beta}\rangle = \delta^{j_1\cdots j_r}_{i_1\cdots i_r}$$
$$\Lambda^r(V) = \text{Span}\,\{e_{i_1} \wedge \cdots \wedge e_{i_r}, 1 \le i_1 < \cdots < i_r \le n\}$$
$$(\Lambda^r(V))^* = \Lambda^r(V^*)$$

8.3 Vector Bundle

Definition 8.7 Fiber bundle

Fiber bundle *In mathematics, and particularly topology, a fiber bundle is a space that is locally a product space, but globally may have a different topological structure. Specifically, the similarity between a space E and a product space B × F is defined using a continuous surjective map $\pi : E \to B$ that in small regions of E behaves just like a projection from corresponding regions of $B \times F$ to B. The map π, called the projection or submersion of the bundle, is regarded as part of the structure of the bundle. The space E is known as the total space of the fiber bundle, B as the base space, and F the fiber.*

Vector Bundle *In mathematics, a vector bundle is a topological construction that makes precise the idea of a family of vector spaces parameterized by another space X (for example X could be a topological space, a manifold, or an algebraic variety): to every point x of the space X we associate (or "attach") a vector space $V(x)$ in such a way that these vector spaces fit together to form another space of the same kind as X (e.g. a topological space, manifold, or algebraic variety), which is then called a vector bundle over X.*

Tangent bundle *In differential geometry, the tangent bundle of a differentiable manifold M is a manifold TM, which assembles all the tangent vectors in M. As a set, it is given by the disjoint union of the tangent spaces of M. That is,*

$$TM = \bigsqcup_{x\in M} T_xM = \bigcup_{x\in M} \{x\} \times T_xM = \bigcup_{x\in M} \{(x, y)|y \in T_xM\}$$

where T_xM denotes the tangent space to M at the point x. So, an element of TM can be thought of as a pair (x, v) , where x is a point in M and v is a tangent vector to M at x. There is a natural projection $\pi : TM \to M$ defined by $\pi(x, v) = x$. This projection maps each tangent space T_xM to the single point x. A section of TM is a vector field on M, and the dual bundle to TM is the cotangent bundle, which is the disjoint union of the cotangent spaces of M.

Cotangent bundle $T^*M = \bigcup_{x\in M} T^*_xM$

Tensor bundle $T^r_sM = \bigcup_{x\in M} T^r_{sx}M$

8.4 Tangent vector field

Theorem 8.1

Let M be a smooth manifold, and let $Y : M \to TM$ be a vector field. If $(U, [x_i])$ is an arbitrary smooth coordinate chart on M, then Y is smooth on U if and only if its component functions with respect to this chart are smooth.

Theorem 8.2

Let M be a m dimensional smooth manifold and v a smooth tangent vector field on M. $v : C^\infty(M) \to C^\infty$ satisfy that
(1) $\forall f, g \in C^\infty(M), v(f+g) = v(f) + v(g)$;
(2) $\forall f \in C^\infty(M), \alpha \in \boldsymbol{R}, v(\alpha f) = \alpha \cdot v(f)$;
(3) $\forall f, g \in C^\infty(M), v(f \cdot g) = f \cdot v(g) + g \cdot v(f)$.
If $\alpha : C^\infty(M) \to C^\infty(M)$ satisfy the three conditions above, there exists a unique smooth vector field v on M that $\forall f \in C^\infty(M), v(f) = \alpha(f)$.

Theorem 8.3

$\forall X, Y \in \mathcal{H}(M), [X, Y] = X \circ Y - Y \circ X \in \mathcal{H}(M)$.

Proposition 8.4

(1) $[aX + bY, Z] = a[X, Z] + b[Y, Z]$;$[Z, aX + bY] = a[Z, X] + b[Z, Y]$;
(2) $[X, Y] = -[Y, X]$;
(3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$;
(4) $[X, Y]|_U = [X|_U, Y|_U] = (X_i \frac{\partial Y^j}{\partial u^i} - Y^i \frac{\partial X^j}{\partial u^i}) \frac{\partial}{\partial u^j}$;
(5) $f_*[X, Y] = [f_*X, f_*Y]$;

Definition 8.8 One parameter differentiable transformation group

Let M be a smooth manifold and $\phi : \boldsymbol{R} \times M \to M$ a smooth mapping, and $\forall (t, p) \in \boldsymbol{R} \times M$,denote $\phi_t(p) = \phi(t, p)$. If ϕ satisfy that
(1) $\phi_0 = \mathrm{id} : M \to M$;
(2) $\forall s, t \in \boldsymbol{R}, \phi_s \circ \phi_t = \phi_{s+t}$;
then ϕ is called a one parameter differentiable transformation group acting on M.
Trajectory of ϕ through p on M:$\gamma_p(t) = \phi(t, p)$.
Vector field induced by ϕ: $X_p(f) = \langle \gamma_p, f \rangle$.

Proposition 8.5

(1) $\gamma_q(t) = \phi(t, \phi(s,p)) = \phi(t+s, p) = \gamma_p(t+s)$;
(2) $(\phi_s)_* X_p = X_{\phi_s(p)}$;
(3) $\psi_* X_p = \tilde{X}_{\psi(p)}$ if X is induced by ϕ and $\tilde{X}$ is induced by $\psi \circ \phi \circ \psi^{-1}$. ψ is a smooth homeomorphism.
(4) $[X, Y] = \lim_{t\to 0} \frac{Y_p - (\phi_t)_* Y_{\phi_{-t}(p)}}{t} = \lim_{t\to 0} \frac{(\phi_{-t})_* Y_{\phi_t(p)} - Y}{t}$ if X is induced by ϕ.

Definition 8.9 Lie derivative

$$\mathcal{L}_X Y \equiv \lim_{t\to 0} \frac{(\phi_{-t})_* Y_{\phi_t(p)} - Y}{t} = [X, Y]$$
$$\mathcal{L}_X f \equiv X(f)$$

Proposition 8.6

$$\mathcal{L}_X(Y + \lambda Z) = \mathcal{L}_X Y + \lambda \mathcal{L}_X Z$$
$$\mathcal{L}_X(f \cdot Y) = \mathcal{L}_X(f) \cdot Y + f \mathcal{L}_X Y$$
$$\mathcal{L}_X([Y, Z]) = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]$$

Theorem 8.4

Let M be a n-dimensional smooth manifold and $X \in \mathcal{H}(M)$. If $p \in M$ and $X_p \neq 0$, $\exists (V, x^i)$ and $p \in V$ that $X|_V = \frac{\partial}{\partial y^1}$

Definition 8.10 Distribution

Distribution *Let M be a C^∞ manifold of dimension m, and let $n \leq m$. Suppose that for each $x \in M$, we assign an n-dimensional subspace $\Delta_x \subset T_x(M)$ of the tangent space in such a way that for a neighbourhood $N_x \subset M$ of x there exist n linearly independent smooth vector fields $X_1, \ldots, X_n$ such that for any point $y \in N_x$, $X_1(y), \ldots, X_n(y)$ span Δ_y. We let Δ refer to the collection of all the Δ_x for all $x \in M$ and we then call Δ a distribution of dimension n on M, or sometimes a C^∞ n-plane distribution on M. The set of smooth vector fields $\{X_1, \ldots, X_n\}$ is called a local basis of Δ.*

Definition 8.11 Involutive distributions

We say that a distribution Δ on M is involutive if for every point $x \in M$ there exists a local basis $\{X_1, \ldots, X_n\}$ of the distribution in a neighbourhood of x such that for all $1 \leq i, j \leq n$, $[X_i, X_j]$ is in the span of $\{X_1, \ldots, X_n\}$.That is, if $[X_i, X_j]$ is a linear combination of $\{X_1, \ldots, X_n\}$. Normally this is written as $[\Delta, \Delta] \subset \Delta$.
Involutive distributions are the tangent spaces to foliations. Involutive distributions are important in that they satisfy the conditions of the Frobenius theorem, and thus lead to integrable systems. A related idea occurs in Hamiltonian mechanics: two functions f and g on a symplectic manifold are said to be in mutual involution if their Poisson bracket vanishes.

Theorem 8.5 Frobenius Theorem

If distribution Δ on M is involutive, then $\forall p \in M$, $\exists (V, x^i)$ and $p \in V$ that $\Delta|_V = \mathrm{Span}\{\frac{\partial}{\partial y^1}, \cdots, \frac{\partial}{\partial y^h}\}$.

Definition 8.12 Integrable manifold

Let L^h be a smooth distribution on M. If $\phi : N \to M$ is an injective immersion manifold, and $\forall p \in N$, $\phi_*(T_pN) \subset L^h(\phi(p))$, then (ϕ, N) is called an integrable manifold of L^h. If $\forall q \in M$, there is an integrable manifold of L^h through it, we say that L^h is completely integrable.

Theorem 8.6

Let

$$\tau : \underbrace{A^1(M) \times \cdots \times A^1(M)}_{p} \times \underbrace{\mathcal{H}(M) \times \cdots \times \mathcal{H}(M)}_{q} \to c^\infty(M)$$

be a $p+q$ multi-linear mapping, if $\forall 1 \leq a \leq p, 1 \leq b \leq q$ and $\mu \in C^\infty(M)$,

$$\begin{aligned} & \tau(\alpha^1, \cdots, \mu\alpha^a, \cdots, \alpha^p, v_1, \cdots, v_q) \\ = \; & \tau(\alpha^1, \cdots, \alpha^p, v_1, \cdots, \mu v_b, \cdots, v_q) \\ = \; & \mu \cdot \tau(\alpha^1, \cdots, \alpha^p, v_1, \cdots, v_q) \end{aligned}$$

then the mapping τ define a (p, q) tensor for all $x \in M$ smoothly.

Definition 8.13 Lie derivatives

Let X be a smooth tangent vector field on M and ϕ_t the one parameter differentiable transformation group inducing it. Denote the trajectory of ϕ_t through x by $\gamma_x(t)$. So we have linear isomorphism

$$(\phi_t^{-1})_* = (\phi_{-t})_* : T_{\gamma_x(t)}M \to T_xM$$

$$(\phi_t)^* : T^*_{\gamma_x(t)} \to T_xM$$

So we can induce the linear isomorphism

$$\Phi_t : T^p_q(\gamma_x(t)) \to T^p_q(x)$$

If S and T are smooth tensor fields on M,
(1) for all t which is small enough, $\Phi_t S$ is a smooth tensor field on M which has the same type as S ,and $\lim_{t\to 0} \Phi_t(S(\gamma_p(t))) = S(p), \forall p \in M$.
(2)$\Phi_t(S \otimes T) = \Phi_t S \otimes \Phi_t T$.
(3)$\Phi_t(C^a_b(S)) = C^a_b(\Phi_t(S))$, C^a_b is a tag for contraction.
So,we can define the Lie derivative for smooth tensor field τ on M as

$$\mathcal{L}_X(\tau) = \lim_{t\to 0} \frac{\Phi_t(\tau) - \tau}{t}$$

Proposition 8.7

$$\mathcal{L}_X(\tau_1 + \lambda\tau_2) = \mathcal{L}_X\tau_1 + \lambda\mathcal{L}_X\tau_2$$

$$\mathcal{L}_X(\tau_1 \otimes \tau_2) = \mathcal{L}_X\tau_1 \otimes \tau_2 + \tau_1 \otimes \mathcal{L}_X\tau_2$$

$$C^r_s(\mathcal{L}_X\tau) = \mathcal{L}_X(C^r_s(\tau))$$

$$(\mathcal{L}_X\omega)(Y) = X(\omega(Y)) - \omega([X,Y])$$

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X$$

$$\mathcal{L}_{X+Y} = \mathcal{L}_X + \mathcal{L}_Y$$

Proposition 8.8

$$((\mathcal{L}_X\tau)|_U)^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,v_q} = X^\alpha\partial_\alpha\tau^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,v_q} - \sum_{i=1}^{p}\tau^{\mu_1,\cdots,\alpha,\cdots,\mu_p}_{v_1,\cdots,v_q}\partial_\alpha X^{\mu_i} + \sum_{j=1}^{q}\tau^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,\alpha,\cdots,v_q}\partial_{v_j}X^\alpha$$

8.5 Exterior differential

Definition 8.14 Exterior form space

$A(M) = \sum_{r=0}^{m} A^r(M)$
For $\tau \in A^r(M)$,

$$\tau|_U = \frac{1}{r!}\tau_{i_1\cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r} = \tau_{|i_1\cdots i_r|} dx^{i_1} \wedge \cdots \wedge dx^{i_r}$$

$$\tau_{i_1\cdots i_r} = \tau(\frac{\partial}{\partial x^{i_1}}, \cdots, \frac{\partial}{\partial x^{i_r}})$$

$$\begin{aligned} \tau(v_1, \cdots, v_r)|_U &= \tau_{|i_1\cdots i_r|} dx^{i_1} \wedge \cdots \wedge dx^{i_r}(v_1, \cdots, v_r) \\ &= \tau_{|i_1\cdots i_r|} \begin{vmatrix} v_1^{i_1} & \cdots & v_r^{i_1} \\ \vdots & & \vdots \\ v_1^{i_r} & \cdots & v_r^{i_r} \end{vmatrix} \end{aligned}$$

It is a r *multi-linear mapping, and for every variable, it is* $C^\infty(M)$ *linear.*

Proposition 8.9 Pullback mapping

$f : M \to N \Rightarrow f_* : T_pM \to T_{f(p)}N \Rightarrow f^* : \wedge^r(T^*_{f(p)}N) \to \wedge^r(T^*_pM)$

$$f^*\phi(v_1, \cdots, v_r) = \phi(f_*v_1, \cdots, f_*v_r)$$
$$f^*\phi|_U = \frac{1}{r!}(\phi_{\alpha_1\cdots\alpha_r} \circ f) \cdot \frac{\partial f^{\alpha_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{\alpha_r}}{\partial x^{i_r}} dx^{i_1} \wedge \cdots \wedge dx^{i_r}$$
$$f^*(\phi \wedge \psi) = f^*\phi \wedge f^*\psi$$

Definition 8.15 Exterior differential

Let M *be a m-dimensional smooth manifold. Then* $\exists$ *a unique mapping* $d : A(M) \to A(M)$ *satisfy that*
(1) $d(A^r(M)) \subset A^{r+1}(M)$
(2) $\forall \omega_1, \omega_2 \in A(M), d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$
(3) if $\omega_1 \in A^r(M)$*,then* $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^r \omega_1 \wedge d\omega_2$
(4) $f \in A^0(M)$*,*df *is just the differential of* f
(5) $\forall f \in A^0(M)$*,*$d(df) = 0$
d *is called exterior differential.*

Theorem 8.7

$\forall \omega \in A^1(M), X, Y \in \mathcal{H}(M)$,

$$d\omega(X,Y) = X\langle Y,\omega\rangle - Y\langle X,\omega\rangle - \langle [X,Y],\omega\rangle$$

$\forall \omega \in A^r(M), X_1, \cdots, X_{r+1} \in \mathcal{H}(M)$,

$$\begin{aligned} d\omega(X_1,\cdots,X_{r+1}) &= \sum_{i=1}^{r+1}(-1)^{i+1}X_i(\langle X_1\wedge\cdots\wedge\hat{X}_i\wedge\cdots\wedge X_{r+1},\omega\rangle) \\ &+ \sum_{1\le i<j\le r+1}(-1)^{i+j}\langle [X_i,X_j]\wedge\cdots\wedge\hat{X}_i\wedge\cdots\wedge\hat{X}_j\wedge\cdots X_{r+1},\omega\rangle \end{aligned}$$

Theorem 8.8

$$f^*(d\omega) = d(f^*\omega)$$

Lemma 1 Poincare Lemma

1. $d^2 = 0$
2. Let $U = B_0(r)$ be a spherical neighbourhood with center origin O and radius r in R^n. $\forall \omega \in A^r(U)$ and $d\omega = 0$, $\exists \tau \in A^{r-1}(U)$,satisfy that $\omega = d\tau$.

Definition 8.16 Pfaff euqations

Let $\omega^\alpha (1 \le \alpha \le r) \in A^1(U)$ and U is an open set of m-dimensional smooth manifold M. Differential equation set $\omega^\alpha = 0$ is called Pfaff equations.

Definition 8.17 Integral manifold of Pfaff equations

If there is an injective immersion submanifold $\phi : N \to U$ satisfying that $\phi^*\omega^\alpha = 0$,(ϕ, N) is called an integral manifold of Pfaff eqation set.

Proposition 8.10 Partial differential equations and Pfaff equations

There is a set of first order partial differential equations

$$\frac{\partial y^\alpha}{\partial x^i} = f_i^\alpha(x^1, \cdots, x^m, y^1, \cdots, y^n) \quad (1 \le i \le m, 1 \le \alpha \le n)$$

$f_i^\alpha(x, y)$ is a smooth function on the open set $U \times V \subset R^m \times R^n$. The equations sets can be written as Pfaff equations on $U \times V$

$$\omega^\alpha \equiv dy^\alpha - f_i^\alpha(x, y)dx^i = 0$$

If the partial differential equations have solution

$$y^\alpha = g^\alpha(x^1, \ldots, x^m)$$

then the submanifold $\phi : U \to U \times V$,

$$\phi(x^1, \ldots, x^m) = (x^1, \ldots, x^m, g^1(x), \ldots, g^n(x))$$

is an integral manifold of the Pfaff equations , i.e. $\phi^*\omega^\alpha = 0$

Proposition 8.11 Distribution and Pfaff equations

Pfaff equations $\omega^\alpha = 0$ on open set $V \in M$ with rank r is equivalent to a $h = m - r$ dimensional smooth distribution locally.

$$\Delta^h(p) = \{v \in T_pM : \omega^\alpha(v) = 0, 1 \le \alpha \le r\}$$

If $\phi : N \to V$ is an integral manifold of ω^α,$\forall X \in T_pN$,$\omega^\alpha(\phi_* X) = \phi^*\omega_\alpha(X) = 0$. So $\phi_* X \in \Delta^h(p)$, and so $\phi : N \to V$ is an integral manifold of Δ^h.

Definition 8.18 Completely integrable

Suppose ω^α is a set of r linearly independent 1 forms defined on an open set $U \subset M$. If $\forall p \in U$, Paffa equations

$$\omega^\alpha = 0 \quad (1 \le \alpha \le r)$$

has an $h = \dim M - r$ dimensional integral manifold $\phi : N \to V$ such that $p \in V$, Paffa equations are called completely integrable.

Definition 8.19 Frobenius condition

Frobenius condition for Pfaff equations $\omega^\alpha = 0 (1 \leq \alpha \leq r)$ *is that*

$$d\omega^\alpha \wedge \omega^1 \wedge \cdots \wedge \omega^r = 0$$

Theorem 8.9 Frobenius theorem

Pfaff equations satisfying Frobenius condition is completely integrable.

Definition 8.20 Orientation of manifold

Let $\alpha : [0,1] \to M$ *be a path on* M. $\forall t \in [0,1]$, *assign an orientation for* $T_{\alpha(t)}M$, *denoted by* μ_t. *If for* $t_0 \in [0,1]$, *there is a local coordinate* $(U; x_i)$ *of* $\alpha(t_0)$ *and a neighbourhood* $[t_0 - \delta_1, t_0 + \delta_2]$ *of* t_0 *that*

$$\alpha([t_0 - \delta_1, t_0 + \delta_2]) \subset U$$

and

$$\left\{ \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m} \right\} |_{\alpha(t)} \in \mu_t, \forall t \in [t_0 - \delta_1, t_0 + \delta_2],$$

μ *is called a continuous topological orientation of* α

Definition 8.21 The propagation of orientation

Let $p, q \in M$ *and* $\alpha : [0,1] \to M$ *a path connecting* p, q. *Assign an orientation* λ *of* T_pM. *If there is a continuous topological orientation of* α μ *satisfying that* $\mu_0 = \lambda$, *then orientation* μ_1 *of* T_qM *is called the propagation of orientation* λ *along* α. *The orientation of* μ_1 *is unique.*

Definition 8.22 Orientable manifold

Let M *be a* m *dimensional smooth manifold. If there is an atlases* $(\mathcal{A}, = \{(U_\alpha, \phi_\alpha)\})$, *making that if* $U_\alpha \cap U_\beta \neq \emptyset$,*the Jacobian of*

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

is positive. Then M *is called orientable manifold.*

Theorem 8.10

Let M be a orientable connected manifold. $\forall p \in M$,assign an orientation λ for T_pM, then for all point $q \in M$, the propagation of λ along an arbitrary path define a unique orientation μ for T_qM.

Definition 8.23 Manifold with boundary

A topological manifold with boundary is a Hausdorff space in which every point has a neighbourhood homeomorphic to an open subset of Euclidean half-space (for a fixed n):

$$\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n \geq 0\}$$

Definition 8.24 Boundary and interior

Let M be a manifold with boundary. The interior of M, denoted $Int\ M$, is the set of points in M which have neighbourhoods homeomorphic to an open subset of $\mathbb{R}^n$. The boundary of M, denoted ∂M, is the complement of $Int\ M$ in M. The boundary points can be characterized as those points which land on the boundary hyperplane ($x^n = 0$) of $\mathbb{R}^n_+$ under some coordinate chart.
If M is a manifold with boundary of dimension n, then $Int\ M$ is a manifold (without boundary) of dimension n and ∂M is a manifold (without boundary) of dimension $n-1$.

Theorem 8.11

Let M be a smooth manifold with boundary and $\partial M \neq \emptyset$. The differential structure of ∂M can be deduced from the M, making ∂M a $m-1$ dimensional smooth manifold and the inclusion map $i : \partial M \to M$ is embedding map. If M is orientable, then ∂M is also orientable.

Definition 8.25 Induced orientation

Let M be an orientable m dimensional smooth manifold with boundary and $\partial M \neq \emptyset$. $\mathcal{A}$ is the orientation of M. For local coordinates $(U; x^i) \in \mathcal{A}$, when

$$\tilde{U} = U \cap \partial M = \{(x^1, \ldots, x^m) \in U : x^m = 0\} \neq \emptyset$$

assign a local coordinate system $((-1)^m \cdot x^1, x^2, \ldots, x^{m-1})$ on $\tilde{U}$. The orientation defined by this local coordinate system is called induced orientation of ∂M.

Definition 8.26 Support set

Let M be a m dimensional orientable smooth manifold. $\omega \in A^r(M)$, the support set of ω can be defined as

$$Supp\,\omega = \overline{\{p \in M : \omega(p) \neq 0\}}$$

All the r-form with compact support set is denoted as $A_0^r(M)$.

Definition 8.27 Partition of unity

Let Σ be an open cover of M. Then there is a family of smooth function g_α on M that

1. $\forall\alpha, 0 \leq g_\alpha \leq 1$, $\operatorname{supp} g_\alpha$ *is compact and there is an open set* $W_i \in \Sigma$ *that* $\operatorname{supp} g_\alpha \subset W_i$.
2. $\forall p \in M$, *it has a neighbourhood U which intersect finite* $\operatorname{supp} g_\alpha$.
3. $\sum_\alpha g_\alpha = 1$

Definition 8.28 Integral of differential form with compact support

$$\phi = (\sum_\alpha g_\alpha) \cdot \phi = \sum_\alpha (g_\alpha \cdot \phi)$$

$$\int_M g_\alpha \cdot \phi = \int_{W_i} g_\alpha \cdot \phi = \int_{W_i} f(u^1, \cdots, u^m) du^1 \wedge \cdots \wedge du^m = \int_{W_i} f(u^1, \cdots, u^m) du^1 \cdots du^m$$

$$\int_M \phi = \sum_\alpha \int_M g_\alpha \cdot \phi$$

Theorem 8.12 Stokes Theorem

Let M be an orientable m dimensional smooth manifold with boundary and $\omega \in A_0^{m-1}(M)$,then

$$\int_M d\omega = \int_{\partial M} i^*\omega$$

Here, ∂M has an orientation induced by M and i is embedding mapping.

8.6 Connection

Definition 8.29 Connection

Let M be a smooth manifold and E a q dimensional real vector bundle on M. $\Gamma(E)$ is the set of all smooth sections of E on M. The connection on E is a mapping:

$$D : \Gamma(E) \to \Gamma(T^*(M) \otimes E)$$

it satisfies that

1. $\forall s_1, s_2 \in \Gamma(E), D(s_1 + s_2) = Ds_1 + Ds_2$
2. $\forall s \in \Gamma(E)$ *and* $\alpha \in C^\infty(M)$, $D(\alpha s) = d\alpha \otimes s + \alpha Ds$

If X is a smooth tangent vector field on M, $s \in \Gamma(E)$,then $D_X s = \langle X, Ds \rangle$, called absolute derivative of s along X.

Proposition 8.12

Local representation of connection is:

$$Ds_\alpha = \sum_{1 \le i \le m, 1 \le \beta \le q} \Gamma^\beta_{\alpha i} du^i \otimes s_\beta$$

$$\omega^\beta_\alpha = \sum_{1 \le i \le m} \Gamma^\beta_{\alpha i} du^i$$

$$Ds_\alpha = \sum_{\beta=1}^{q} \omega^\beta_\alpha \otimes s_\beta$$

$$DS = \omega \otimes S$$

Transformation law of connection is:

$$S' = A \cdot S$$

$$\begin{aligned} DS' &= dA \otimes S + A \cdot DS \\ &= (dA + A \cdot \omega) \otimes S \\ &= (dA \cdot A^{-1} + A \cdot \omega \cdot A^{-1}) \otimes S' \end{aligned}$$

$$\omega' = dA \cdot A^{-1} + A \cdot \omega \cdot A^{-1}$$

Theorem 8.13

For an arbitrary vector bundle, connection always exists.

Theorem 8.14

Let D be a connection of vector bundle E. $\forall p \in M$, there exists a local frame field S on the neighbourhood of p that $\omega(p) = 0$.

Definition 8.30 Curvature matrix

$$\Omega \equiv d\omega - \omega \wedge \omega$$

Proposition 8.13 Transformation law of curvature matrix

$$\Omega' = A \cdot \Omega \cdot A^{-1}$$

Definition 8.31 Curvature operator

$$s = \sum_{\alpha=1}^{q} \lambda^\alpha s_{\alpha|p}$$

$$R(X,Y)s = \sum_{\alpha,\beta=1}^{q} \lambda^\alpha \Omega_\alpha^\beta(X,Y) s_{\beta|p}$$

$$(R(X,Y)s)(p) = R(X_p, Y_p)s_p$$

Proposition 8.14

$$R(X,Y) = D_X D_Y - D_Y D_X - D_{[X,Y]}$$

Theorem 8.15 Bianchi equation

$$d\Omega = \omega \wedge \Omega - \Omega \wedge \omega$$

Definition 8.32 Induced connection

$$d\langle s, s^*\rangle = \langle Ds, s^*\rangle + \langle s, Ds^*\rangle$$

$$\langle s_\alpha, s^{*\beta}\rangle = \delta^\beta_\alpha \Rightarrow Ds^{*\beta} = -\sum_{\alpha=1}^{q} \omega^\beta_\alpha \otimes s^{*\alpha}$$

$$D(s_1 \oplus s_2) \equiv Ds_1 \oplus Ds_2$$

$$D(s_1 \otimes s_2) \equiv Ds_1 \otimes s_2 + s_1 \otimes Ds_2$$

Definition 8.33 Affine connection

$$D\frac{\partial}{\partial u^i} \equiv \omega^j_i \otimes \frac{\partial}{\partial u^j} \equiv \Gamma^j_{ik} du^k \otimes \frac{\partial}{\partial u^j}$$

Proposition 8.15

$$\Gamma'^j_{ik} = \Gamma^q_{pr} \frac{\partial w^j}{\partial u^q}\frac{\partial u^p}{\partial w^i}\frac{\partial u^r}{\partial w^k} + \frac{\partial^2 u^p}{\partial w^i \partial w^k}\frac{\partial w^j}{\partial u^p}$$

$$DX = (dX^i + x^j\omega^i_j) \otimes \frac{\partial}{\partial u^i} = (X^i_{,j} + X^k\Gamma^i_{kj})du^j \otimes \frac{\partial}{\partial u^i} = X^i_{;j} du^j \otimes \frac{\partial}{\partial u^i}$$

$$D\alpha = (d\alpha_i - \alpha_j\omega^j_i) \otimes du^i = (\alpha_{i,j} - \alpha_k\Gamma^k_{ij})du^j \otimes du^i = \alpha_{i;j} du^j \otimes du^i$$

Definition 8.34 Geodesic equation

$$\frac{D(\frac{du^i(t)}{dt}\frac{\partial}{\partial u^i})}{dt} = 0$$

$$\frac{d^2u^i}{dt^2} + \Gamma^i_{jk}\frac{du^j}{dt}\frac{du^k}{dt} = 0$$

Definition 8.35 Curvature tensor

$$\Omega^j_i = \frac{1}{2}R^j_{ikl} du^k \wedge du^l$$

$$R \equiv R^j_{ikl} du^i \otimes \frac{\partial}{\partial u^j} \otimes du^k \otimes du^l$$

Proposition 8.16

$$R^j_{ikl} = \frac{\partial \Gamma^j_{il}}{\partial u^k} - \frac{\partial \Gamma^j_{ik}}{\partial u^l} + \Gamma^h_{il}\Gamma^j_{hk} - \Gamma^h_{ik}\Gamma^j_{hl}$$

$$R'^j_{ikl} = R^q_{prs}\frac{\partial w^j}{\partial u^q}\frac{\partial u^p}{\partial w^i}\frac{\partial u^r}{\partial w^k}\frac{\partial u^s}{\partial w^l}$$

$$R(X, \alpha_Y, Z, W) = \langle \alpha_Y, R(Z, W)X\rangle$$

$$R^j_{ikl} = \langle R(\frac{\partial}{\partial u^k}\frac{\partial}{\partial u^l})\frac{\partial}{\partial u^i}, du^j\rangle$$

Definition 8.36 Torsion tensor

$$T^j_{ik} = \Gamma^j_{ki} - \Gamma^j_{ik}$$

$$T = T^j_{ik}\frac{\partial}{\partial u^j} \otimes du^i \otimes du^k$$

Proposition 8.17

$$T(X, Y) = T^k_{ij}X^iY^j\frac{\partial}{\partial u^k}$$

$$T(X, Y) = D_XY - D_YX - [X, Y]$$

Theorem 8.16

Let D be an affine connection without torsion on M. $\forall p \in M$, there is a local coordinate system that $\Gamma^j_{ik}(p)$ vanishes.

Theorem 8.17

Let D be an affine connection without torsion on M. Then we have Bianchi equation

$$R^j_{ikl;h} + R^j_{ihk;l} + R^j_{ilh;k} = 0$$

8.7 Riemann manifold

Definition 8.37 Riemann manifold

Let M be a smooth manifold equipped with a smooth non-degenerate symmetric second order covariant tensor field G, then M is called general Riemann manifold and G is called the metric tensor of M.
If G is positive definite, then M is called Riemann manifold.

Theorem 8.18

There must be a Riemann metric on m dimensional manifold M.

Definition 8.38 Index lifting

$$f : T_p(M) \to T_p^*(M) \quad \alpha_X(Y) \equiv G(X, Y)$$

Definition 8.39 Adapted connection

Let (M, G) be a general Riemann manifold and D a connection on M. If $DG = 0$,then D is called adapted connection on M.

Proposition 8.18 Christoffel-Levi-Civita connection

Let M be a general Riemann manifold, then there is a unique adapted connection without torsion on M, called Christoffel-Levi-Civita connection.
As $\omega_i^j = \Gamma_{ik}^j du^k$, $dg_{ij} = g_{ik}\omega_j^k + g_{kj}\omega_i^k$. If we denote that $\omega_{ij} = \omega_i^j g_{jk}$ and $\Gamma_{ijk} = \Gamma_{ik}^l g_{lj}$, we have $\omega_{ij} = \Gamma_{ijk} du^k$ and $dg_{ij} = \omega_{ji} + \omega_{ij}$. At last, we have

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left(\frac{\partial g_{il}}{\partial u^j} + \frac{\partial g_{jl}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l} \right)$$

Proposition 8.19 Curvature tensor

If we denote that $\Omega_{ij} = \Omega_i^k g_{kj}$ and $R_{ijkl} = R_{ikl}^h g_{hj}$, we will have that

$$\Omega_{ij} + \Omega_{ji} = 0, \;\; \Omega_{ij} = d\omega_{ij} + \omega_i^l \wedge \omega_{jl}, \;\; \Omega_{ij} = \frac{1}{2} R_{ijkl} du^k \wedge du^l$$

The properties of curvature tensor:

$$R_{ijkl} = -R_{jikl} = -R_{ijlk}$$

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0$$

$$R_{ijkl} = R_{klij}$$

Definition 8.40 Normal coordinates

In differential geometry, normal coordinates at a point p in a differentiable manifold equipped with a symmetric affine connection are a local coordinate system in a neighbourhood of p obtained by applying the exponential map to the tangent space at p. In a normal coordinate system, the Christoffel symbols of the connection vanish at the point p, thus often simplifying local calculations. In normal coordinates associated to the Levi-Civita connection of a Riemann manifold, one can additionally arrange that the metric tensor is the Kronecker delta at the point p, and that the first partial derivatives of the metric at p vanish.

The properties of normal coordinates often simplify computations. In the following, assume that U is a normal neighbourhood centred at p in M and (x_i) are normal coordinates on U.

Let V be some vector from T_pM with components V^i in local coordinates, and γ_V be the geodesic with starting point p and velocity vector V, then γ_V is represented in normal coordinates by $\gamma_V(t) = (tV^1, ..., tV^n)$ as long as it is in U.

The coordinates of p are $(0, \cdots, 0)$

In Riemann normal coordinates at p the components of the Riemann metric g simplify to δ_{ij}.

The Christoffel symbols vanish at p. In the Riemann case, so do the first partial derivatives of g_{ij}.

Theorem 8.19

Let M be a differentiable manifold equipped with a symmetric affine connection. $\forall x_0 \in M$, there is a neighbourhood W that for every point in W, there is a neighbourhood equipped with a normal coordinate system which contains W.

Theorem 8.20

Let M be a Riemann manifold. $\forall O \in M$,there is a neighbourhood with normal coordinates W that:
(1) For every point in W, there is a neighbourhood equipped with a normal coordinates which contains W.
(2) The geodesic connecting O and $p \in W$ is the only shortest path connecting these two points in W.

Theorem 8.21

Let U be the neighbourhood with normal coordinates of O. $\exists \epsilon > 0, \forall \delta \in (0, \epsilon)$,the surface

$$\Sigma_\delta = p \in U | \sum_{i=1}^{m} (u^i(p))^2 = \delta^2$$

has following properties:
(1) $\forall p \in \Sigma_\delta$, there is a unique shortest geodesic connecting p and O in U.
(2) For all geodesics tangent to Σ_δ, there is a neighbourhood of the cut point in which the geodesics lies outside of Σ_δ

Theorem 8.22

Let M be a Riemann manifold and $\forall p \in M$, there is a η-spherical neighbourhood W that for arbitrary two points in W, there is a unique geodesic connecting these two points.

Definition 8.41 Cross section curvature

$$R(X, Y, Z, W) \equiv R_{ijkl} X^i Y^j Z^k W^l$$

$$R(X, Y, Z, W) = (R(Z, W)X) \cdot Y$$

$$G(X, Y, Z, W) \equiv G(X, Z)G(Y, W) - G(X, W)G(Y, Z)$$

Let E be a two dimensional subspace of $T_p(M)$ and X, Y two linearly independent tangent vector of E,then

$$K(E) = -\frac{R(X, Y, X, Y)}{G(X, Y, X, Y)}$$

is a function of E, which is independent of the choice of X, Y,called cross section curvature.

Theorem 8.23

Let M be a Riemann space. The curvature tensor of $p \in M$ is uniquely determined by the cross section curvature of all the two dimensional subspace of $T_p(M)$.

Definition 8.42 Constant curvature Riemann manifold

Let M be a Riemann manifold. If all of $K(E)$ on p is constant, then M is called isotropic on p.
If M is isotropic every where and $K(p)$ is constant over M, then M is called constant curvature Riemann manifold.

Theorem 8.24 F.Schur theorem

Let M be a m-dimensional connected Riemann manifold that is isotropic every where. If $m \geq 3$, then M is constant curvature Riemann manifold.

Chapter 9
A Geometrical Description of Newton Theory

9.1 Introduction

We choose Euclidean coordinates for our absolute space and an absolute time t, than the equation of motion can be written as

$$\frac{d^2t}{d\lambda^2} = 0$$

$$\frac{d^2x^i}{d\lambda^2} + \frac{\partial\Phi}{\partial x^i}(\frac{d\lambda}{dt})^2 = 0$$

It is convenient to define that $\Gamma^i_{00} = \frac{\partial\Phi}{\partial x^i}$,and all other $\Gamma^\alpha_{\beta\gamma}$ vanish.Then we can write the equation of motion as

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0$$

Next, we can get the Riemann tensor given the connection above

$$R^i_{0j0} = -R^i_{00j} = \frac{\partial\Phi}{\partial x^i\partial x^j},$$

and all other terms vanish. It is straight forward to derive the expression of Ricci tensor,

$$R_{00} = \Phi_{ii} = \nabla^2\Phi,$$

and all other terms vanish. So, newton gravity law can be written as

$$R_{00} = 4\pi\rho$$

9.2 Geometry structure of Newtonian Space-time

Stratification of space-time

Regard absolute time t as a scalar field defined once and for all in Newtonian space-time $t = t(\mathcal{P})$.The layers of space-time are the slices of constant t-the "space slices"-each of which has an identical geometric structure: the old "absolute space."

Flat Euclidean space

A given space slice is endowed with basis vectors $\boldsymbol{e}_i = \frac{\partial}{\partial x^i}$; and this basis has vanishing connection coefficients,$\Gamma^i_{jk} = 0$. Consequently,the geometry of each space slice is completely flat. Absolute space is Euclidean in its geometry.Each space slice is endowed with a three-dimensional metric, and its Galilean coordinate basis is orthonormal,$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$.

Curvature of space-time

Parallel transport a vector around a closed curve lying entirely in a space slice; it will return to its starting point unchanged. But transport it forward in time by Δt, northerly in space by Δx_k, back in time by $-\Delta t$, and southerly by $-\Delta x_k$ to its starting point; it will return changed by

$$\delta \boldsymbol{A} = -\mathcal{R}(\Delta t \frac{\partial}{\partial t}, \Delta x_k \frac{\partial}{\partial x_k})\boldsymbol{A}$$

Geodesics of a space slice (Euclidean straight lines) that are initially parallel remain always parallel. But geodesics of space-time (trajectories of freely falling particles) initially parallel get pried apart or pushed together by space-time curvature,

$$\nabla_{\boldsymbol{u}} \nabla_{\boldsymbol{u}} \boldsymbol{n} + \mathcal{R}(\boldsymbol{n}, \boldsymbol{u})\boldsymbol{u} = 0$$

9.3 Geometry formulation of Newtonian gravity

1. There exists a function t called "universal time", and a symmetric covariant derivative ∇.

2. The 1-form $\boldsymbol{d}t$ is covariant constant, i.e.,

$$\nabla_{\boldsymbol{u}} \boldsymbol{d}t = 0 \text{ for all } \boldsymbol{u}.$$

Note: *if $\boldsymbol{w}$ is a spatial vector field, then $\nabla_{\boldsymbol{u}} \boldsymbol{w}$ is also spatial for every $\boldsymbol{u}$.*

3. Spatial vectors are unchanged by parallel transport around infinitesimal closed curves; i.e.,

$$\mathcal{R}(\boldsymbol{n}, \boldsymbol{u})\boldsymbol{w} = 0 \text{ if } \boldsymbol{w} \text{ is spatial, for every } \boldsymbol{u} \text{ and } \boldsymbol{n}.$$

4. All vectors are unchanged by parallel transport around infinitesimal, spatial, closed curves; i.e.,

$$\mathcal{R}(\boldsymbol{v}, \boldsymbol{w}) = 0 \text{ for every spatial } \boldsymbol{v} \text{ and } \boldsymbol{w}.$$

5. The Ricci curvature tensor has the form

$$\boldsymbol{Ricci} = 4\pi\rho \boldsymbol{d}t \otimes \boldsymbol{d}t$$

where ρ is the density of mass.

6. There exists a metric $\cdot$ defined on spatial vectors only, which is compatible with the covariant derivative in this sense: for any spatial $\boldsymbol{w}$ and $\boldsymbol{v}$, and for any $\boldsymbol{u}$ whatsoever,

$$\nabla_{\boldsymbol{u}}(\boldsymbol{w} \cdot \boldsymbol{v}) = (\nabla_{\boldsymbol{u}} \boldsymbol{w}) \cdot \boldsymbol{v} + \boldsymbol{w} \cdot (\nabla_{\boldsymbol{u}} \boldsymbol{v}).$$

Note: *Axioms (1), (2), and (3) guarantee that such a spatial metric can exist.*

7. The Jacobi curvature operator $\mathcal{J}(\boldsymbol{u},\boldsymbol{v})$, defined for any vectors $\boldsymbol{u},\boldsymbol{n},\boldsymbol{p}$ by

$$\mathcal{J}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{p} = \frac{1}{2}[\mathcal{R}(\boldsymbol{p},\boldsymbol{n})\boldsymbol{u} + \mathcal{R}(\boldsymbol{p},\boldsymbol{u})\boldsymbol{n}]$$

is "self-ad-joint" when operating on spatial vectors,i.e.,

$$\boldsymbol{v}\cdot[\mathcal{R}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{w}] = \boldsymbol{w}\cdot[\mathcal{R}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{v}] \text{ for all spactial } \boldsymbol{v},\boldsymbol{w};\ \text{ and for any } \boldsymbol{u},\boldsymbol{n}.$$

8. "Ideal rods" measure the lengths that are calculated with the spatial metric; "ideal clocks" measure universal time t (or some multiple thereof); and "freely falling particles" move along geodesics of ∇.

9.4 Standard formulation of Newtonian gravity

1. There exist a universal time t, a set of Cartesian space coordinates x_i (called "Galilean coordinates"), and a Newtonian gravitational potential Φ.

2. The density of mass ρ generates the Newtonian potential by Poisson's equation,

$$\nabla^2\Phi = \frac{\partial\Phi}{\partial x^i\partial x^i} = 4\pi\rho.$$

3. The equation of motion for a freely falling particle is

$$\frac{d^2x^i}{dt^2} + \frac{\partial\Phi}{\partial x^i} = 0.$$

4. "Ideal rods" measure the Galilean coordinate lengths; "ideal clocks" measure universal time.

9.5 Galilean coordinate system

The features of Galilean coordinate systems are

$$x^0(\mathcal{P}) = t(\mathcal{P})$$

$$\frac{\partial}{\partial x^i}\cdot\frac{\partial}{\partial x^j} = \delta_{ij}$$

$$\Gamma^j_{00} = \Phi_{,j} \text{ for some scalar field } \Phi,\ \text{ and all other } \Gamma^\alpha_{\beta\gamma}\text{vanish.}$$

Consider following coordinate transformation:
(1)$x^{0'} = x^0 = t$,both time coordinates must be universal time.
(2)at fixed t,both sets of space coordinates must be Euclidean, so they must be related by a rotation and a translation:

$$\bar{x}^{i'}(t) = A_{i'j}(t)x^j + \bar{a}^{i'}(t)$$

We can get

$$\bar{\Gamma}^{i'}_{0j'} = \bar{\Gamma}^{i'}_{j'0} = A_{i'l}\dot{A}_{j'l}$$

$$\bar{\Gamma}^{i'}_{00} = \Phi_{,i'} + A_{i'j}(\ddot{A}_{l'j}\bar{x}^{l'} - \ddot{a}^j), \text{ here, } a^j = \bar{a}^{l'}A_{l'j}$$

and all other terms vanish.So, new coordinates have the standard Galilean form if and only if

$$\dot{A}_{i'j} = 0, \;\; \Phi' = \Phi - \ddot{\bar{a}}^{i'}x^{i'} + C$$

Were all the matter in the universe concentrated in a finite region of space and surrounded by emptiness ("island universe"), then one could impose the global boundary condition $\Phi \to 0$ as $r \equiv (x^i x^i)^{\frac{1}{2}} \to \infty$. This would single out a subclass of Galilean coordinates ("absolute" Galilean coordinates), with a unique, common Newtonian potential. The transformation from one absolute Galilean coordinate system to any other is called Galilean transformation.

9.6 Coordinate transformation in space

We now consider a coordinate transformation of Galilean coordinate system purely in space without any terms related with time. That means that $\bar{x}^{i'} = y^{i'}(x^i)$ and $t' = t$. We can calculate the connection term in the new coordinate system.

$$\bar{\Gamma}^{i'}_{00} = \Gamma^i_{00}\frac{\partial y^{i'}}{\partial x^i}$$

$$\bar{\Gamma}^{i'}_{j'k'} = \Gamma^i_{jk}\frac{\partial^2 x^m}{\partial y^{i'}\partial y^{k'}}\frac{\partial y^{i'}}{\partial x^m}$$

The equation of motion of free fall body is that

$$\frac{d^2t'}{d\lambda^2} = 0$$

$$\frac{d^2\bar{x}^{i'}}{d\lambda^2} + \bar{\Gamma}^{i'}_{j'k'}\frac{d\bar{x}^{j'}}{d\lambda}\frac{d\bar{x}^{k'}}{d\lambda} + \bar{\Gamma}^{i'}_{00}\frac{dt'}{d\lambda}\frac{dt'}{d\lambda} = 0$$

We can write it compactly as

$$\frac{d^2\bar{x}^{i'}}{dt^2} + \bar{\Gamma}^{i'}_{j'k'}\frac{d\bar{x}^{j'}}{dt}\frac{d\bar{x}^{k'}}{dt} + \bar{\Gamma}^{i'}_{00} = 0$$

We can demonstrate that

$$\bar{\Gamma}^{i'}_{j'k'} = \frac{1}{2}\bar{g}^{i'p'}(\partial_{k'}\bar{g}_{j'p'} + \partial_{j'}\bar{g}_{k'p'} - \partial_{p'}\bar{g}_{j'k'})$$

and

$$\bar{\Gamma}^{i'}_{00} = \bar{g}^{i'j'}\partial_{j'}\Phi$$

Here,$\bar{g}$ is the metric of the space in new coordinate system.

Chapter 10
Geometry of Space-time

10.1 Hodge dual operator

10.1.1 Hodge dual

Definition 10.1 Hodge dual operator

The Hodge star operator on a vector space V with a non-degenerate symmetric bilinear form (herein referred to as the inner product) is a linear operator on the exterior algebra of V, mapping k-vectors to $(n-k)$-vectors where $n = \dim V$, for $0 \leq k \leq n$. It has the following property, which defines it completely: given two k-vectors α, β,

$$\alpha \wedge (\star\beta) = \langle \alpha, \beta \rangle \omega$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on k-vectors and ω is the preferred unit n-vector. The inner product $\langle \cdot, \cdot \rangle$ on k-vectors is extended from that on V by requiring that

$$\langle \alpha, \beta \rangle = \det \left[\langle \alpha_i, \beta_j \rangle \right]$$

for any decomposable k-vectors $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_k$ and $\beta = \beta_1 \wedge \cdots \wedge \beta_k$. The unit n-vector ω is unique up to a sign. The preferred choice of ω defines an orientation on V.

Given an orthonormal basis $(e_1, \cdots, e_n)$ ordered such that $\omega = e_1 \wedge \cdots \wedge e_n$, we see that

$$\star(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}) = e_{i_{k+1}} \wedge e_{i_{k+2}} \wedge \cdots \wedge e_{i_n}$$

where $(i_1, i_2, \cdots, i_n)$ is an even permutation of $\{1, 2, \cdots, n\}$.Of these $\frac{n!}{2}$, only $\binom{n}{k}$ are independent. The first one in the usual lexicographical order reads

$$\star(e_1 \wedge e_2 \wedge \cdots \wedge e_k) = e_{k+1} \wedge e_{k+2} \wedge \cdots \wedge e_n$$

10.1.2 Levi-Civita tensor

Definition 10.2 Levi-Civita tensor

$$\epsilon_{i_1,\cdots,i_n} \equiv |g|^{\frac{1}{2}} \tilde{\epsilon}_{i_1,\cdots,i_n},$$

where ϵ is Levi-Civita symbol.

Proposition 10.1

$$\epsilon^{i_1,\cdots,i_n} = g^{i_1 j_1} \cdots g^{i_n j_n} \epsilon_{j_1,\cdots,j_n} = \frac{|g|^{\frac{1}{2}}}{g} \tilde{\epsilon}^{i_1,\cdots,i_n} = sgn(g) \frac{1}{|g|^{\frac{1}{2}}} \tilde{\epsilon}^{i_1,\cdots,i_n}$$

Using tensor index notation, the Hodge dual is obtained by contracting the indices of a k-form with the n-dimensional completely antisymmetric Levi-Civita tensor.

Proposition 10.2

$$(\star\eta)_{i_1,i_2,\ldots,i_{n-k}} = \frac{1}{(n-k)!} \eta^{j_1,\ldots,j_k} \epsilon_{j_1,\ldots,j_k,i_1,\ldots,i_{n-k}},$$

where η is an arbitrary antisymmetric tensor in k indices.

10.2 Metric-induced properties of Riemann curvature tensor

1. In a n dimensional manifold with torsion-free affine connection, the number of independent components of Riemann tensor is

$$\frac{n^3(n-1)}{2} - \frac{n^2(n-1)(n-2)}{6} = \frac{(n^2-1)n^2}{3}$$

   In a n dimensional Riemann manifold, the number of independent components of Riemann tensor is

$$\left[\frac{n(n-1)}{2}\right]^2 - \frac{n^2(n-1)(n-2)}{6} = \frac{(n^2-1)n^2}{12}$$

2. The double dual of Riemann tensor

$$\bar{G}^{\alpha\beta}{}_{\gamma\delta} = \frac{1}{2}\tilde{\epsilon}^{\alpha\beta\mu\nu} R_{\mu\nu}{}^{\rho\sigma} \frac{1}{2}\tilde{\epsilon}_{\rho\sigma\gamma\delta} = -\frac{1}{4}\delta^{\alpha\beta\mu\nu}_{\rho\sigma\gamma\delta} R_{\mu\nu}{}^{\rho\sigma}$$

   contains precisely the same amount of information as Riemann tensor, and satisfies precisely the same set of symmetries.

3. The Einstein curvature tensor, which is symmetric

$$G^{\beta}{}_{\delta} = \bar{G}^{\mu\beta}{}_{\mu\delta}; \quad G_{\beta\delta} = G_{\delta\beta}$$

4. The Bianchi identity takes a particularly simple form when rewritten in Bianchi identities terms of the double dual $\bar{G}$:

$$\bar{G}^{\alpha\beta}{}_{\gamma\delta}{}^{;\delta} = 0$$

and it has the obvious consequence

$$G_{\beta\delta}{}^{;\delta} = 0$$

5. The Ricci curvature tensor, which is symmetric, and the curvature scalar

$$R^{\beta}{}_{\delta} = R^{\mu\beta}{}_{\mu\delta}; \quad R_{\beta\delta} = R_{\delta\beta}; \quad R = R^{\beta}{}_{\beta}$$

which are related to the Einstein tensor by

$$G^{\beta}{}_{\delta} = R^{\beta}{}_{\delta} - \frac{1}{2}\delta^{\beta}_{\delta} R$$

6. The Weyl conformal tensor

$$C^{\alpha\beta}{}_{\gamma\delta} = R^{\alpha\beta}{}_{\gamma\delta} - 2\delta^{[\alpha}{}_{[\gamma} R^{\beta]}{}_{\delta]} + \frac{1}{3}\delta^{[\alpha}{}_{[\gamma}\delta^{\beta]}{}_{\delta]} R$$

possesses the same symmetries as the Riemann tensor. Weyl tensor is completely "trace-free"; i.e., that contraction of $C_{\alpha\beta\gamma\delta}$ on any pair of slots vanishes. Thus, $C_{\alpha\beta\gamma\delta}$ can be regarded as the trace-free part of Riemann, and $R_{\alpha\beta}$ can be regarded as the trace of Riemann. Riemann is determined entirely by its trace-free part $C_{\alpha\beta\gamma\delta}$ and its trace $R_{\alpha\beta}$

10.3 The coordinates of observer

10.3.1 Riemann normal coordinates

$$g_{\alpha\beta}(\mathcal{P}_0) = \eta_{\alpha\beta}$$

$$g_{\alpha\beta,\gamma}(\mathcal{P}_0) = 0$$

$$\Gamma^{\alpha}{}_{\beta\gamma}(\mathcal{P}_0) = 0$$

$$\Gamma^{\alpha}{}_{\beta\gamma,\mu}(\mathcal{P}_0) = -\frac{1}{3}(R^{\alpha}{}_{\beta\gamma\mu} + R^{\alpha}{}_{\gamma\beta\mu})$$

$$g_{\alpha\beta,\mu\nu}(\mathcal{P}_0) = -\frac{1}{3}(R_{\alpha\mu\beta\nu} + R_{\alpha\nu\beta\mu})$$

$$R_{\alpha\beta\gamma\delta}(\mathcal{P}_0) = g_{\alpha\delta,\beta\gamma} - g_{\alpha\gamma,\beta\mu}$$

10.3.2 The proper reference frame of an accelerated observer

1. Let τ be proper time as measured by the accelerated observer's clock .Let $\mathcal{P} = \mathcal{P}_0(\tau)$ be the observer's world line.

2. The observer carries with himself an orthonormal tetrad $\{\boldsymbol{e}_{\hat{\alpha}}\}$ with

$$\boldsymbol{e}_{\hat{0}} = \boldsymbol{u} = \frac{d\mathcal{P}_0}{d\tau}$$

and with

$$\boldsymbol{e}_{\hat{\alpha}} \cdot \boldsymbol{e}_{\hat{\beta}} = \eta_{\alpha\beta}$$

3. The tetrad changes from point to point along the observer's world line, relative to parallel transport:

$$\nabla_{\boldsymbol{u}} \boldsymbol{e}_{\hat{\alpha}} = -\boldsymbol{\Omega} \cdot \boldsymbol{e}_{\hat{\alpha}}$$

$$\Omega^{\mu\nu} = a^\mu u^\nu - u^\mu a^\nu + u_\alpha \omega_\beta \epsilon^{\alpha\beta\mu\nu}$$

This transport law has the same form in curved space-time as in flat because curvature can only be felt over finite distances, not over the infinitesimal distance involved in the "first time-rate of change of a vector" (equivalence principle).

$$\boldsymbol{a} = \nabla_{\boldsymbol{u}} \boldsymbol{u}$$

$$\boldsymbol{u} \cdot \boldsymbol{a} = \boldsymbol{u} \cdot \boldsymbol{\omega} = 0$$

If ω were zero, the observer would be Fermi-Walker-transporting his tetrad (gyroscope-type transport). If both $\boldsymbol{a}$ and $\boldsymbol{\omega}$ were zero, he would be freely falling (geodesic motion) and would be parallel-transporting his tetrad.

4. The observer constructs his proper reference frame (local coordinate system) in a manner analogous to the Riemann-normal construction. From each event $\mathcal{P}_0(\tau)$ on his world line, he sends out purely spatial geodesics (geodesics orthogonal to $\boldsymbol{u}$), with affine parameter equal to proper length. The tangent vector has unit length, because the chosen affine parameter is proper length.

5. Each event near the observer's world line is intersected by precisely one of the geodesics $\mathcal{G}[\tau, \boldsymbol{n}, s]$. [Far away, this is not true; the geodesics may cross, either because of the observer's acceleration].

6. Pick an event $\mathcal{P}$ near the observer's world line. The geodesic through it originated on the observer's world line at a specific time τ, had original direction $\boldsymbol{n} = n^{\hat{j}} \boldsymbol{e}_{\hat{j}}$; and needed to extend a distance s before reaching $\mathcal{P}$. Hence, the four numbers

$$(x^{\hat{0}}, x^{\hat{1}}, x^{\hat{2}}, x^{\hat{3}}) \equiv (\tau, sn^{\hat{1}}, sn^{\hat{2}}, sn^{\hat{3}})$$

are a natural way of identifying the event $\mathcal{P}$. These are the coordinates of $\mathcal{P}$ in the observer's proper reference frame.

Along the world line of observer:

$$\frac{\partial}{\partial x^{\hat{\alpha}}} = \boldsymbol{e}_{\hat{\alpha}}, \;\; g_{\hat{\alpha}\hat{\beta}} = \boldsymbol{e}_{\hat{\alpha}} \cdot \boldsymbol{e}_{\hat{\beta}} = \eta_{\hat{\alpha}\hat{\beta}}$$

$$\Gamma^{\hat{\beta}}_{\hat{\alpha}\hat{0}} = -\Omega^{\hat{\beta}}_{\hat{\alpha}}, \;\; \Gamma^{\hat{0}}_{\hat{0}\hat{0}} = 0$$

$$\Gamma^{\hat{0}}_{\hat{j}\hat{0}} = \Gamma^{\hat{j}}_{\hat{0}\hat{0}} = a^{\hat{j}}, \;\; \Gamma^{\hat{j}}_{\hat{k}\hat{0}} = -\omega^{\hat{i}}\epsilon_{0\hat{i}\hat{j}\hat{k}}$$

$$\Gamma^{\hat{\alpha}}_{\hat{j}\hat{k}} = 0$$

$$g_{\hat{\alpha}\hat{\beta},\hat{0}} = 0, \;\; g_{\hat{j}\hat{k},\hat{l}} = 0$$

$$g_{\hat{0}\hat{0},\hat{j}} = -2a^{\hat{j}}, \;\; g_{\hat{0}\hat{j},\hat{k}} = -\epsilon_{0\hat{j}\hat{k}\hat{l}}\omega^{\hat{l}}$$

10.4 Hypersurfaces

10.4.1 Description of hypersurfaces

Note: *We only discuss timelike and spacelile hypersurfaces in this section.*

Normal vector

$$n^\alpha n_\alpha = \epsilon \equiv \begin{cases} -1 \text{ if } \Sigma \text{ is spacelike} \\ +1 \text{ if } \Sigma \text{ is timelike} \end{cases}$$

Induced metric

Suppose that the hypersurface is parametrized with equation $x^\alpha = x^\alpha(y^a)$, Then

$$e^\alpha_a = \frac{\partial x^\alpha}{\partial y^a}.$$

For displacements within Σ, we have

$$\begin{aligned} ds^2_\Sigma &= g_{\alpha\beta}dx^\alpha dx^\beta \\ &= g_{\alpha\beta}(\frac{\partial x^\alpha}{\partial y^a}dy^a)(\frac{\partial x^\beta}{\partial y^b}dy^b) \\ &= h_{ab}dy^a dy^b \end{aligned}$$

where $h_{ab} = g_{\alpha\beta}e^\alpha_a e^\beta_b$. The completeness relation can be written as

$$g^{\alpha\beta} = \epsilon n^\alpha n^\beta + h^{ab}e^\alpha_a e^\beta_b$$

10.4.2 Integration on hypersurfaces

The positive volume element of the whole space time is $dx^0 \wedge \cdots \wedge dx^{m-1}$, the positive volume element of the hypersurfaces is $dy^1 \wedge \cdots \wedge dy^{m-1}$. Suppose that the coordinate in hypersurfaces is compatible with the coordinate of the whole space-time, which means that $-dy^m \wedge dy^1 \wedge \cdots \wedge dy^{m-1}$ has the same orientation as $dx^0 \wedge \cdots \wedge dx^{m-1}$. Then we have that

$$\tilde{\epsilon}_{\alpha_m \alpha_1 \cdots \alpha_{m-1}} \frac{\partial x^{\alpha_m}}{\partial y^m} e_1^{\alpha_1} \cdots e_{m-1}^{\alpha_{m-1}} < 0$$

If we demand that the direction of n^α is the opposite of $\frac{\partial x^\alpha}{\partial y^m}$, then we have

$$\tilde{\epsilon}_{\alpha_m \alpha_1 \cdots \alpha_{m-1}} n^{\alpha_m} e_1^{\alpha_1} \cdots e_{m-1}^{\alpha_{m-1}} > 0.$$

Surface element

We define the surface element of a hypersurface as

$$d\Sigma_\mu = \epsilon_{\mu\alpha\beta\gamma} e_1^\alpha e_2^\beta e_3^\gamma dy^1 \wedge dy^2 \wedge dy^3$$

It is easy to verify that

$$f^*(\sqrt{-g} dx^1 \wedge dx^2 \wedge dx^3) = d\Sigma_0$$
$$f^*(-\sqrt{-g} dx^0 \wedge dx^2 \wedge dx^3) = d\Sigma_1$$

and so on. We can demonstrate that

$$d\Sigma_\mu = \epsilon n_\mu |h|^{\frac{1}{2}} dy^1 \wedge dy^2 \wedge dy^3$$

Element of two-surface

Within the hypersurface Σ, we can define a two-surface S, which is parametrized with $y^a = y^a(\theta_A)$, then

$$e_A^a = \frac{\partial y^a}{\partial \theta^A}, \quad e_A^\alpha = \frac{\partial x^\alpha}{\partial \theta^A} = e_a^\alpha e_A^a$$
$$\sigma_{AB} = h_{AB} e_A^a e_B^b = g_{\alpha\beta} e_A^\alpha e_B^\beta$$
$$h^{ab} = \epsilon_r r^a r^b + \sigma^{AB} e_A^a e_B^b$$
$$g^{\alpha\beta} = \epsilon_n n^\alpha n^\beta + \epsilon_r r^\alpha r^\beta + \sigma^{AB} e_A^\alpha e_B^\beta$$

If we demand that the direction r^a is the opposite of that of $\frac{\partial y^a}{\partial \theta^1}$, then the condition of compatibility can be written as

$$\epsilon_{\mu\nu\beta\gamma} n^\mu r^\nu e_2^\beta e_3^\gamma > 0$$

We define the surface element of a two-surface as

$$dS_{\mu\nu} = \epsilon_{\mu\nu\beta\gamma} e_2^\beta e_3^\gamma d\theta^2 \wedge d\theta^3$$

It is easy to verify that

$$f^*(\sqrt{-g} dx^2 \wedge dx^3) = dS_{01}$$

and so on. We can demonstrate that

$$dS_{\alpha\beta} = \epsilon_n \epsilon_r (n_\alpha r_\beta - n_\beta r_\alpha) \sqrt{\sigma} d\theta^2 \wedge d\theta^3$$

Gauss-Stokes theorem

Linear algebra theory tells us that

$$\frac{\partial \det A}{\partial A_{ab}} = (\det A) A^{-1}_{ab}$$

Then we can deduce that

$$\frac{\partial g}{\partial x^\mu} = g g^{\alpha\beta} g_{\alpha\beta,\mu}$$

If A^μ is a vector, we have

$$A^\mu_{;\mu}\sqrt{-g} = (A^\mu\sqrt{-g})_{,\mu}$$

If $B^{\mu\nu}$ is an antisymmetric tensor, we have

$$B^{\mu\nu}_{;\nu}\sqrt{-g} = (B^{\nu\mu}\sqrt{-g})_{,\nu}$$

Apply the general Stokes theorem in differential geometry,

$$\int_M d\omega = \int_{\partial M} i^*\omega$$

we can derive the following theorem,

Theorem 10.1 Gauss-Stokes theorem

1. $$\int_{\mathcal{V}} A^\alpha_{;\alpha}\sqrt{-g}dx^4 = \oint_{\partial\mathcal{V}} A^\alpha d\Sigma_\alpha = \oint_{\partial\mathcal{V}} \epsilon A^\alpha n_\alpha \sqrt{h} dy^3$$

2. $$\int_\Sigma B^{\alpha\beta}{}_{;\beta} d\Sigma_\alpha = \frac{1}{2}\oint_{\partial\Sigma} B^{\alpha\beta} dS_{\alpha\beta} = \oint_{\partial\Sigma} \epsilon_n \epsilon_r B^{\alpha\beta} n_\alpha r_\beta \sqrt{\sigma} d\theta^2$$

   where $B_{\alpha\beta}$ is an antisymmetric tensor

♣

10.4.3 Differentiation of tangent vector fields

Tangent tensor field

$$A^{\alpha\beta\cdots} = A^{ab\cdots} e^\alpha_a e^\beta_b \cdots$$

$$A_{\alpha\beta\ldots} e^\alpha_a e^\beta_b \cdots = A_{ab\ldots} = h_{am} h_{bn} \cdots A^{mn\cdots}$$

Projection tensor

$$h^{\alpha\beta} \equiv h^{ab} e^\alpha_a e^\beta_b = g^{\alpha\beta} - \epsilon n^\alpha n^\beta$$

Intrinsic covariant derivative

$$A_{a|b} \equiv A_{\alpha;\beta} e^\alpha_a e^\beta_b = A_{a,b} - \Gamma^c_{ab} A_c$$

Here, the connection Γ^c_{ab} is compatible with h_{ab}.

Extrinsic curvature

$$K_{ab} \equiv n_{\alpha;\beta} e^\alpha_a e^\beta_b$$
$$A^\alpha_{;\beta} e^\beta_b = A^a_{|b} e^\alpha_a - \epsilon A^a K_{ab} n^\alpha$$
$$e^\alpha_{a;\beta} e^\beta_b = \Gamma^c_{ab} e^\alpha_c - \epsilon K_{ab} n^\alpha$$
$$K_{ab} = \frac{1}{2}(\mathcal{L}_n g_{\alpha\beta}) e^\alpha_a e^\beta_b$$
$$K \equiv h^{ab} K_{ab} = n^\alpha_{;\alpha}$$

Theorem 10.2 Gauss-Codazzi theorem

1.
$$R^\mu_{\alpha\beta\gamma} e^\alpha_a e^\beta_b e^\gamma_c = R^m_{abc} e^\mu_m + \epsilon(K_{ab|c} - K_{ac|b}) n^\mu + \epsilon K_{ab} n^\mu_{;\gamma} e^\gamma_c - \epsilon K_{ac} n^\mu_{;\beta} e^\beta_b$$

2.
$$-2\epsilon G_{\alpha\beta} n^\alpha n^\beta = \ {}^3R + \epsilon(K^{ab} K_{ab} - K^2)$$
$$G_{\alpha\beta} e^\alpha_a n^\beta = K^b_{a|b} - K_{,a}$$

3.
$$R = \ {}^3R + \epsilon(K^2 - K^{ab} K_{ab}) + 2\epsilon(n^\alpha_{;\beta} n^\beta - n^\alpha n^\beta_{;\beta})_{;\alpha}$$

10.5 Killing vector

A vector filed that satisfies the Killing equation

$$\mathcal{L}_\xi g = 0$$

is called a Killing vector field; it keeps the metric invariant and therefore corresponds to a space-time symmetry. A few lines of algebra can show that it is equivalent to define killing vector field by

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0$$

Derivatives of killing vectors can be related to Riemann tensor by

$$\nabla_\mu \nabla_\sigma \xi^\rho = R^\rho_{\sigma\mu\nu} \xi^\nu$$

This shows that from the value of ξ^μ and $\xi^\mu_{;\nu}$ at a given point one can determine the Killing vector field uniquely. One should then specify N values for ξ^ν and $\frac{N(N-1)}{2}$ values for $\xi^\mu_{;\nu}$, so that there are at most $N(N+1)/2$ linearly independent Killing vector fields. For $N = 4$, this is at most 10 Killing vectors, which is precisely the dimension of the Poincare group of Minkowski space.

A space-time enjoying the maximum number of Killing vector fields is called a maximally

symmetric space-time. It can be shown from the Killing equations the Riemann tensor must then satisfy

$$R^{\rho}_{\sigma\nu\nu} = \frac{R}{N(N-1)}(\delta^{\rho}_{\mu} g_{\sigma\nu} - \delta^{\rho}_{\nu} g_{\sigma\mu})$$

After contraction, we have

$$R_{\sigma\nu} = \frac{R}{N} g_{\sigma\nu}$$

Also, the Ricci scalar must be a constant. Maximally symmetric spaces are thus spaces of constant curvature. For $N = 4$, there are three maximally symmetric space-times: Minkowski , de Sitter and anti-de Sitter.

Now, consider a geodesic with tangent vector u^{μ}. We have

$$u^{\nu}(u^{\mu}\xi_{\mu})_{;\nu} = 0$$

So, the quantity $u^{\mu}\xi_{\mu}$ is constant along the trajectory. Similarly, if $T^{\mu\nu}_{;\nu} = 0$ and $T^{\mu\nu} = T^{\nu\mu}$, we have

$$(T^{\mu\nu}\xi_{\nu})_{;\mu} = 0$$

This relation has an important implication. Define $I^{\mu} \equiv T^{\mu\nu}\xi_{\nu}$. Gauss-Stokes theorem implies that

$$\oint_{\partial V} I^{\mu} n_{\mu} \sqrt{h} d^3y = 0$$

So

$$Q = \int_{\Sigma_t} I^{\mu} n_{\mu} \sqrt{h} d^3y$$

is a conservation charge.

Chapter 11
Formulation of General Relativity

Give the fields that generate mass-energy, and their time-rates of change, and give 3-geometry of space and its time-rate of change, all at one time, and solve for the 4-geometry of spacetime at that one time. Four of the ten components of Einstein's law connect the curvature of space here and now with the distribution of mass-energy here and now, and the other six equations tell how the geometry as thus determined then proceeds to evolve.

11.1 Basic assumptions of general relativity

1. Space-time is a four dimensional pseudo-Riemann manifold.

2. The metric of the manifold is governed by the Einstein field equation

$$\boldsymbol{G} = 8\pi \boldsymbol{T}.$$

3. All special relativistic laws of physics are valid in local Lorentz frames of metric.

11.2 Lagrangian formulation

11.2.1 Mechanics

$$S[q] = \int_{\tau_1}^{\tau_2} L(x^\alpha, \frac{dx^\alpha}{d\tau})d\tau$$

$$\delta x^\alpha(\tau_1) = 0, \delta x^\alpha(\tau_2) = 0$$

$$\delta S = 0 \Rightarrow \frac{d}{d\tau}\frac{\partial L}{\partial u^\alpha} - \frac{\partial L}{\partial x^\alpha} = 0$$

Example:

$$L = -m(-g_{\mu\nu}u^\mu u^\nu)^{1/2} + eA_\mu u^\mu$$

$$\Downarrow$$

$$m(\frac{du_\alpha}{d\tau} - \frac{1}{2}\frac{\partial g_{uv}}{\partial x^\alpha}u^\mu u^\nu) = e(A_{\mu,\alpha} - A_{\alpha,\mu})u^\mu \Rightarrow ma_\alpha = eF_{\alpha\mu}u^\mu$$

11.2.2 Field Theory

$$S[q] = \int_{\mathcal{V}} \mathcal{L}(q, q_\alpha)\sqrt{-g}d^4x$$
$$\delta q|_{\partial\mathcal{V}} = 0$$
$$\delta S = 0 \Rightarrow \nabla_\alpha(\frac{\partial \mathcal{L}}{\partial q_{,\alpha}}) - \frac{\partial \mathcal{L}}{\partial q} = 0$$

Example:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + A_\mu j^\mu$$
$$\Downarrow$$
$$F^{\mu\nu}{}_{;\nu} = j^\mu$$

11.2.3 General relativity

$$S_H[g] = \frac{1}{16\pi}\int_{\mathcal{V}} R\sqrt{-g}d^4x$$
$$S_B[g] = \frac{1}{8\pi}\oint_{\partial\mathcal{V}} \epsilon K|h|^{\frac{1}{2}}d^3y$$
$$S_0 = \frac{1}{8\pi}\oint_{\partial\mathcal{V}} \epsilon K_0|h|^{\frac{1}{2}}d^3y$$
$$S_M[\phi; g] = \int_{\partial\mathcal{V}} \mathcal{L}(\phi, \phi_{,\alpha}; g_{\alpha\beta})\sqrt{-g}d^4x$$

Variation of Hilbert term

$$(16\pi)\delta S_H = \int_{\mathcal{V}} G_{\alpha\beta}\delta g^{\alpha\beta}\sqrt{-g}d^4x - \oint_{\partial\mathcal{V}} \epsilon h^{\alpha\beta}\delta g_{\alpha\beta,\mu}n^\mu|h|^{\frac{1}{2}}d^3y$$

Variation of boundary term

$$16\pi\delta S_B = \oint_{\partial\mathcal{V}} \epsilon h^{\alpha\beta}\delta g_{\alpha\beta,\mu}n^\mu|h|^{\frac{1}{2}}d^3y$$

Variation of matter action

$$\delta S_M = \int_{\mathcal{V}} \left(\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} - \frac{1}{2}\mathcal{L}g_{\alpha\beta}\right)\delta g^{\alpha\beta}\sqrt{-g}d^4x$$

Define

$$T_{\alpha\beta} \equiv -2\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} + \mathcal{L}g_{\alpha\beta}$$

Example:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$$
$$\Downarrow$$

$$T_{\alpha\beta} = F_{\mu\alpha}F^{\mu}{}_{\beta} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}g_{\alpha\beta}$$

Nondynamical term

$K_0 =$ extrinsic curvature of $\partial\mathcal{L}$ embedded in flat space-time.

11.3 Hamiltonian formulation

11.3.1 3+1 decomposition

Figure 11.1: Foliation of space-time by space-like hypersurfaces

Figure 11.2: Decomposition of t^α into lapse and shift

The space-time is foliated by spacelike hypersurfaces Σ_t that is described by scalar function $t(x^\alpha)$. t is a single valued function and the unit normal to the hypersurfaces $n_\alpha \propto \partial_\alpha t$ is a future directed timelike vector field.
Consider a congruence of curves γ intersecting Σ_t. We use t as a parameter on the curves and the vector t^α is tangent to the congruence (so, $t^\alpha\partial_\alpha t = 1$).Install coordinates y^a on Σ_t and impose $y^a(P'') = y^a(P') = y^a(P)$, so y^a is held constant on each member of the congruence. This construction defines a coordinate system (t, y^a) in $\mathcal{V}$.

base vector

$$t^\alpha = \left(\frac{\partial x^\alpha}{\partial t}\right)_{y^a}, \quad e^\alpha_a = \left(\frac{\partial x^\alpha}{\partial y^a}\right)_t, \quad \mathcal{L}_t e^\alpha_a = 0$$

Normal vector

$$n_\alpha = -N\partial_\alpha t, \quad n_\alpha e^\alpha_a = 0$$

Decomposition of t^α

$$t^\alpha = Nn^\alpha + N^a e^\alpha_a$$

Metric

$$\begin{aligned} ds^2 &= g_{\alpha\beta}dx^\alpha dx^\beta \\ &= g_{\alpha\beta}(t^\alpha dt + e^\alpha_a dy^a)(t^\beta dt + e^\beta_b dy^b) \\ &= -N^2dt^2 + h_{ab}(dy^a + N^a dt)(dy^b + N^b dt) \\ \sqrt{-g} &= N\sqrt{h} \end{aligned}$$

11.3.2 Field theory

$$\dot{q} = \frac{\partial q}{\partial t}, \quad p = \frac{\partial}{\partial \dot{q}}(\sqrt{-g}\mathcal{L})$$

$$\mathcal{H}(p, q, q_a) = p\dot{q} - \sqrt{-g}\mathcal{L}$$

$$H = \int_{\Sigma_t} \mathcal{H}(p, q, q_{,a})d^3y$$

$$S = \int_{t_1}^{t_2} dt \int_{\Sigma_t} (p\dot{q} - \mathcal{H})d^3y$$

$$\delta S = 0 \Rightarrow \dot{p} = -\frac{\partial \mathcal{H}}{\partial q} + \left(\frac{\partial \mathcal{H}}{\partial q_{,a}}\right)_{,a}, \quad \dot{q} = \frac{\partial \mathcal{H}}{\partial p}$$

Example: For electromagnetic field in 3+1 decomposition form, we define the electrical field as $E_a = F_{\alpha\beta}n^\beta e^\alpha_a$, the magnetic field as $\epsilon_{abc}B^c = F_{\alpha\beta}e^\alpha_a e^\beta_b$. In this definition, the equation of motion of particles in electromagnetic field can be written as

$$mA_a = \gamma e(E_a + \epsilon_{abc}v^b B^c)$$

Here, $A_a = u_{\alpha;\beta}u^\beta e^\alpha_a$, $\gamma = \frac{1}{\sqrt{1-v^2}}$. So, the three force

$$\boldsymbol{f} = \frac{d\boldsymbol{p}}{dt} = e(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B})$$

If we adopt the coordinates (t, y^a), it is easy to verify that

$$E^a = NF^{0a}, \quad B^a = \frac{1}{2}\epsilon^{abc}F_{bc}$$

We further define

$$\mathcal{E}^a = \sqrt{h}E^a, \quad \mathcal{B}^a = \sqrt{h}B^a, \quad \phi = -A_0, \quad \rho_e = -j^\alpha n_\alpha = Nj^0$$

If we notice that

$$F_{0a} = -h_{ab}N^2F^{0b} - F_{ab}N^b, \quad \tilde{\epsilon}_{abc}\tilde{\epsilon}_{ijk}h^{ai}h^{bj} = \frac{2h_{ck}}{h}$$

It is easy to verify that

$$\sqrt{-g}\mathcal{L} = -\mathcal{E}^a \dot{A}_a + \phi \mathcal{E}^a_{,a} - \frac{1}{2} N h^{-\frac{1}{2}} h_{ab}(\mathcal{E}^a \mathcal{E}^b + \mathcal{B}^a \mathcal{B}^b) + \tilde{\epsilon}_{abc} N^a \mathcal{E}^b \mathcal{B}^c - \sqrt{h}\phi\rho_e + N\sqrt{h} A_a j^a$$

So,$\pi^a = -\mathcal{E}^a$, and we can get the Hamilton density

$$\mathcal{H} = \phi \pi^a_{,a} + \frac{1}{2} N h^{-\frac{1}{2}} h_{ab}(\pi^a \pi^b + \mathcal{B}^a \mathcal{B}^b) + \tilde{\epsilon}_{abc} N^a \pi^b \mathcal{B}^c + \sqrt{h}\phi\rho_e - N\sqrt{h} A_a j^a$$

Then, the Hamilton equation can be written as

$$\dot{A}_a = -\phi_{,a} + N h^{-\frac{1}{2}} h_{ab} \pi^b + \tilde{\epsilon}_{abc} N^a \mathcal{B}^c$$

$$\dot{\pi}^a = -\tilde{\epsilon}^{jab}(N h^{-\frac{1}{2}} h_{ij} \mathcal{B}^i)_{,b} - \tilde{\epsilon}^{cab}(\tilde{\epsilon}_{ijc} N^i \pi^j)_{,b} + N\sqrt{h} j^a$$

and also the constraint equation $\pi^a_{,a} + \sqrt{h}\rho_e = 0$. After simplification, the Maxwell equations are

$$\begin{aligned}
\frac{1}{\sqrt{h}}\frac{\partial}{\partial t}(\sqrt{h}\boldsymbol{E}) &= \nabla \times (N\boldsymbol{B} - \boldsymbol{N} \times \boldsymbol{E}) - N\boldsymbol{J} \\
\frac{1}{\sqrt{h}}\frac{\partial}{\partial t}(\sqrt{h}\boldsymbol{B}) &= -\nabla \times (N\boldsymbol{E} + \boldsymbol{N} \times \boldsymbol{B}) \\
\nabla \cdot \boldsymbol{E} &= \rho_e \\
\nabla \cdot \boldsymbol{B} &= 0
\end{aligned}$$

11.3.3 General relativity

$$S_G = \frac{1}{16\pi}\int_{t_1}^{t_2} dt \left\{ \int_{\Sigma_t} \left({}^3R + K^{ab}K_{ab} - K^2 \right) N\sqrt{h} d^3y + 2\oint_{\Sigma_t} (k - k_0) N\sqrt{\sigma} d^2\theta \right\}$$

$k_0 =$ extrinsic curvature of S_t embedded in flat space.

Gravitational Hamiltonian

$$\dot{h}_{ab} \equiv \mathcal{L}_t h_{ab} = \mathcal{L}_t(g_{\alpha\beta} e^\alpha_a e^\beta_b) = 2NK_{ab} + N_{a|b} + N_{b|a})$$

$$K_{ab} = \frac{1}{2N}(\dot{h}_{ab} - N_{a|b} - N_{b|a})$$

$$p^{ab} = \frac{\partial}{\partial \dot{h}_{ab}}(\sqrt{-g}\mathcal{L}_G) = \frac{\sqrt{h}}{16\pi}(K^{ab} - K h^{ab})$$

$$\sqrt{h} K^{ab} = 16\pi(p^{ab} - \frac{1}{2} p h^{ab})$$

$$\mathcal{H}_G = p^{ab}\dot{h}_{ab} - \sqrt{-g}\mathcal{L}_G$$

$$\begin{aligned}
16\pi H_G &= \int_{\Sigma_t} \left[N(K^{ab}K_{ab} - K^2 - {}^3R) - 2N_a(K^{ab} - K h^{ab})_{|b} \right] \sqrt{h} d^3y \\
&\quad -2\oint_{S_t} \left[N(k - k_0) - N_a(K^{ab} - K h^{ab}) r_b \right] \sqrt{\sigma} d^2\theta
\end{aligned}$$

Variation of gravitational Hamiltonian

$$\delta N = \delta N^a = \delta h_{ab} = 0 \text{ on } S_t$$

$$\delta H_G = \int_{\Sigma_t} (\mathcal{P}^{ab}\delta h_{ab} + \mathcal{H}_{ab}\delta p^{ab} - \mathcal{C}\delta N - 2\mathcal{C}_a\delta N^a) d^3y$$

$$\begin{aligned}
(16\pi)\mathcal{P}^{ab} &= N\sqrt{h}G^{ab} - \sqrt{h}(N^{|ab} - h^{ab}N^{|c}{}_c) \\
&\quad +(16\pi)[2p^{c(a}N^{b)}{}_{|c} - \sqrt{h}(\frac{1}{\sqrt{h}}p^{ab}N^c)_{|c}] \\
&\quad +(16\pi)^2[\frac{2N}{\sqrt{h}}(p^a_c p^{bc} - \frac{1}{2}pp^{ab}) - \frac{N}{2\sqrt{h}}(p^{cd}p_{cd} - \frac{1}{2}p^2)h^{ab}] \\
\mathcal{H}_{ab} &= (16\pi)\frac{2N}{\sqrt{h}}(p_{ab} - \frac{1}{2}ph_{ab}) + 2N_{(a|b)} \\
\mathcal{C} &= \frac{\sqrt{h}}{16\pi}(\ {}^3R + K^2 - K^{ab}K_{ab}) \\
\mathcal{C}^a &= \frac{\sqrt{h}}{16\pi}(K_a{}^b - K\delta_a{}^b)_{|b}
\end{aligned}$$

Variation of electromagnetic Hamiltonian

$$\delta H_E = \int_{\Sigma_t} (-\frac{1}{2}N\sqrt{h}\mathcal{I}^{ab}\delta h_{ab} + \sqrt{h}\rho\delta N - \sqrt{h}s_a\delta N^a)$$

$$\begin{aligned}
\mathcal{I}^{ab} &= \frac{1}{2}(E^cE_c + B^cB_c)h^{ab} - E^aE^b - B^aB^b \\
\rho &= \frac{1}{2}(E^cE_c + B^cB_c) \\
s_a &= \epsilon_{abc}E^bB^c
\end{aligned}$$

Hamilton's equations

$$\dot{h}_{ab} = \mathcal{H}_{ab}, \quad \dot{p}^{ab} = -\mathcal{P}^{ab} + \frac{1}{2}N\sqrt{h}\mathcal{I}^{ab}$$

$${}^3R + K^2 - K^{ab}K_{ab} = 16\pi\rho$$

$$(K_a{}^b - K\delta_a{}^b)_{|b} = -8\pi s_a$$

Chapter 12
Weak Field Limit

12.1 The linearized theory of gravity

In a weak-field situation

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad |h_{\mu\nu}| \ll 1,$$

one can expand the field equations in powers of $h_{\mu\nu}$ using a coordinate frame where $|h_{\mu\nu}| \ll 1$ holds; and without much loss of accuracy, one can keep only linear terms. The resulting formalism is often called "the linearized theory of gravity". The resulting connection coefficients, when linearized in the metric perturbation $h_{\mu\nu}$, read

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}\eta^{\mu\nu}(h_{\alpha\nu,\beta} + (h_{\beta\nu,\alpha} - h_{\alpha\beta,\nu}) \equiv \frac{1}{2}(h_{\alpha}{}^{\mu}{}_{,\beta} + h_{\beta}{}^{\mu}{}_{,\alpha} + h_{\alpha\beta}{}^{,\mu})$$

Whenever one expands in powers of $h_{\mu\nu}$, indices of $h_{\mu\nu}$ are raised and lowered using $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$, not $g^{\mu\nu}$ and $g_{\mu\nu}$. A similar linearization of the Ricci tensor yields

$$R_{\mu\nu} = \Gamma^{\alpha}{}_{\mu\nu,\alpha} - \Gamma^{\alpha}{}_{\mu\alpha,\nu} = \frac{1}{2}(h_{\mu}{}^{\alpha}{}_{,\nu\alpha} + h_{\nu}{}^{\alpha}{}_{,\mu\alpha} - h_{\mu\nu,\alpha}{}^{\alpha} - h_{,\mu\nu})$$

where

$$h \equiv \eta_{\mu\nu}h^{\mu\nu}$$

After a further contraction to form $R \equiv g^{\mu\nu}R_{\mu\nu} \approx \eta^{\mu\nu}R_{\mu\nu}$, one finds the Einstein equations read

$$h_{\mu\alpha,\nu}{}^{\alpha} + h_{\nu\alpha,\mu}{}^{\alpha} - h_{\mu\nu,\alpha}{}^{\alpha} - h_{,\mu\nu} - \eta_{\mu\nu}(h^{\alpha\beta}{}_{,\alpha\beta} - h_{,\alpha}{}^{\alpha}) = 16\pi T_{\mu\nu}$$

Define

$$\overline{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$$

the linearized field equations become

$$H^{\mu\alpha\nu\beta}{}_{,\alpha\beta} = 16\pi T_{\mu\nu}$$

where

$$-H^{\mu\alpha\nu\beta} \equiv \overline{h}^{\mu\nu}\eta^{\alpha\beta} + \overline{h}^{\alpha\beta}\eta^{\mu\nu} - \overline{h}^{\alpha\nu}\eta^{\mu\beta} - \overline{h}^{\mu\beta}\eta^{\alpha\nu}$$

Two different types of coordinate transformations connect nearly globally Lorentz systems to each other: global Lorentz transformations, and infinitesimal coordinate transformations.
Global Lorentz Transformations:

$$x^{\mu} = \Lambda^{\mu}{}_{\alpha'}x^{\alpha'} \quad \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}\eta_{\mu\nu} = \eta_{\alpha'\beta'}$$

We can verify that $h_{\mu\nu}$ and $\bar{h}_{\mu\nu}$ transform like components of a tensor in flat spacetime

$$h_{\alpha'\beta'} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}h_{\mu\nu}$$

Infinitesimal Coordinate Transformations:

$$x^{\mu'} = x^{\mu} + \xi^{\mu}$$

where ξ^{μ} are four arbitrary functions small enough to leave $|h_{\mu'\nu'}| \ll 1$. We can verify that the metric perturbation functions in the new $x^{\mu'}$ and old x^{μ} coordinate systems are related by

$$h^{\text{new}}_{\mu\nu} = h^{\text{old}}_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu}$$

whereas the functional forms of all other scalars, vectors, and tensors which is of order $O(h)$, such as $R_{\mu\nu}$, $T_{\mu\nu}$ and R, are unaltered, to within the precision of linearized theory.
For any physical situation, one can specialize the gauge so that

$$\overline{h}^{\mu\alpha}{}_{,\alpha} = 0$$

called Lorentz gauge. The Lorentz gauge is not fixed uniquely. The gauge condition is left unaffected by any gauge transformation for which

$$\Box\xi \equiv \xi^{\alpha,\beta}{}_{\beta} = 0$$

The field equations then become

$$\Box h_{\mu\nu} = -16\pi T_{\mu\nu}$$

Once the gauge has been fixed by fiat for a given system, one can regard $h_{\mu\nu}$ and $h_{\mu\nu}$ as components of tensors in flat spacetime; and one can regard the field equations and the chosen gauge conditions as geometric, coordinate-independent equations in flat spacetime. This viewpoint allows one to use curvilinear coordinates, if one wishes. But in doing so, one must everywhere replace the Lorentz components of the metric $\eta_{\mu\nu}$ by the metric's components $g_{\mu\nu_{\text{flat}}}$ in the flat-spacetime curvilinear coordinate system; and one must replace all ordinary derivatives in the field equations and gauge conditions by covariant derivatives whose connection coefficients come from $g_{\mu\nu_{\text{flat}}}$.

12.2 Nearly Newtonian gravitational fields

The general solution to the linearized field equations in Lorentz gauge lends itself to expression as a retarded integral of the form familiar from electromagnetic theory:

$$\overline{h}_{\mu\nu}(t, \boldsymbol{x}) = \int \frac{4T_{\mu\nu}(t - R, \boldsymbol{x}')}{R} d^3x' \quad R = |\boldsymbol{x} - \boldsymbol{x}'|$$

Here focus attention on a nearly Newtonian source: $T_{00} \gg |T_{0i}|$ and $T_{00} \gg |T_{ij}|$, and velocities slow enough that retardation is negligible. In this case, we have

$$\overline{h}_{00} = -4\Phi \quad \overline{h}_{0i} = \overline{h}_{ij} = 0$$

$$\Phi(t,\boldsymbol{x}) = -\int \frac{T_{00}(t,\boldsymbol{x}')}{R} d^3x' = \text{ Newtonian potential}$$

The corresponding metric is

$$ds^2 = -(1+2\Phi)dt^2+(1-2\Phi)(dx^2+dy^2+dz^2) \approx -(1-\frac{2M}{r})dt^2+(1+\frac{2M}{r})(dx^2+dy^2+dz^2)$$

For a test particle whose velocity $v \ll 1$, the geodesic equation can be written as

$$\frac{d^2v^i}{dt^2} + \Gamma^i_{00} = \frac{d^2v^i}{dt^2} + \Phi_{,i} = 0$$

So, we reproduce the classical Newtonian gravitation theory.

12.3 Gravitational wave

Let us decompose $h_{\mu\nu}$ as

$$h_{00} = -2A \quad h_{0i} = \partial_i B + \bar{B}_i \quad h_{ij} = 2C\delta_{ij} + 2\partial_i\partial_j E + \partial_i \bar{E}_j + \partial_j \bar{E}_i + \tilde{E}_{ij}$$

with

$$\partial_i \bar{B}_i = 0 \quad \partial_i \bar{E}_i = 0 \quad \partial_i \tilde{E}_{ij} = 0 \quad \tilde{E}^i_i = 0$$

Then we decompose the displacement vector for gauge transformation as

$$\xi^0 = -T \quad \xi^i = -\partial^i L - \bar{L}^i$$

with

$$\partial_i \bar{L}^i = 0$$

Under such coordinate transformation , the metric transform as

$$A \to A + \dot{T} \quad B \to B + \dot{L} - T \quad C \to C \quad E \to E + L$$

for the scalar modes,

$$\bar{B}_i \to \bar{B}_i + \dot{\bar{L}}_i \quad \bar{E}_i \to \bar{E}_i + \bar{L}_i$$

for the vector modes, and the tensor modes remain unchanged

$$\tilde{E}_{ij} \to \tilde{E}_{ij}$$

The tensor modes are therefore gauge invariant since they do not depend on the choice of the coordinate system. This is not the case for the vector and scalar modes. However, we can define a combination of these modes that are gauge invariant. For the scalar modes, we define

$$\Phi = A + \dot{B} - \ddot{E} \quad \Psi = -C$$

and for the vector modes

$$\bar{\Phi}_i = \dot{\bar{E}}_i - \bar{B}_i$$

We therefore have defined 2 scalar quantities and 1 vector quantity (2 degrees of freedom) which are gauge invariant. All together , we therefore have 6 degrees of freedom, once the 4

arbitrary degrees of freedom related to the gauge choice have been absorbed.

For the scalar modes, the Einstein equation $G_{\mu\nu} = 0$ imply that

$$\nabla^2\Psi = 0 \quad \partial_i\dot{\Psi} = 0 \quad \Phi - \Psi = 0$$

The only regular solutions are

$$\Phi = \Psi = 0$$

which means that no scalar mode can propagate.
For the vector modes, we have

$$\nabla^2\bar{\Phi}_i = 0$$

the only regular solution is $\bar{\Phi}_i = 0$. Just as for scalar modes, no vector modes can propagate.
For tensor modes, we have

$$\Box\tilde{E}_{ij} = 0$$

So, the only perturbations that can propagate in a Minkowski space-time are the gravitational waves and they satisfy

$$\Box\tilde{E}_{ij} = 0 \quad \partial_i\tilde{E}_{ij} = 0 \quad \tilde{E}^i_i = 0$$

The three conditions $\Phi = \Psi = \bar{\Phi}_i = 0$ define a gauge equivalence class. We can choose a gauge in this family by imposing some conditions on the perturbations. For instance, setting $E = B = 0$ and $\bar{E}_i = 0$, we define what is called a transverse and traceless (TT) gauge in which the metric is completely determined. In this case, the only non-vanishing component of $h^{\mathrm{TT}}_{\mu\nu}$ is $\tilde{E}_{ij}$.

Let us consider a plane wave propagating along the axis z in the TT gauge. Since the wave is transverse and traceless, it satisfies $h^{\mathrm{TT}}_{xz} = h^{\mathrm{TT}}_{yz} = h^{\mathrm{TT}}_{zz} = 0$. Therefore, it has only two degrees of freedom (two polarizations) that should satisfy $h^{\mathrm{TT}}_{xx} = -h^{\mathrm{TT}}_{yy}$ ans $h^{\mathrm{TT}}_{xy} = h^{\mathrm{TT}}_{yx}$. We hence can decompose it as following

$$h^{\mathrm{TT}}_{ij} = E_+(t-z)\epsilon^+_{ij} + E_\times(t-z)\epsilon^\times_{ij}$$

with the two polarization tensors being define by

$$\epsilon^+ \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \epsilon_\times \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The solution of the propagation equation is $E_{+/\times} = A_{+/\times}\cos(k(z-t)+\phi)$.

Consider the geodesic equation of a test particle in the gravitational field of a gravitational wave in the TT gauge. Since to leading order in perturbation we have $\Gamma^i_{00} = 0$, a particle initially at rest will remain at rest.
Of course, this does not mean that nothing happens , but rather that the frame of reference is co moving with the test particle. To see if anything happens , we should look at the relative motion of two neighbouring particles, which can be done using the geodesic deviation equation. The relative acceleration is given by

$$\nabla_{\boldsymbol{u}}\nabla_{\boldsymbol{u}}\boldsymbol{n} = -R(\boldsymbol{n},\boldsymbol{u})\boldsymbol{u}$$

To leading order in v, we have $u^\mu = (1, 0)$ and $n^\mu = (0, n^i)$, $R^0_{i0j} = -\frac{1}{2}\ddot{h}^{\rm TT}_{ij}$. So that

$$\frac{d^2 n^i}{dt^2} = -\frac{1}{2}\ddot{h}^{\rm TT}_{ij} n^j$$

Following figure illustrates the deformation of a ring of particles under the influence of a gravitational plane wave propagating along the axis z for each polarization.

Figure 12.1: Effects of a gravitational plane wave propagating along the axis z on a ring of particles located in the plane xy, depending on the wave polarization.

Chapter 13
Mass And Angular Momentum of A Gravitating System

13.1 External field of a gravitating source

Consider an isolated system with gravity so weak that in calculating its structure and motion one can completely ignore self-gravitational effects. Calculate the weak gravitational field

$$\bar{\bar{h}}_{\mu\nu}(t,\boldsymbol{x}) \equiv h_{\mu\nu}(t,\boldsymbol{x}) = \int \frac{4\bar{T}_{\mu\nu}(t-R,\boldsymbol{x}')}{R} d^3x'$$

produced by such a system. Restrict attention to the spacetime region far outside the system, and expand $h_{\mu\nu}$ in powers $\frac{\boldsymbol{x}'}{r} \equiv \frac{\boldsymbol{x}'}{r}$, we can get

$$\begin{aligned} ds^2 &= -\left[1-\frac{2M}{r}+O(\frac{1}{r^3})\right]dt^2 - \left[4\epsilon_{jkl}S^k\frac{x^l}{r^3}+O(\frac{1}{r^3})\right]dtdx^i \\ &+ \left[(1+\frac{2M}{r})\delta_{jk} + \text{ gravitational radiation terms that die out as } O(\frac{1}{r}))\right]dx^jdx^k \end{aligned}$$

where

$$M \equiv \int T^{00}d^3x \quad S_k \equiv \int \epsilon_{klm}x^lT^{m0}d^3x$$

It is plausible to say the mass of the system is M and the angular momentum $\boldsymbol{S}$. With an appropriate choice of gauge, g_{00} far from any weak source are time-independent and are determined uniquely by the source's mass M; g_{0j} is time-independent and is fixed by the source's intrinsic angular momentum S^j; but g_{jk} has time-dependent terms (gravitational waves!) of $O(\frac{1}{r})$.

The values of a system's mass and angular momentum can be measured by probing the imprint they leave in its external gravitational field. Of all tools one might use to probe, the simplest is a test particle in a gravitationally bound orbit. If the particle is sufficiently far from the source, its motion is affected hardly at all by the source's angular momentum or by the gravitational waves; only the spherical, Newtonian part of the gravitational field has a significant influence. Hence, the particle moves in an elliptical Keplerian orbit. To determine the source's mass M, one need only apply Kepler's third law

$$M = \left(\frac{2\pi}{\text{orbital period}}\right)^2 (\text{ Semi-major axis of ellipse })^3$$

Angular momentum can be probed by a gyroscope. Place a gyroscope at rest in the source's gravitational field. By a force applied to its center of mass, prevent it from falling. As time passes, the g_{0j} term in the metric will force the gyroscope to precess relative to the basis vectors $\frac{\partial}{\partial x^j}$; and since these basis vectors are "tied" to the coordinate system, which in turn is tied to the Lorentz frames at infinity, which in turn are tied to the "fixed stars" , the precession is relative to the "fixed stars." The angular velocity of precession is

$$\boldsymbol{\Omega} = \frac{1}{r^3}\left[-\boldsymbol{S} + \frac{3(\boldsymbol{S}\cdot\boldsymbol{x})\boldsymbol{x}}{r^3}\right]$$

One sometimes says that the source's rotation "drags the inertial frames near the source," thereby forcing the gyroscope to precess.

Consider an isolated, gravitating system inside which spacetime may or may not be highly curved. We restrict attention to the weak gravitational field far from the source, and analyze it using linearized theory in vacuum. Expand $h_{\mu\nu}$ in multipole moments and powers of $\frac{1}{r}$; and adjust the gauge, the Lorentz frame, and the origin of coordinates to simplify the resulting metric. We have

$$\begin{aligned} ds^2 &= -\left[1 - \frac{2M}{r} + \frac{2M^2}{r^2} + O(\frac{1}{r^3})\right] dt^2 - \left[4\epsilon_{jkl}S^k\frac{x^l}{r^3} + O(\frac{1}{r^3})\right] dtdx^i \\ &+ \left[(1 + \frac{2M}{r} + \frac{3M^2}{2r^2})\delta_{jk} + \text{ gravitational radiation terms that die out as } O(\frac{1}{r}))\right] dx^j dx^k \end{aligned}$$

However, M and $\boldsymbol{S}$ is not a simple integration of T^{00} and $\epsilon_{klm}x^l T^{0m}$ of the system as in the case of weakly gravitating sources. However, since this integration is impossible to measure in practice. We always determine the mass of the star by studying the orbits of planets in its external gravitational field. So, one defines the "total mass-energy" M of the sun or other body to be the constant that appears in the line element for its distant external spacetime geometry. Similarly, one defines the body's intrinsic angular momentum as the constant 3-vector $\boldsymbol{S}$ appearing in its line element.

13.2 Conservation laws for 4-momentum and angular momentum

Recall in the weak field limit, the Einstein equation can be written as

$$H^{\mu\alpha\nu\beta}{}_{,\alpha\beta} = 16\pi T^{\mu\nu}$$

So we have

$$T^{\mu\nu}{}_{,\nu} = \frac{1}{16\pi}H^{\mu\alpha\nu\beta}{}_{,\alpha\beta\nu} = 0$$

The source's total 4-momentum can be defined as

$$P^\mu \equiv \int T^{\mu 0}d^3x = \frac{1}{16\pi}\oint_S H^{\mu\alpha 0j}{}_{,\alpha}d^2S_j$$

Particularly, we have

$$P^0 = \frac{1}{16\pi}\oint_S (g_{jk,k} - g_{kk,j})d^2S_j$$

The angular momentum of the source can be defined as

$$J^{\mu\nu} \equiv \int (x^\mu T^{0\nu} - x^\nu T^{0\mu})d^3x = \frac{1}{16\pi}\oint_S (x^\mu H^{\nu\alpha 0j}{}_{,\alpha} - x^\nu H^{\mu\alpha 0j}{}_{,\alpha} + H^{\mu j0\nu} - H^{\nu j0\nu})d^2S_j$$

Note: *To evaluate the flux integrals (by contrast with the volume integrals), one need utilize only the gravitational field far outside the source. Since that gravitational field has the same form in full general relativity for strong sources as in linearized theory for weak sources, the flux integrals can be used to calculate P^μ and $J^{\mu\nu}$ for any isolated source whatsoever, weak or strong. So we will use the flux integrals as the definition.*

Knowing P^μ and $J^{\mu\nu}$, one can calculate the source's total mass-energy M intrinsic angular momentum S_{rho} by

$$M = (-P^\mu P_\mu)^{-1/2}$$

$$Y^\mu = -J^{\mu\nu}P_\nu/M^2$$

$$S_\rho = \frac{1}{2}\epsilon_{\mu\nu\sigma\rho}(J^{\mu\nu} - Y^\mu P^\nu + Y^\nu P^\mu)P^\sigma/M$$

Note especially that the integrands of the flux integrals are not gauge-invariant. In any local inertial frame at an event they vanish. However, the total integrals P^μ and $J^{\mu\nu}$ are. They have meaning and significance independent of any coordinate system and gauge. They are tensors in the asymptotically flat region surrounding the source.

In full general relativity, though $|h_{\mu\nu}| \ll 1$ breaks down, we can still define formally

$$-H^{\mu\alpha\nu\beta} \equiv \overline{h}^{\mu\nu}\eta^{\alpha\beta} + \overline{h}^{\alpha\beta}\eta^{\mu\nu} - \overline{h}^{\alpha\nu}\eta^{\mu\beta} - \overline{h}^{\mu\beta}\eta^{\alpha\nu} \quad h_{\mu\nu} \equiv g_{\mu\nu} - \eta_{\mu\nu}$$

And we define the effective energy-momentum pseudotensor by

$$H^{\mu\alpha\nu\beta}{}_{,\alpha\beta} = 16\pi T^{\mu\nu}_{\text{eff}}$$

So, we have

$$T^{\mu\nu}_{\text{eff},\nu} = 0$$

$$P^\mu = \frac{1}{16\pi}\int d^3x T^{\mu 0}_{\text{eff}}$$

$$J^{\mu\nu} = \int (x^\mu T^{0\nu}_{\text{eff}} - x^\nu T^{0\mu}_{\text{eff}})d^3x$$

for both strong or weak source. Define gravitation energy-momentum pseudotensor as

$$16\pi t^{\mu\nu} \equiv H^{\mu\alpha\nu\beta}{}_{,\alpha\beta} - 2G^{\mu\nu}$$

so we have

$$T^{\mu\nu}_{\text{eff}} = T^{\mu\nu} + t^{\mu\nu}$$

All the quantities $H^{\alpha\mu\nu\beta}$, $t^{\mu\nu}$ and $T^{\mu\nu}_{\text{eff}}$ depend for their definition and existence on the choice of coordinates. There is, nevertheless, adequate invariance under general coordinate transformations to give the values P^μ and $J^{\mu\nu}$ of the volume integrals geometric, coordinate-free significance in the asymptotically flat region far outside the source. Although this invariance is hard to see in the volume integrals themselves, it is clear from the surface-integral forms that no coordinate transformation which changes the coordinates only inside some spatially bounded region can influence the values of the integrals. For coordinate changes in the distant, asymptotically flat regions, linearized theory guarantees that under Lorentz transformations the integrals for P^μ and $J^{\mu\nu}$ will transform like special relativistic tensors, and that under infinitesimal coordinate transformations (gauge changes) they will be invariant.
It is clear that any quantities $H^{\alpha\mu\nu\beta}_{\text{new}}$ which agree with the original $H^{\alpha\mu\nu\beta}$ in the asymptotic weak-field region will give the same values as $H^{\alpha\mu\nu\beta}$ does for the P^μ and $J^{\mu\nu}$ surface integrals. One especially convenient choice is Landau-Lifshitz pseudotensor. The details can be found in section 96 of *The Classical Theory of Fields (L.D.Landau & E.M.Lifshitz)*

For a system of gravitating bodies, we have

$$\frac{dP^\mu}{dt} = -\oint T^{\mu j}_{\text{eff}} d^2 S_j$$

$$\frac{dJ^{\mu\nu}}{dt} = -\oint (x^\mu T^{\nu j}_{\text{eff}} - x^\nu T^{\mu j}_{\text{eff}}) d^2 S_j$$

The flux is integrated over the surface in asymptomatic flat region. Although the pseudotensor $t^{\mu\nu}$ in the interbody region and outside the system, contributes negligibly to the total 4-momentum and angular momentum, its contribution via gravitational waves to the time derivatives can be important when added up over astronomical periods of time. Thus, one must not ignore it in the flux integrals.

In evaluating these flux integrals, it is especially convenient to use the Landau- Lifshitz form of $T^{\mu\nu}_{\text{eff}}$, since that form contains no second derivatives of the metric. Thus set

$$T^{\mu\nu}_{\text{eff}} = (-g)(T^{\mu\nu} + t^{\mu\nu}_{\text{L-L}}) \approx T^{\mu\nu} + t^{\mu\nu}_{\text{L-L}}$$

Only those portions of $t^{\mu\nu}_{\text{L-L}}$ that die out as $\frac{1}{r^2}$ at large r can contribute to the flux integrals. For static solutions, $t^{\mu\nu}_{\text{L-L}}$ dies out as $\frac{1}{r^4}$. Hence, the only contributions come from dynamic parts of the metric, which, at these large distances, are entirely in the form of gravitational waves. The study of gravitational waves will reveal that when $t^{\mu\nu}_{\text{L-L}}$ is averaged over several wavelengths, it becomes a stress-energy tensor $T^{\mu\nu}_{\text{GW}}$ for the waves, which has all the properties one ever requires of any stress-energy tensor. One can freely make in these integrals the replacement

$$T^{\mu\nu}_{\text{eff}} = T^{\mu\nu} + T^{\mu\nu}_{\text{GW}}$$

So, the rate of loss of 4-momentum and angular momentum from the system, as measured gravitationally, is precisely equal to the rate at which matter, fields, and gravitational waves carry off 4-momentum and angular momentum.

Part IV

Quantum Mechanics

Chapter 14
Linear Algebra

14.1 Linear Vector Space

14.1.1 Definition

Definition 14.1 Linear vector space

A linear vector space is a set of elements, called vectors, which is closed under addition and multiplication by scalars. That is to say, if ϕ and ψ are vectors then so is $a\phi + b\psi$, where a and b are arbitrary scalars. If the scalars belong to the field of complex (real) numbers, we speak of a complex (real) linear vector space. Henceforth the scalars will be complex numbers unless otherwise stated.

Example:

1. Discrete vectors, which may be represented as columns of complex numbers.

2. Spaces of functions of some type, for example the space of all differentiable functions

14.1.2 Linear independence

Definition 14.2 Linear independence

A set of vectors $\{\phi_n\}$ is said to be linearly independent if no non-trivial linear combination of them sums to zero; that is to say, if the equation $\sum_n c_n\phi_n$ can hold only when $c_n = 0$ for all n. If this condition does not hold, the set of vectors is said to be linearly dependent, in which case it is possible to express a member of the set as a linear combination of the others.

Definition 14.3 Dimension

The maximum number of linearly independent vectors in a space is called the dimension of the space.

Definition 14.4 Base

A maximal set of linearly independent vectors is called a basis for the space. Any vector in the space can be expressed as a linear combination of the basis vectors.

14.1.3 Inner product

Definition 14.5 Inner product

An inner product (or scalar product) for a linear vector space associates a scalar (ϕ, ψ) *with every ordered pair of vectors. It must satisfy the following properties:*

1. $(\phi, \psi) =$ *a complex number*
2. $(\phi, \psi) = (\psi, \phi)^*$
3. $(\phi, c_1\psi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2)$
4. $(\phi, \phi) geq 0$*,with equality holding if and only if* $\phi = 0$

Example:

1. If ψ is the column vector with elements a_1, a_2, $\cdots$, and ϕ is the column vector with elements b_1, b_2, $\cdots$, then

$$(\psi, \phi) = a_1^* b_1 + a_2^* b_2 + \cdots$$

2. If ψ and ϕ are functions of x, then

$$(\phi, \psi) = \int \phi^*(x)\psi(x)w(x)dx$$

where $w(x)$ is some non-negative weight function.

Definition 14.6 Norm

$$||\phi|| = (\phi, \phi)^{\frac{1}{2}}$$

Theorem 14.1 Schwarz's inequality

$$|(\psi, \phi)|^2 \leq (\psi, \psi)(\phi, \phi)$$

Theorem 14.2 triangle inequality

$$||(\psi + \phi)|| \leq ||\phi|| + ||\psi||$$

Definition 14.7 Orthonormal

A set of vectors $\{\phi_n\}$ is said to be orthonormal if the vectors are pairwise orthogonal and of unit norm; that is to say, their inner products satisfy $(\psi_m, \phi_n) = \delta_{mn}$.

14.1.4 Dual space

Definition 14.8 Dual vector

Corresponding to any linear vector space V there exists the dual space of linear functionals on V . A linear functional F assigns a scalar $F(\phi)$ to each vector ϕ, such that

$$F(a\phi + b\psi) = aF(\phi) + bF(\psi)$$

for any vectors for ϕ and ψ, and any scalars a and b. The set of linear functionals may itself be regarded as forming a linear space V' if we define the sum of two functionals as

$$(F_1 + F_2)(\phi) = F_1(\phi) + F_2(\phi)$$

Theorem 14.3 Riesz theorem

There is a one-to-one correspondence between linear functionals F in V' and vectors f in V, such that all linear functionals have the form

$$F(\phi) = (f, \phi)$$

f being a fixed vector, and ϕ being an arbitrary vector. Thus the spaces V and V' are essentially isomorphic.

14.1.5 Dirac's bra and ket notation

In Dirac' s notation, which is very popular in quantum mechanics, the vectors in V are called ket vectors, and are denoted as $|\phi\rangle$. The linear functionals in the dual space V' are called bra vectors, and are denoted as $\langle F|$. The numerical value of the functional is denoted as

$$F(\phi) = \langle F|\phi\rangle$$

According to the Riesz theorem, there is a one-to-one correspondence between bras and kets. Therefore we can use the same alphabetic character for the functional (a member of V') and the vector (in V) to which it corresponds, relying on the bra, $\langle F|$, or ket, $|F\rangle$, notation to determine which space is referred to.So

$$\langle F|\phi\rangle = (F, \phi)$$

Note that the Riesz theorem establishes, by construction, an antilinear correspondence between bras and kets. If $\langle F| \leftrightarrow |F\rangle$, then

$$c_1^*\langle F_1| + c_2^*\langle F_2| \leftrightarrow c_1|F_1\rangle + c_2|F_2\rangle$$

14.2 Linear Operators

Definition 14.9 Linear operators

An operator on a vector space maps vectors onto vectors. A linear operator satisfies

$$A(c_1\psi_1 + c_2\psi_2) = c_1A(\psi_1) + c_2A(\psi_2)$$

Define the sum and product of operators,

$$\begin{aligned}(A+B)\psi &= A\psi + B\psi \\ AB\psi &= A(B\psi)\end{aligned}$$

Define their action to the left on bra vectors as

$$(\langle\phi|A)\psi\rangle = \langle\phi|(A|\psi\rangle)$$

So we may define the operation of A on the bra space of functionals as

$$AF_\phi(\psi) = F_\phi(A\psi)$$

According to the Riesz theorem there must exist a ket vector χ such that

$$AF_\phi(\psi) = (\chi, \psi) = F_\chi(\psi)$$

Define operator $A^\dagger$ as

$$AF_\phi = F_{A^\dagger\chi}$$

Therefore,

$$(A^\dagger\phi, \psi) = (\phi, A\psi)$$
$$\langle\psi|A^\dagger|\phi\rangle^* = \langle\phi|A|\psi\rangle$$

Definition 14.10 Outer product

$$(|\psi\rangle\langle\phi|)|\lambda\rangle \equiv |\psi\rangle(\langle\phi|\lambda\rangle)$$

Definition 14.11 Trace

$$\mathrm{Tr}A \equiv \sum \langle u_j|A|u_j\rangle$$

where $\{u_j\}$ may be any orthonormal basis. It can be shown that the value of $\mathrm{Tr}A$ is independent of the particular orthonormal basis that is chosen for its evaluation.

Proposition 14.1

$$\begin{aligned}(cA)^\dagger &= c^* A^\dagger \\ (A+B)^\dagger &= A^\dagger + B^\dagger \\ (AB)^\dagger &= B^\dagger A^\dagger\end{aligned}$$

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|$$

14.3 Self-Adjoint operators

Definition 14.12 Self-Adjoint operators

An operator A that is equal to its adjoint $A^\dagger$ is called self-adjoint. This means that it satisfies

$$\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^*$$

and that the domain of A coincides with the domain of $A^\dagger$. An operator that only satisfies above equation is called Hermitian.

Theorem 14.4

If $\langle\psi|A|\psi\rangle = \langle\psi|A|\psi\rangle^*$ for all $|\psi\rangle$, then it follows that $\langle\phi_1|A|\phi_2\rangle = \langle\phi_2|A|\phi_1\rangle^*$ for all $|\phi_1\rangle$ and $|\phi_2\rangle$, and hence that $A = A^\dagger$.
If an operator acting on a certain vector produces a scalar multiple of that same vector,

$$A|\phi\rangle = a|\phi\rangle$$

we call the vector $|\phi\rangle$ an eigenvector and the scalar a an eigenvalue of the operator A. The antilinear correspondence between bras and kets, and the definition of the adjoint operator $A^\dagger$, imply that the left-handed eigenvalue equation

$$\langle\phi|A^\dagger = a^*\langle\phi|$$

Theorem 14.5

If A is a Hermitian operator then all of its eigenvalues are real.

Theorem 14.6

Eigenvectors corresponding to distinct eigenvalues of a Hermitian operator must be orthogonal.

If the orthonormal set of vectors $\{\phi_i\}$ is complete, then we can expand an arbitrary vector $|v\rangle$ in terms of it:

$$|v\rangle = \sum |\phi_i\rangle(\langle\phi_i|v\rangle) = \left(\sum |\phi_i\rangle\langle\phi_i|\right)|v\rangle$$

So,

$$\sum |\phi_i\rangle\langle\phi_i| = I$$

If $A|\phi_i\rangle = a_i|\phi_i\rangle$ and the eigenvectors form a complete orthonormal set, then the operator can be reconstructed in a useful diagonal form in terms of its eigenvalues and eigenvectors:

$$A = \sum a_i|\phi_i\rangle\langle\phi_i|$$

We can define a function of an operator

$$f(A) = \sum f(a_i)|\phi_i\rangle\langle\phi_i|$$

The Hermitian operators in a finite N-dimensional vector space have complete sets of eigenvectors. But This statement does not carry over to infinite-dimensional spaces. A Hermitian operator in an infinite dimensional vector space may or may not possess a complete set of eigenvectors, depending upon the precise nature of the operator and the vector space. Instead, we have spectral theorem.

Theorem 14.7

To each self-adjoint operator A there corresponds a unique family of projection operators, $E(\lambda)$, for real λ, with the properties:

1. If $\lambda_1 < \lambda_2$ then $E(\lambda_1)E(\lambda_2) = E(\lambda_2)E(\lambda_1) = E(\lambda_1)$
2. If $\epsilon > 0$, then $E(\lambda + \epsilon)|\psi\rangle \to E(\lambda)|\psi\rangle$ as $\epsilon \to 0$
3. $E(\lambda)|\psi\rangle \to 0$ as $\lambda \to -\infty$
4. $E(\lambda)|\psi\rangle \to |\psi\rangle$ as $\lambda \to \infty$
5. $\int_{-\infty}^{\infty} \lambda dE(\lambda) = A$

We can define a function of an operator

$$f(A) = \int_{-\infty}^{\infty} f(\lambda) dE(\lambda)$$

Following Dirac's pioneering formulation, it has become customary in quantum mechanics to write a formal eigenvalue equation for an operator such as Q that has a continuous spectrum,

$$Q|q\rangle = q|q\rangle$$

The orthonormality condition for the continuous case takes the form

$$\langle q'|q''\rangle = \delta(q - q')$$

Evidently the norm of these formal eigenvectors is infinite, since $\langle q|q\rangle \to \infty$. Instead of the spectral theorem for Q, Dirac would write

$$Q = \int_{-\infty}^{\infty} q|q\rangle\langle q| dq$$

Dirac's formulation does not fit into the mathematical theory of Hilbert space, which admits only vectors of finite norm. The projection operator formally given by

$$E(\lambda) = \int_{-\infty}^{\lambda} |q\rangle\langle q| dq$$

is is well defined in Hilbert space, but its derivative does not exist within the Hilbert space framework.

Theorem 14.8

If A and B are self-adjoint operators, each of which possesses a complete set of eigenvectors, and if $AB = BA$, then there exists a complete set of vectors which are eigenvectors of both A and B.

Let $(A, B, \cdots)$ be a set of mutually commutative operators that possess a complete set of common eigenvectors. Corresponding to a particular eigenvalue for each operator, there may be more than one eigenvector. If, however, there is no more than one eigenvector (apart from the arbitrary phase and normalization) for each set of eigenvalues $(a_n, b_m, \cdots)$, then the operators $(A, B, \cdots)$ are said to be a complete commuting set of operators.

Theorem 14.9

Any operator that commutes with all members of a complete commuting set must be a function of the operators in that set. ♣

14.4 Rigged Hilbert space

Definition 14.13 Rigged Hilbert spece

Formally, a rigged Hilbert space consists of a Hilbert space $\mathcal{H}$, together with a subspace Φ which carries a finer topology, that is one for which the natural inclusion $\Phi \subseteq \mathcal{H}$ is continuous. It is no loss to assume that Φ is dense in $\mathcal{H}$ for the Hilbert norm. We consider the inclusion of conjugate space $\mathcal{H}^X$ in Φ^X. Φ^X is the space of τ_Φ continuous antilinear functional on Φ.
For any $\phi \in \Phi$, $F \in \Phi^X$,we define

$$\langle \phi | F \rangle \equiv F(\phi)$$

$$\langle F | \phi \rangle \equiv [F(\phi)]^*$$

Now by applying the Riesz representation theorem we can identify $\mathcal{H}^X$ with $\mathcal{H}$. Therefore, the definition of rigged Hilbert space is in terms of a sandwich:

$$\Phi \subseteq \mathcal{H} \subseteq \Phi^X$$

There may or may not exist any solutions to the eigenvalue equation $A|a_n\rangle = a_n|a_n\rangle$ for a self-adjoint operator A on an infinite-dimensional vector space. However, the generalized spectral theorem asserts that if A is self-adjoint in $\mathcal{H}$ then a complete set of eigenvectors exists in the extended space Φ^X. The precise conditions for the proof of this theorem are rather technical, so the interested reader is referred to *Gel'fand and Vilenkin (1964)* for further details.

There are many examples of rigged-Hilbert-space triplets. A Hilbert space $\mathcal{H}$ is formed by those functions that are square-integrable. That is, $\mathcal{H}$ consists of those functions $\psi(x)$ for which

$$\langle \psi | \psi \rangle = \int_{-\infty}^{\infty} |\psi(x)|^2 dx \text{ is finite}$$

A nuclear space Φ is made up of functions $\psi(x)$ which satisfy the infinite set of conditions,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 (1 + |x|)^m dx \text{ is finite for } m = 0, 1, 2, \cdots$$

The functions $\psi(x))$ which make up Φ must vanish more rapidly than any inverse power of x in the limit $|x| \to \infty$. The extended space Φ^X, which is conjugate to Φ, consists of those functions $\chi(x)$ for which

$$\langle\chi|\psi\rangle = \int_{-\infty}^{\infty} \chi^*(x)\psi(x)dx \text{ is finite for any } \psi \text{ in } \Phi$$

In addition to the functions of finite norm, which also lie in $\mathcal{H}$, Φ^X will contain functions that are unbounded at infinity provided the divergence is no worse than a power of x. Hence Φ^X contains e^{ikx}, which is an eigenfunction of the operator $D = i\frac{d}{dx}$. It also contains the Dirac delta function, $\delta(x-\lambda)$, which is an eigenfunction of the operator X, defined by $X\psi(x) = x\psi(x)$. These two examples suffice to show that rigged Hilbert space seems to be a more natural mathematical setting for quantum mechanics than is Hilbert space.

14.5 Unitary operators

Definition 14.14 Unitary operator

A unitary operator is a bounded linear operator $U : H \to H$ on a Hilbert space H that satisfies $UU^\dagger = U^\dagger U = I$, where $U^\dagger$ is the adjoint of U, and $I : H \to H$ is the identity operator.

Consider a family of unitary operators, $U(s)$, that depend on a single continuous parameter s. Let $U(0) = I$ be the identity operator, and let $U(s_1+s_2) = U(s_1)U(s_2)$. We can demonstrate that

$$\left.\frac{dU}{ds}\right|_{s=0} = iK \text{ with } K = K^\dagger$$

The Hermitian operator K is called the generator of the family of unitary operators because it determines $U(s)$, not only for infinitesimal s, but for all s. This can be shown by differentiating

$$U(s_1+s_2) = U(s_1)U(s_2)$$

with respect to s_2 and we can get

$$\left.\frac{dU}{ds}\right|_{s=s_1} = U(s_1)iK$$

This first order differential equation with initial condition $U(0) = I$ has the unique solution

$$U(s) = e^{iKs}$$

14.6 Antiunitary operators

Definition 14.15 Antiunitary operator

In mathematics, an antiunitary transformation, is a bijective antilinear map

$$U : H_1 \to H_2$$

between two complex Hilbert spaces such that

$$\langle Ux, Uy \rangle = \overline{\langle x, y \rangle}$$

for all x and y y in H_1, where the horizontal bar represents the complex conjugate. If additionally one has $H_1 = H_2$ then U is called an antiunitary operator.

Proposition 14.2

1. $\langle Ux, Uy \rangle = \overline{\langle x, y \rangle} = \langle y, x \rangle$ holds for all elements x, y of the Hilbert space and an antiunitary U.
2. When U is antiunitary then U^2 is unitary. This follows from
$$\langle U^2x, U^2y \rangle = \overline{\langle Ux, Uy \rangle} = \langle x, y \rangle$$
3. For unitary operator V the operator VK, where K is complex conjugate operator, is antiunitary. The reverse is also true, for antiunitary U the operator UK is unitary.
4. For antiunitary U the definition of the adjoint operator U^* is changed into
$$\langle U^*x, y \rangle = \overline{\langle x, Uy \rangle}$$
5. The adjoint of an antiunitary U is also antiunitary and $UU^* = U^*U = 1$.

Chapter 15
Formulation of quantum mechanics

15.1 Axioms of quantum mechanics

1. The properties of a quantum system are completely defined by specification of its state vector $|\psi\rangle$. The state vector is an element of a complex Hilbert space $\mathcal{H}$ called the space of states.

2. With every physical property A (energy, position, momentum, angular momentum, ...) there exists an associated linear, Hermitian operator A (usually called observable), which acts in the space of states. The eigenvalues of the operator are the possible values of the physical properties.

3.
   - If $|\psi\rangle$ is the vector representing the state of a system and if $|\phi\rangle$ represents another physical state, there exists a probability $P(|\psi\rangle, |\phi\rangle)$ of finding $|\psi\rangle$ in state $|\phi\rangle$, which is given by the squared modulus of the scalar product on $\mathcal{H}$: $P(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$ (Born Rule)
   - If A is an observable with eigenvalues a_k and eigenvectors $|k\rangle$, given a system in the state $|\psi\rangle$, the probability of obtaining a_k as the outcome of the measurement of A is $|\langle k|\psi\rangle|^2$. After the measurement the system is left in the state projected on the subspace of the eigenvalue a_k (Wave function collapse).

4. The evolution of a closed system is unitary. The state vector $|\psi(t)\rangle$ at time t is derived from the state vector $|\psi(t_0)\rangle$ at time t_0 by applying a unitary operator $U(t, t_0)$, called the evolution operator: $\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$.

15.2 Transformations of States

A transformation of states can be described by $|\psi\rangle \to U(\tau)|\psi\rangle \equiv |\psi'\rangle$. And we demand that

$$|\langle\phi|\psi\rangle| = |\langle\phi'|\psi'\rangle|$$

Theorem 15.1 Wigner Theorem

Any mapping of the vector space onto itself that preserves the value of $|\langle\phi|\psi\rangle|$ may be implemented by an operator U with U being either unitary (linear) or antiunitary (antilinear). ♣

Continuous transformation

Only linear operators can describe continuous transformations because every continuous transformation has a square root. Suppose, for example, that $U(l)$ describes a displacement through the distance l. This can be done by two displacements of $U(l/2)$, and hence $U(l) = U(l/2)U(l/2)$. The product of two antilinear operators is linear, since the second complex conjugation nullifies the effect of the first. Thus, regardless of the linear or antilinear character of $U(l/2)$, it must be the case that $U(l)$ is linear. A continuous operator cannot change discontinuously from linear to antilinear as a function of l, so the operator must be linear for all l.

Transformations of observables

For an observable Q,

$$\langle\phi'|Q|\phi'\rangle = \langle\phi|U^{-1}QU|\phi\rangle$$

If $U(\tau)^{-1}QU(\tau) = \tau Q$, we can prove that

$$U|q\rangle = |\tau q\rangle$$

Here, $|q\rangle$ is the eigenvector of Q with eigenvalue q.

15.3 Schrödinger equation

$U(t, t_0)$ is unitary and $U(t_2, t_0) = U(t_2, t_1)U(t_1, t_0)$. We can define $H(t_0)$ as

$$\frac{d}{dt}U(t, t_0)\bigg|_{t=t_0} = -iH(t_0) \text{ with } H(t_0) = H(t_0)^\dagger$$

We can demonstrate that

$$\frac{dU(t, t_0)}{dt}\bigg|_{t=t_1} = -iH(t_1)U(t_1, t_0)$$

The formal solution of the differential equation is

$$U(t, t_0) = I + (-i)^n \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n)$$

Suppose that T stands for time ordering, placing all operators evaluated at later times to the left, the above equation can be written as

$$U(t, t_0) = I + \frac{(-i)^n}{n!} \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T\{H(t_1)H(t_2)\cdots H(t_n)\} \equiv \exp\left[-iT\left\{\int_{t_0}^{t} H(t')dt'\right\}\right]$$

If the Hamiltonian operator H is time-dependent but the H's at different times commute. The equation above can be simplified to

$$U(t, t_0) = \exp\left[-i\int_{t_0}^{t} H(t')dt'\right]$$

If the H is time-independent, then

$$U(t,t_0) = \exp\left[-iH(t-t_0)\right]$$

Since $|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle$, we can derive the Schrödinger equation

$$\frac{d|\psi(t)\rangle}{dt} = -iH(t)|\psi(t)\rangle$$

The expectation value of an observable Q is $\langle\psi|Q|\psi\rangle$, denoted by $\langle Q\rangle$. We can then derive that

$$\frac{d\langle Q\rangle}{dt} = -i\left\{\langle[Q,H]\rangle + \langle\frac{\partial Q}{\partial t}\rangle\right\}$$

This is called Ehrenfest's theorem.

15.4 Position operators

In three dimensional space, for a particle, we have three operators corresponding to the observations of its position in space, $\boldsymbol{X} = (X_1, X_2, X_3)$. If the particle has some other internal degrees of freedom, then $\boldsymbol{X}$ plus some other observables S's will form a complete commuting set of operators. The eigenstate state will be denoted by $|\boldsymbol{x}, s\rangle$, satisfying that

$$X_i|\boldsymbol{x}, s\rangle = x_i|\boldsymbol{x}, s\rangle$$

It describes a particle posited in $\boldsymbol{x}$ with internal state s. And we will normalize $|\boldsymbol{x}, s\rangle$ by

$$\langle\boldsymbol{x}, s'|\boldsymbol{x}, s\rangle = \delta_{ss'}\delta(\boldsymbol{x}-\boldsymbol{x}')$$

15.5 Momentum operators and canonical quantization

Since $\boldsymbol{X}$ plus some other observables S's form a complete commuting set of operators. So, the momentum operators can not be independent of them. Numerous experiments shows that the position and momentum of particles can not be measured simultaneously. So, we expect $[X,P] \neq 0$.

Guess For a system which has a classical correspondence, the classical equation of motion of a particle is

$$\begin{aligned}\dot{x} &= [x, H_C(x,p,t)]_C \\ \dot{p} &= [p, H_C(x,p,t)]_C\end{aligned}$$

$[\quad]_C$ is the Poisson bracket in classical mechanics. In quantum mechanics,

$$\begin{aligned}\frac{d\langle X\rangle}{dt} &= -i\langle[X,H]\rangle \\ \frac{d\langle P\rangle}{dt} &= -i\langle[P,H]\rangle\end{aligned}$$

If we assume that the classical equation of motion of a particle is an approximation of quantum mechanics, we may expect

$$[\quad] = i[\quad]_C$$

Since the Poisson bracket in classical mechanics and commutation bracket in quantum mechanics have the same algebra structure. To get the right classical equation of motion of the particle, we demand that

$$[X_i, X_j] = 0 \quad [X_i, X_j] = 0 \quad [X_i, P_j] = i\delta_{ij}$$

and

$$H = H_C(X, P, t)$$

For a general system, we formally define momentum operator $\boldsymbol{P}$ by

$$[X_i, P_j] = i\delta_{ij}$$

The form of H can not be given as a priori, which can be specified only by the hints from classical theory and experiments.

15.6 Momentum operators and translation of states

Theorem 15.2

$$\exp(iG\lambda)A\exp(-iG\lambda) = A + i\lambda[G, A] + \cdots + \frac{i^n\lambda^n}{n!}[G, [G, [G, \cdots [G, A]]] \cdots] + \cdots$$

Define $T(\boldsymbol{a}) \equiv e^{-i\boldsymbol{P}\cdot\boldsymbol{a}}$ We can get

$$T(\boldsymbol{a})^{-1}\boldsymbol{X}T(\boldsymbol{a}) = \boldsymbol{X} + \boldsymbol{a}$$

$$T(\boldsymbol{a})|\boldsymbol{x}\rangle = |\boldsymbol{x} + \boldsymbol{a}\rangle$$

So, $T(\boldsymbol{a})$ is the space translation operator. Now, we can also define the momentum operator as the generator of space translation.

15.7 Angular momentum operators and rotation of states

We define the angular momentum operators $\boldsymbol{J}$ as the generator of rotation.

$$R(\boldsymbol{\theta}) \equiv e^{-i\boldsymbol{J}\cdot\boldsymbol{n}\theta}$$

If the operator $\boldsymbol{M} = (M_1, M_2, M_3)$ is a vector in configuration space and can be rotated by R, then we can demonstrate that

$$[J_i, M_j] = i\epsilon_{ijk}M_k$$

Especially,

$$[J_i, J_j] = i\epsilon_{ijk}J_k$$

Orbital angular momentum

Orbital angular momentum of a particle is defined as $\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P}$. It is the generator of rotation of the position of the particle, since

$$[L_i, X_j] = i\epsilon_{ijk}X_k \quad [L_i, P_j] = i\epsilon_{ijk}P_k \quad [L_i, L_j] = i\epsilon_{ijk}L_k$$

Spin angular momentum

Experiments show that some microscopic particles possess a property called spin. The state of the spin is denoted by $|s\rangle$. The corresponding operators are $\boldsymbol{S} = [S_1, S_2, S_3]$, which measure the spin along the $\mathbf{1}, \mathbf{2}, \mathbf{3}$ direction. Spin operator is the generator of rotation of the spin of the particle, so we have

$$[S_i, S_j] = i\epsilon_{ijk}S_k$$

And the rotation of position and spin is independent, so

$$[S_i, L_j] = 0$$

Total angular momentum

The total angular momentum of the particle is

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$$

It is the generator of the rotation of the entire system, which is equivalent to the rotation of the coordinates in opposite direction.

15.8 Heisenberg picture

Define

$$Q_H = U^\dagger(t, t_0)QU(t, t_0)$$

We can derive that

$$\frac{dQ_H(t)}{dt} = -i[Q_H(t), H_H(t)] + \left(\frac{\partial Q}{\partial t}\right)_H$$

Here, $H_H(t) = U^\dagger(t, t_0)H(t)U(t, t_0)$ If the state of the system at t_0 is $|\phi_0\rangle$, then

$$\langle Q\rangle_t = \langle\phi(t)|Q|\phi(t)\rangle = \langle\phi_0|Q_H(t)|\phi_0\rangle$$

If the state $|q\rangle$ is the eigenstate of the Q with the eigenvalue q, then $U^\dagger(t, t_0)|q\rangle$ is the eigenstate of the Q_H with eigenvalue q, which can be denoted by $|q_H(t)\rangle$, so we have

$$\langle q|\phi(t)\rangle = \langle q_H(t)|\phi_0\rangle$$

15.9 Symmetries and conservation laws

Let $U = e^{iKs}$ be a continuous unitary transformation with generator$K = K^\dagger$. To say that the Hamiltonian operator H is invariant under this transformation means that

$$U(s)^{-1}H(t)U(s) = H(t)$$

Then we can deduce that

$$[K, H(t)] = 0$$

Usually, K does not depend on time explicitly. If the above equation hold for all t, then in Heisenberg picture,

$$K_H(t) = K \quad |k_H(t)\rangle = |k\rangle$$

So,

$$\langle K\rangle_t = \langle K\rangle_{t_0} \quad \langle k|\phi(t)\rangle = \langle k|\phi_0\rangle$$

The probability distribution of the measurement of the observable K will not change with time for an arbitrary initial state. We can assume that the K is a constant of motion.

Note: *The concept of a constant of motion should not be confused with the concept of a stationary state. Suppose that the Hamiltonian operator H is independent of t, and that the initial state vector is an eigenvector of H, $|\phi_0\rangle = |E_n\rangle$ with $H|E_n\rangle = E_n|E_n\rangle$. This describes a state having a unique value of energy E_n. So*

$$|\phi(t)\rangle = e^{-iE_nt}|\phi_0\rangle$$

From this result it follows that the average of any dynamical variable R,

$$\langle\phi(t)|R|\phi(t)\rangle = \langle E_n|R|E_n\rangle$$

is independent of t for such a state. By considering functions of R we can further show that the probability distribution is independent of time. In a stationary state the averages and probabilities of all dynamical variables are independent of time, whereas a constant of motion has its average and probabilities independent of time for all states.

Chapter 16
Coordinate and Momentum Representation

16.1 Coordinate representation

To form a representation of an abstract linear vector space, one chooses a complete orthonormal set of basis vectors $\{|u_i\rangle\}$ and represents an arbitrary vector $|\psi\rangle$ by its expansion coefficients $\{c_i\}$, where $|\psi\rangle = \sum c_i|u_i\rangle$. The array of coefficients $\langle u_i|\psi\rangle$ can be regarded as a column vector (possibly of infinite dimension), provided the basis set is discrete.
Coordinate representation is obtained by choosing as the basis set the eigenvectors $\{|\boldsymbol{x}\rangle\}$ of the position operator . Since this is a continuous set, the expansion coefficients define a function of a continuous variable,

$$\psi(\boldsymbol{x}) \equiv \langle \boldsymbol{x}|\psi\rangle$$

We can show that the inner product of the state vector in coordinate representation is

$$\langle\phi|\psi\rangle = \int \phi^*(\boldsymbol{x})\psi(\boldsymbol{x})d^3\boldsymbol{x}$$

It is a matter of taste whether one says that the set of functions forms a representation of the vector space, or that the vector space consists of the functions $\psi(\boldsymbol{x})$.
The action of an operator A on the function space is related to its action on the abstract vector space by the rule

$$A\psi(\boldsymbol{x}) \equiv \langle \boldsymbol{x}|A|\psi\rangle$$

The action of an position operator in coordinate representation is

$$\boldsymbol{X}\psi(\boldsymbol{x}) = \boldsymbol{x}\psi(\boldsymbol{x})$$

The action of an momentum operator in coordinate representation is

$$\boldsymbol{P}\psi(\boldsymbol{x}) = -i\boldsymbol{\nabla}\psi(\boldsymbol{x})$$

For a spin-less particle in the scalar potential $W(\boldsymbol{x})$, $H = \frac{\boldsymbol{P}^2}{2m} + W(\boldsymbol{X})$. The equation of motion in the coordinate representation is

$$\left[-\frac{1}{2m}\boldsymbol{\nabla}^2 + W(\boldsymbol{x})\right]\psi(\boldsymbol{x},t) = i\frac{\partial}{\partial t}\psi(\boldsymbol{x},t)$$

16.2 Galilei transformation of Schrödinger equation

For simplicity we shall treat only one spatial dimension. Let us consider two frames of reference: F with coordinates x and t, and F' with coordinates x' and t'. F' is moving uniformly with velocity v relative to F, so that

$$x = x' + vt' \quad t = t'$$

The potential energy is given by $W(x,t)$ in F , and by $W'(x',t')$ in F', with

$$W(x,t) = W'(x',t')$$

Because the requirement of invariance under Galilei transformation, we expect in F' the Schrödinger equation has the form

$$\left[-\frac{1}{2m}\frac{\partial^2}{\partial x'^2} + W'(x')\right]\psi'(x',t') = i\frac{\partial}{\partial t'}\psi'(x',t')$$

where $\psi'(x',t')$ is the wave function in F'. The probability density at a point in space−time must be the same in the two frames of reference

$$|\psi(x,t)|^2 = |\psi'(x',t')|^2$$

and hence we must have

$$\psi(x,t) = e^{if}\psi'(x',t')$$

where f is a real function of the coordinates. Put all the conditions above together, we can derive

$$f(x,t) = mvx - \frac{1}{2}mv^2t$$

apart from an irrelevant constant term.

16.3 Probability flux and conditions on wave functions

Define the probability flux vector

$$\boldsymbol{J}(\boldsymbol{x},t) = \frac{1}{m}\mathrm{Im}(\psi^*\boldsymbol{\nabla}\psi)$$

We can get a continuity equation

$$\frac{d}{dt}|\psi(\boldsymbol{x},t)|^2 + \boldsymbol{\nabla}\cdot\boldsymbol{J}(\boldsymbol{x},t) = 0$$

Applying the divergence theorem, we obtain

$$\frac{\partial}{\partial t}\int_\Omega |\psi(\boldsymbol{x},t)|^2 d^3x = -\oint_\sigma \boldsymbol{J}\cdot d\boldsymbol{s}$$

The equations of continuity require that the probability flux $\boldsymbol{J}(\boldsymbol{x},t)$ be continuous across any surface, since otherwise the surface would contain sources or sinks. Although this condition applies to all surfaces, implying that $\boldsymbol{J}(\boldsymbol{x},t)$ must be everywhere continuous, its practical applications are mainly to surfaces separating regions in which the potential has different analytic forms. Usually, we have the following conditions,

1.
$$\psi(x)|_{x+0} = \psi(x)|_{x-0} \quad \frac{d\psi}{dx}|_{x+0} = \frac{d\psi}{dx}|_{x-0}$$

2.
$$\psi(x)|_{x+0} = \psi(x)|_{x-0} = 0 \quad \frac{d\psi}{dx}|_{x+0} - \frac{d\psi}{dx}|_{x-0} \text{ is finite}$$

Consider next the behavior at a singular point, assumed for convenience to be the origin of coordinates. Let S be a small sphere of radius r surrounding the singularity. The probability that the particle is inside S must be finite. Suppose that $\psi = \frac{u}{r^\alpha}$, where u is a smooth function that does not vanish at $r = 0$. Then we must have $|\psi|^2 r^3$ convergent at the origin, which implies that $\alpha < \frac{3}{2}$.
The net outward flow through the surface S is $F = \oint_S J \cdot dS$. It must vanish in the limit $r \to 0$, since otherwise the origin would be a point source or sink. One has $\frac{\partial \psi}{\partial r} = r^{-\alpha}\frac{\partial u}{\partial r} - \alpha u r^{-\alpha-1}$. The second term does not contribute to the flux, so we obtain

$$F \propto r^{2-2\alpha}$$

where the integration is over solid angle. If the integral does not vanish, then we must have $\alpha < 1$ in order for F to vanish in the limit $r \to 0$. This is a stronger condition than that derived from the probability density.
Since $|\psi|^2$ is a probability density, it must vanish sufficiently rapidly at infinity so that its integral over all configuration space is convergent and equal to 1.
The conditions that we have discussed apply to wave functions $\psi(x)$ which represent physically realizable states, but they need not apply to the eigenfunctions of operators that represent observables. Those eigenfunctions, $\chi(x)$, which play the role of filter functions in computing probabilities, are only required to lie in the extended space, Φ^X, of the rigged-Hilbert-space triplet. It has been suggested that $\psi(x)$ be restricted to the nuclear space Φ, rather than merely to the Hilbert space $\mathcal{H}$. In many cases this would amount to requiring that $\psi(x)$ should vanish at infinity more rapidly than any inverse power of the distance.

16.4 Path integrals

Theorem 16.1 Gaussian integration

$$\int dx e^{-\frac{1}{2}ax^2+Jx} = \left(\frac{2\pi}{a}\right)^{\frac{1}{2}} e^{\frac{J^2}{2a}}$$

♣

The time evolution of a quantum state vector, $|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle$, can be regarded as the propagation of an amplitude in configuration space,

$$\psi(x,t) = \int G(x,t;x',t_0)\psi(x',t_0)dx'$$

where

$$G(x,t;x',t_0) = \langle x,t|U(t,t_0)|x',t_0\rangle$$

is often called the propagator.
Making use of the multiplicative property of the time development operator, it follows that the propagator can be written as

$$G(x,t;x_0,t_0) = \int \cdots \int G(x,t;x_N,t_N) \cdots G(x_1,t_1;x_0,t_0) dx_N \cdots dx_1$$

The N-fold integration is equivalent to a sum over zigzag paths that connect the initial point (x_0,t_0) to the final point (x,t). If we now pass to the limit of $N \to \infty$ and $\Delta t = t_i - t_{i-1} \to 0$, we will have the propagator expressed as a sum (or, rather, as an integral) over all paths that connect the initial point to the final point. We can show that

$$\langle x|e^{-iH\Delta t}|x'\rangle = \sqrt{\frac{m}{2i\pi\Delta t}} \exp\left\{ i\left[\frac{m(x-x')^2}{2\Delta t^2} - V(x')\right]\Delta t\right\} \quad \Delta t \to 0$$

So,

$$G(x,t;x_0,t_0) = \lim_{N\to\infty} \int \cdots \int \left(\frac{m}{2i\pi\Delta t}\right)^{\frac{N+1}{2}} \exp\left\{ i\sum_{j=0}^{N}\left[\frac{m(x_{j+1}-x_j)^2}{2\Delta t^2} - V(x_{j+1})\right]\Delta t\right\} dx_1 \cdots dx_N$$

The final result can be expressed as

$$G(x,t;x_0,t_0) = \int \mathcal{D}[x(\tau)] e^{iS[x(\tau)]}$$

Here, $S[x(\tau)]$ is the action associated with the path

$$S[x(\tau)] = \int_{x(\tau)} L(x,\dot{x}) d\tau$$

The integral is a functional integration over all paths $x(\tau)$ which connect the initial point (x_0,t_0) to the final point (x,t).
To conclude this section, let us generalize our path-integral formula to a more complicated systems. Consider a very general quantum system, described by arbitrary set of of coordinates q_i, conjugate momentum p^i, and Hamiltonian $H(q,p)$. We can show that

$$\langle q_{k+1}|e^{-i\epsilon H}|q_k\rangle = \left(\prod_i \int \frac{dp_k^i}{2\pi}\right) \exp\left[-i\epsilon H\left(\frac{q_{k+1}+q_k}{2}, p_k\right)\right] \exp\left[i\sum_i p_k^i (q_{i,k+1} - q_{i,k})\right]$$

so,

$$\langle q_N|U(t,t_0)|q_0\rangle = \left(\prod_{i,k} \int \frac{dp_k^i dq_{i,k}}{2\pi}\right) \exp\left[i\sum_k \left(\sum_i p_k^i (q_{i,k+1} - q_{i,k}) - \epsilon H\left(\frac{q_{k+1}+q_k}{2}, p_k\right)\right)\right]$$

There is one momentum integral for each k from 0 to N, and on coordinate integral for each k from 1 to N. The final result can be expressed as

$$\langle q_N|U(t,t_0)|q_0\rangle = \left(\prod_i \int \mathcal{D}q(t)\mathcal{D}p(t)\right) \exp\left[i\int_0^T dt \left(\sum_i p^i \dot{q}_i - H(q,p)\right)\right]$$

where the functions $q(t)$ are constrained at the endpoints, but $p(t)$ are not. The details of this generalization can be found in chapter 9.1 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

16.5 Momentum representation

Momentum representation is obtained by choosing as the basis set the eigenvectors $\{|\boldsymbol{p}\rangle\}$ of the momentum operator. The orthonormality condition takes the form

$$\langle \boldsymbol{p}|\boldsymbol{p}'\rangle = \delta(\boldsymbol{p}-\boldsymbol{p}')$$

Then we can derive that

$$\langle \boldsymbol{x}|\boldsymbol{p}\rangle = \frac{1}{(2\pi)^{3/2}} e^{i\boldsymbol{p}\cdot\boldsymbol{x}}$$

and

$$\phi(\boldsymbol{p}) \equiv \langle \boldsymbol{p}|\psi\rangle = \frac{1}{(2\pi)^{3/2}} \int e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} \langle \boldsymbol{x}|\psi\rangle d^3x$$

The effect of position operator in momentum representation is

$$\boldsymbol{X}\phi(\boldsymbol{p}) = i\boldsymbol{\nabla}\phi(\boldsymbol{p})$$

Bloch's Theorem

A crystal is unchanged by translation through a vector displacement of the form

$$\boldsymbol{R}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$$

where n_1, n_2 and n_3 are integers, and $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_2$ form the edges of a unit cell of the crystal. Corresponding to such a translation, there is a unitary operator, $U(\boldsymbol{R}_n) = \exp(-i\boldsymbol{P}\cdot\boldsymbol{R}_n)$, which leaves the Hamiltonian of the crystal invariant:

$$U^{-1}(\boldsymbol{R}_n)HU(\boldsymbol{R}_n) = H$$

These unitary operators for translations commute with each other, as well as with H, so there must exist a complete set of common eigenvectors for all of these operators,

$$H|\psi\rangle = E|\psi\rangle \quad U(\boldsymbol{R}_n)|\psi\rangle = c(\boldsymbol{R}_n)|\psi\rangle$$

By the composition relation of the translation operators, we can deduce that

$$c(\boldsymbol{R}_n) = \exp(-i\boldsymbol{p}\cdot\boldsymbol{R}_n)$$

So, in coordinate representation, we have

$$\psi(\boldsymbol{x}-\boldsymbol{R}_n) = U(\boldsymbol{R}_n)\psi(\boldsymbol{x}) = \exp(-i\boldsymbol{p}\cdot\boldsymbol{R}_n)\psi(\boldsymbol{x})$$

The vector $\boldsymbol{p}$ is called the Bloch wave vector of the state. If we expand a function of the Bloch form in a series of plane waves,

$$\psi(\boldsymbol{x}) = \sum_{\boldsymbol{k}'} a(\boldsymbol{k}')e^{i\boldsymbol{k}'\cdot\boldsymbol{x}}$$

we can show that for all $\boldsymbol{R}_n$

$$\exp(i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{R}_n) = 1$$

So $\boldsymbol{G}_n \equiv \boldsymbol{k}'-\boldsymbol{k}$ is a vector of the reciprocal lattice. So, the expansion can be written as

$$\psi(\boldsymbol{x}) = \sum_{\boldsymbol{G}_m} a(\boldsymbol{k}+\boldsymbol{G}_m)e^{i(\boldsymbol{k}+\boldsymbol{G}_m)\cdot\boldsymbol{x}}$$

16.6 Harmonic oscillator

A harmonic oscillator is an object that is subject to a quadratic potential energy, which produces a restoring force against any displacement from equilibrium that is proportional to the displacement. The Hamiltonian for such an object whose motion is confined to one dimension is

$$H = \frac{1}{2m}P^2 + \frac{m\omega^2}{2}Q^2$$

where P is the momentum, Q is the position, and m is the mass.

16.6.1 Algebraic solution

We have the commutation relation

$$[Q, P] = i$$

and the self-adjointness of the operators P and Q,

$$P = P^\dagger \quad Q = Q^\dagger$$

Define

$$p \equiv (m\omega)^{-1/2}P \quad q \equiv (m\omega)^{1/2}Q$$

So

$$[q, p] = i \quad H = \frac{1}{2}\omega(p^2 + q^2)$$

We further define

$$a \equiv \frac{q + ip}{\sqrt{2}}$$

We can verify that

$$[a, a^\dagger] = 1 \quad H = \frac{1}{2}\omega(aa^\dagger + a^\dagger a) = \omega(aa^\dagger - \frac{1}{2}) = \omega(a^\dagger a + \frac{1}{2})$$

Introduce $N \equiv a^\dagger a$, we have

$$[N, a] = -a \quad [N, a^\dagger = a^\dagger]$$

Let $N|\nu\rangle = \nu|\nu\rangle$, with $\langle\nu|\nu\rangle \neq 1$. We have

$$Na|\nu\rangle = a(N-1)|\nu\rangle = (\nu - 1)a|\nu\rangle$$

If $\nu \neq 0$, then $a|\nu\rangle$ cannot be 0, $a|\nu\rangle$ must be an eigenvector of N with eigenvalue $\nu - 1$. On the other hand, we have

$$(\langle\nu|a^\dagger)(a|\nu\rangle) = \langle\nu|N|\nu\rangle = \nu\langle\nu|\nu\rangle$$

Since the norm must be nonnegative, it follows that $\nu \geq 0$, and thus an eigenvalue cannot be negative. By applying the operator a repeatedly, it would appear that one could construct an indefinitely long sequence of eigenvectors having the eigenvalues $\nu - 1$, $\nu - 2$, $\nu - 3$ and so on. The contradiction can be avoided only if the sequence terminates with the value $\nu = 0$

and $a|0\rangle = 0$.
We also have that

$$Na^\dagger|\nu\rangle = a^\dagger(N+1)|\nu\rangle = (\nu+1)a^\dagger|\nu\rangle$$

The squared norm of the vector $a^\dagger|\nu\rangle$ is

$$(\langle\nu|a)(a^\dagger|\nu\rangle) = \langle\nu|N+1|\nu\rangle = (\nu+1)\langle\nu|\nu\rangle$$

which never vanishes because $\nu \geq 0$. Thus $a^\dagger|\nu\rangle$ is an eigenvector of N with eigenvalue $\nu+1$. By repeatedly applying the operator $a^\dagger$, one can construct an unlimited sequence of eigenvectors, each having an eigenvalue one unit greater than that of its predecessor. The sequence begins with the eigenvalue $\nu = 0$. Therefore the spectrum of N consists of the nonnegative integers, $\nu = n$.
The orthonormal eigenvectors of N will be denoted as $|n\rangle$, and we can verify that

$$|n\rangle = n^{-1/2}a^\dagger|n-1\rangle = (n!)^{-1/2}(a^\dagger)^n|0\rangle$$

The matrix elements of $a^\dagger$ and a are

$$\langle n'|a^\dagger|n\rangle = (n+1)^{1/2}\delta_{n',n+1} \quad \langle n'|a|n\rangle = (n)^{1/2}\delta_{n',n-1}$$

Finally we note that the eigenvalues and eigenvectors of the harmonic oscillator Hamiltonian are

$$H|n\rangle = E_n|n\rangle \quad E_n = (n+\frac{1}{2})\omega$$

16.6.2 Solution in coordinate representation

In the coordinate representation, we have

$$-\frac{1}{2m}\frac{d^2}{dx^2}\psi(x) + \frac{m\omega^2}{2}x^2\psi(x) = E\psi(x)$$

Define

$$q \equiv (m\omega)^{1/2}x \quad \lambda \equiv \frac{2E}{\omega}$$

The differential equation becomes

$$\frac{d^2u}{dq^2} + (\lambda - q^2)u = 0$$

When $q \to \pm\infty$, we have $u(q) \sim e^{q^2/2}$ or $e^{-q^2/2}$. The first of these is unacceptable, because it diverges so severely as to be outside of both Hilbert space and rigged Hilbert space. We would seek solutions of the form

$$u(q) = H(q)e^{-\frac{1}{2}q^2}$$

We then have

$$H'' - 2qH' + (\lambda - 1)H = 0$$

It is the so-called Hermite differential equation. When $\lambda - 1 = 2n$, we have regular solutions. Solutions are Hermite polynomials, and will be denoted as $H_n(q)$. After appropriate normalization, we have

$$\psi_n(x) = \left[\frac{\alpha}{\pi^{1/2}2^n n!}\right]^{1/2} H_n(\alpha x)e^{-\frac{1}{2}\alpha^2x^2} \quad E_n = (n+\frac{1}{2})\omega$$

where $\alpha \equiv (m\omega)^{1/2}$.

16.6.3 Path integral solution

The propagator of harmonic oscillator in terms of path integral are

$$\begin{aligned} G(x_b,t_b;x_a,t_a) &= \int \mathcal{D}[x(t)]e^{i\int_{t_a}^{t_b}(\frac{1}{2}m\dot{x}^2-\frac{1}{2}m\omega^2x^2)dt} \\ &= \lim_{N\to\infty}\int\left(\frac{m}{2i\pi\Delta t}\right)^{\frac{N+1}{2}}\exp\left\{i\sum_{j=0}^{N}\left[\frac{m(x_{j+1}-x_j)^2}{2\Delta t^2}-\frac{1}{2}m\omega^2x_{j+1}^2\right]\Delta t\right\}dx_1\cdots dx_N \end{aligned}$$

where

$$x_0 = x_a \quad x_{N+1} = x_b \quad \Delta t = \frac{t_b - t_a}{N+1}$$

Suppose $x_c(t)$ is the classical path of harmonic oscillator and we define

$$\delta x(t) \equiv x(t) - x_c(t)$$

Substitute it into the equation above, the terms which is linear in δx can be dropped, because

$$\frac{\delta S}{\delta x}|_{x(t)=x_c(t)}$$

And since $x_c(t_a) = x_a$, $x_c(t_b) = x_b$, we have

$$\delta x_0 = \delta x_{N+1} = 0$$

At last, we have

$$\begin{aligned} G(x_b,t_b;x_a,t_a) &= e^{iS_c}\lim_{N\to\infty}\int\left(\frac{m}{2i\pi\Delta t}\right)^{\frac{N+1}{2}}\exp\left\{i\sum_{j=1,k=1}^{N}\delta x_jS_{jk}\delta x_k\right\}d\delta x_1\cdots d\delta x_N \\ &= e^{iS_c}\lim_{N\to\infty}\left(\frac{m}{2i\pi\Delta t}\right)^{\frac{N+1}{2}}\sqrt{\frac{\pi^N}{\det(-iS)}} \end{aligned} \tag{16.1}$$

where

$$-iS = \frac{m}{2i\Delta t}\begin{pmatrix} 2-\omega^2\Delta t^2 & -1 & 0 & \cdots \\ -1 & 2-\omega^2\Delta t^2 & -1 & \cdots \\ 0 & -1 & 2-\omega^2\Delta t^2 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

The details of the following calculation will be eliminated here, and some discussion can be found in section 2.1.4 of *Quantum Field Theory of Many-body Systems (Xiao-Gang Wen)*. The final result is

$$G(x_b,t;x_a,0) = \left(\frac{m\omega}{2\pi i\sin\omega t}\right)^{1/2}\exp\left[\frac{im\omega}{2\sin\omega t}\left((x_a^2+x_b^2)\cos\omega t - 2x_ax_b\right)\right]$$

In the limit $\omega\to 0$, we have

$$G(x_b,t;x_a,0) = \left(\frac{m}{2\pi it}\right)^{1/2}\exp\left[\frac{im}{2t}(x_b-x_a)^2\right]$$

This is just the propagator of free particle.

16.7 Quantum-mechanical particles in classical electromagnetic field

16.7.1 General discussion

In classical electrodynamics, if the velocity of the charged particle is much smaller than that of light, then the Hamiltonian of charged particles in a given EM field is

$$H = \frac{(\boldsymbol{\pi} - e\boldsymbol{A})^2}{2m} + e\phi$$

In corresponding quantum theory, we will suppose the Hamiltonian operator to be

$$H = \frac{[\boldsymbol{P} - e\boldsymbol{A}(\boldsymbol{X})]^2}{2M} + e\phi(\boldsymbol{X})$$

The corresponding field operators are

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t} \quad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$$

In Heisenberg picture, we have

$$\frac{d\boldsymbol{X}}{dt} = -i[\boldsymbol{X}, H] = \frac{1}{M}(\boldsymbol{P} - e\boldsymbol{A})$$

Define kinetic momentum $\boldsymbol{K}$ by

$$\boldsymbol{K} = \boldsymbol{P} - e\boldsymbol{A}$$

We have

$$[K_i, K_j] = ie(\partial_i A_j - \partial_j A_i) = ie\epsilon_{ijk}B_k$$

So

$$M\frac{d^2\boldsymbol{X}}{dt^2} = -i[\boldsymbol{K}, H] + \frac{\partial \boldsymbol{K}}{\partial t} = e\left[\boldsymbol{E} + \frac{1}{2}\left(\frac{d\boldsymbol{X}}{dt} \times \boldsymbol{B} - \boldsymbol{B} \times \frac{d\boldsymbol{X}}{dt}\right)\right]$$

In coordinate representation of Schrödinger picture, we have

$$\frac{1}{2M}\left[-i\boldsymbol{\nabla} - e\boldsymbol{A}\right] \cdot \left[-i\boldsymbol{\nabla} - e\boldsymbol{A}\right]\psi(\boldsymbol{x}, t) + e\phi(\boldsymbol{x})\psi(\boldsymbol{x}, t) = i\frac{\partial\psi(\boldsymbol{x}, t)}{\partial t}$$

Define probability current $\boldsymbol{j}$ as

$$\boldsymbol{j} \equiv \frac{1}{M}\mathrm{Im}(\psi^*\boldsymbol{\nabla}\psi) - \frac{e}{M}\boldsymbol{A}|\psi|^2$$

we can verify that

$$\boldsymbol{\nabla} \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t} = 0$$

Transformation

$$\phi \to \phi - \frac{\partial \Lambda}{\partial t} \quad \boldsymbol{A} \to \boldsymbol{A} + \boldsymbol{\nabla}\Lambda$$

will leave $\boldsymbol{E}$ and $\boldsymbol{B}$ unchanged. The transformation is called gauged transformation. In classical electrodynamics, gauge transformation will not change the trajectory of particles, (which is the only thing we can observed in experiment). In quantum theory, suppose the state vector $|\psi\rangle$ will transform as

$$|\psi(t)\rangle \to O(t)|\psi(t)\rangle$$

where $O(t)$ is an unitary operator. If the Schrödinger equation is always satisfied, we can derive that

$$H'O - OH = i\frac{\partial O}{\partial t}$$

where H' is the Hamiltonian operator after gauge transformation. Generally, we have

$$O(t) = \exp\left[ie\Lambda(\boldsymbol{X}, t)\right]$$

And so

$$O^{-1}\boldsymbol{X}O = \boldsymbol{X} \quad O^{-1}\boldsymbol{P}O = \boldsymbol{P} + e\boldsymbol{\nabla}\Lambda \quad O^{-1}(\boldsymbol{P} - e\boldsymbol{A}')O = \boldsymbol{P} - e\boldsymbol{A}$$

So, the expectation value of $\boldsymbol{X}$ and $\boldsymbol{K}$ is invariant under gauge transformation. We can also verify that $\boldsymbol{j}$ is also invariant under gauge transformation. A special case is that

$$\phi \to \phi + \phi_0(t) \quad \boldsymbol{A} \to \boldsymbol{A}$$

In this case, we have

$$O(t) = \exp\left[-i\int_{t_0}^{t} dt' e\phi_0(t')\right]$$

If ϕ_0 is a constant, then

$$O(t) = \exp\left[-ie\phi_0(t - t_0)\right]$$

16.7.2 Motion in a uniform static magnetic field

Suppose that the magnetic field be of magnitude B in the z direction. The Hamiltonian $H = H_{xy} + H_z$ with $H_{xy} = \frac{K_x^2 + K_y^2}{2m}$ and $H_z = \frac{P_z^2}{2m}$. Since $B_x = B_y = 0$, then P_z commutes with P_x and P_y. Hence the operators H_{xy} and H_z are commutative, and every eigenvalue of H is just the sum of an eigenvalue of H_{xy} and an eigenvalue of H_z.
Define

$$Q' \equiv \frac{K_x}{\gamma} \quad P' \equiv \frac{K_y}{\gamma} \quad \gamma \equiv \sqrt{|eB|}$$

Then we have

$$H_{xy} = \frac{1}{2}\frac{|eB|}{m}(Q'^2 + P'^2) \text{ with } [Q', P'] = i \text{ or } -i$$

therefore the eigenvalues of H_{xy} must be equal to $(n + \frac{1}{2})\frac{|eB|}{m}$, where n is any non-negative integer.
The spectrum of K_z can be shown to be gauge-invariant. Because the magnetic field is uniform and in the z direction, it is possible to choose the vector potential such that $A_z = 0$. Therefore

the spectrum of K_z is continuous from $-\infty$ to ∞, like that of P_z.
Thus the energy eigenvalues for a charged particle in a uniform static magnetic field B are

$$E_n(p_z) = \frac{(n+\frac{1}{2})|eB|}{m} + \frac{p_z^2}{2m}$$

The motion parallel to the magnetic field is not coupled to the transverse motion, and is unaffected by the field. The classical motion in the plane perpendicular to the field is in a circular orbit with angular frequency $\omega_c = \frac{eB}{m}$, and it is well known that periodic motions correspond to discrete energy levels whose separation is ω_c.

Now let us choose the vector potential to be $A_x = -yB$,$A_y = A_z = 0$. The Hamiltonian now becomes

$$H = \frac{(P_x + yeB)^2 + P_y^2 + P_z^2}{2m}$$

P_x and P_z commute with H, so it is possible to construct a complete set of common eigenvectors of H, P_x and P_z. In coordinate representation, the eigenvalue equation now takes the form

$$-\frac{1}{2m}\nabla^2\psi - \frac{ieB}{m}y\frac{\partial}{\partial x}\psi + \frac{e^2B^2}{2m}y^2\psi = E\psi$$

Substitute

$$\psi(x,y,z) = \exp(ik_x x + ik_z z)\phi(y)$$

The equation then takes the form

$$-\frac{1}{2m}\frac{d^2\phi(y)}{dy^2} + \left[\frac{m\omega_c^2}{2}(y-y_0)^2 - E'\right]\phi(y) = 0$$

where $\omega_c = \frac{eB}{m}$ is the classical cyclotron frequency, and $E' = E - \frac{k_z^2}{2m}$ is the energy associated with motion in the xy plane. This is just the energy eigenvalue equation for a simple harmonic oscillator with angular frequency $\omega = |\omega_c|$, whose eigenvalues are $E' = (n+1/2)\omega$. Thus the energies for the charged particle in the magnetic field must be $E = (n+1/2)|\omega_c| + \frac{k_z^2}{2m}$. Apart from a normalization constant, the eigenfunction will be

$$\psi = \exp(ik_x x + ik_z z)H_n[\alpha(y-y_0)]\exp[-\frac{1}{2}\alpha^2(y-y_0)^2]$$

with $\alpha = \sqrt{m\omega} = \sqrt{|eB|}$, and $y_0 = -\frac{k_x}{eB}$.
For fixed n and k_z, the energy eigenvalue is highly degenerate. For convenience, we assume that the system is confined to a rectangle of dimension $D_x \times D_y$ and subject to periodic boundary conditions. The allowed values of k_x are $k_x = \frac{2\pi n_x}{D_x}$, with $n_x = 0, \pm 1, \cdots$.The orbit center coordinate $y_0 = -\frac{2\pi n_x}{D_x eB}$ must lie in the range $[0, D_y]$. In the limit as D_x and D_y become large, we may ignore problems associated with orbits lying near the boundary, since they will be a negligible fraction of the total. In this limit the number of degenerate states corresponding to fixed n and k_z will be $\frac{D_x D_y |eB|}{2\pi}$.

16.7.3 The Aharonov-Bohm effect

A long solenoid is placed perpendicular to the plane of the figure, so that a magnetic field can be created inside the solenoid while the region external to the solenoid remains field-free. The solenoid is located in the unilluminated shadow region so that no particles will reach it, and moreover it may be surrounded by a cylindrical shield that is impenetrable to the charged particles. Nevertheless it can be shown that the interference pattern depends upon the magnetic flux through the cylinder.

Figure 16.1: The Aharonov–Bohm experiment

Let $\Psi^{(0)}(\boldsymbol{x}, t)$ be the solution of the Schrödinger equation and boundary conditions of this problem for the case in which the vector potential is everywhere zero. Now let us consider the case in which the magnetic field is non-zero inside the cylinder but zero outside of it. The vector potential $\boldsymbol{A}$ will not vanish everywhere in the exterior region, even though $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} = 0$ outside of the cylinder. This follows by applying Stokes's theorem to any path surrounding the cylinder

$$\oint \boldsymbol{A} \cdot d\boldsymbol{x} = \int\int (\boldsymbol{\nabla} \times \boldsymbol{A}) \cdot d\boldsymbol{S} = \int\int \boldsymbol{B} \cdot d\boldsymbol{S} = \Phi$$

If the flux through the cylinder is not zero, then the vector potential must be nonzero on every path that encloses the cylinder. However in any simply connected region outside of the cylinder, it is possible to express the vector potential as the gradient of a scalar, from the zero potential solution by means of a gauge transformation, $\Psi = \Psi^{(0)} \exp(ie\Lambda)$.
In region L, which contains the slit on the left, the wave function can be written as $\Psi_L = \Psi_L \exp(ie\Lambda_1)$, where Ψ_L is the zero potential solution in region L, and $\Lambda_1(\boldsymbol{x}, t) = \int \boldsymbol{A} \cdot d\boldsymbol{x}$, with the integral taken along a path within region L. A similar form can be written for the wave function in the region R, which contains the slit on the right.
At the point b, in the overlap of regions L and R, the wave function is a superposition of contributions from both slits. Hence we have

$$\Psi(\boldsymbol{b}) = \Psi_L \exp(ie\Lambda_1) + \Psi_R \exp(ie\Lambda_2)$$

The interference pattern depends $\exp(ie(\Lambda 1 - \Lambda_2)) = \exp(ie\Phi)$. Therefore the interference pattern is sensitive to the magnetic flux inside of the cylinder, even though the particles never pass through the region in which the magnetic field is nonzero. The AB effect is a topological effect, in that the effect depends on the flux encircled by the paths available to the particle,

even though the paths may never approach the region of the flux.
At last, we conclude that in quantum theory, the potentials themselves are physically significant; however, they are subject to the requirement that all observable effects be invariant under gauge transformations.

Chapter 17
Angular Momentum

17.1 Eigenvalues of angular momentum operator

The commutation relations among the angular momentum operators are

$$[J_i, J_j] = \epsilon_{ijk} J_k$$

And these three operators are self-adjoint. We first introduce the operator $J^2 = J_x^2 + J_y^2 + J_z^2$. We can verify that $[J^2, \boldsymbol{J}] = 0$. Thus there exists a complete set of common eigenvectors of J^2 and any one component of $\boldsymbol{J}$. Particularly, we have the pair of eigenvalue equations

$$J^2|\beta, m\rangle = \beta|\beta, m\rangle \quad J_z|\beta, m\rangle = m|\beta, m\rangle$$

Since

$$\langle\beta, m|J^2|\beta, m\rangle = \langle\beta, m|J_x^2|\beta, m\rangle + \langle\beta, m|J_y^2|\beta, m\rangle + \langle\beta, m|J_z^2|\beta, m\rangle$$

we have $m^2 \leq \beta$. Thus for a fixed value of β there must be maximum and minimum values for m.
Define

$$J_+ \equiv J_x + iJ_y \quad J_- \equiv J_x - iJ_y$$

we have the commutation relations

$$[J_z, J_+] = J_+ \quad [J_z, J_-] = -J_- \quad [J_+, J_-] = 2J_z$$

So

$$J_z J_+|\beta, m\rangle = J_+(J_z + 1)|\beta, m\rangle = (m+1)J_+|\beta, m\rangle$$

Therefore, either $J_+|\beta, m\rangle$ is an eigenvector of J_z with the raised eigenvalue $m+1$, or $J_+|\beta, m\rangle = 0$. Now for fixed β there is a maximum value of m, which we shall denote as j. It must be the case that

$$J_+|\beta, j\rangle = 0$$

Since

$$J_- J_+ = J^2 - J_z^2 - J_z$$

it is obvious that $\beta = j(j+1)$. By similar method, we can show the minimum eigenvalue of J_z for fixed β satisfy that $\beta = k(k-1)$. So, we have $k = -j$.
We have thus shown the existence of a set of eigenvectors corresponding to integer spaced m values in the range $-j \leq m \leq j$. Since the difference between the maximum value j and the minimum value $-j$ must be an integer, it follows that $j =$ integer$/2$. Henceforth we shall

adopt the common and more convenient notation of labeling the eigenvectors by j instead of by β. Thus the vector that was previously denoted as $|\beta, m\rangle$ will now be denoted as $|j, m\rangle$.

To find the matrix element of angular momentum operator, we notice that

$$\langle j, m|J_- J_+|j, m\rangle = j(j+1) - m(m+1)$$

so, we can get

$$J_+|j, m\rangle = \sqrt{(j+m+1)(j-m)}|j, m+1\rangle$$

Similarly, we have

$$J_-|j, m\rangle = \sqrt{(j-m+1)(j+m)}|j, m-1\rangle$$

The matrix element of J_+, J_- and J_z are

$$\langle j', m'|J_+|j, m\rangle = \sqrt{(j+m+1)(j-m)}\delta_{jj'}\delta_{m',m+1}$$

$$\langle j', m'|J_-|j, m\rangle = \sqrt{(j-m+1)(j+m)}\delta_{jj'}\delta_{m',m-1}$$

$$\langle j', m'|J_z|j, m\rangle = m\delta_{jj'}\delta_{m',m}$$

17.2 Orbital Angular Momentum and Spin

Let $\psi(\boldsymbol{x})$ be a one-component state function in coordinate representation. When it is subjected to a rotation it is transformed into

$$\boldsymbol{R}\psi(\boldsymbol{x}) = \psi(R^{-1}\boldsymbol{x})$$

where $\boldsymbol{R}$ is the rotation operator generated by $\boldsymbol{R} = \exp(-i\boldsymbol{J}\cdot\boldsymbol{n}\theta)$. For a rotation through infinitesimal angle ϵ about the z axis, we have

$$\boldsymbol{R}_z(\epsilon)\psi(x, y, z) = \psi(x+\epsilon y, y-\epsilon x, z) = \psi(x, y, x) + \epsilon(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y})$$

On the other hand,

$$\boldsymbol{R}_z(\epsilon) = I - i\epsilon J_z$$

so, we have $J_z = -i(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})$. This is just the z component of the orbital angular momentum operator $L = \boldsymbol{X}\times\boldsymbol{P}$.
For a multicomponent state function, we have

$$\boldsymbol{R}\begin{pmatrix}\psi_1(\boldsymbol{x})\\ \psi_2(\boldsymbol{x})\\ \vdots\end{pmatrix} = D\begin{pmatrix}\psi_1(R^{-1}\boldsymbol{x})\\ \psi_2(R^{-1}\boldsymbol{x})\\ \vdots\end{pmatrix}$$

Thus the general form of the rotation operator will be

$$\boldsymbol{R}_n(\theta) = e^{-i\boldsymbol{L}\cdot\boldsymbol{n}\theta}D_n(\theta)$$

The two factors commute because the first acts only on the coordinate and the second acts only on the components of the column vector. The matrix D must be unitary, and so it can be written as

$$D_n(\theta) = e^{-i\boldsymbol{S}\cdot\boldsymbol{n}\theta}$$

The angular momentum operator $\boldsymbol{J}$ has the form

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$$

with $\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P}$ and $[L_\alpha, S_\beta] = 0$. In the particular representation used in this section, we have $\boldsymbol{L} = -i\boldsymbol{x} \times \boldsymbol{\nabla}$, and the components of S are discrete matrices. The operators $\boldsymbol{L}$ and $\boldsymbol{S}$ are called the orbital and spin parts of the angular momentum.

Orbital angular momentum

The form of the gradient operator in spherical coordinates is

$$\boldsymbol{\nabla} = \boldsymbol{e}_r \frac{\partial}{\partial r} + \boldsymbol{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \boldsymbol{e}_\phi \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}$$

The orbital angular momentum operator then has the form

$$\boldsymbol{L} = r\boldsymbol{e}_r \times (-i\boldsymbol{\nabla}) = (-i)\left[\boldsymbol{e}_\phi \frac{\partial}{\partial \theta} - \boldsymbol{e}_\theta \frac{1}{\sin\theta}\frac{\partial}{\partial \phi}\right]$$

So,we have

$$L_z = \boldsymbol{L}\cdot\boldsymbol{e}_z = -i\frac{\partial}{\partial \phi}$$

$$L^2 = \boldsymbol{L}\cdot\boldsymbol{L} = -\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$$

We must now solve the two coupled differential equations,

$$L_z Y(\theta,\phi) = mY(\theta,\phi) \quad L^2 Y(\theta,\phi) = l(l+1)Y(\theta,\phi)$$

Apart from normalization, we have $Y(\theta,\phi) = e^{im\phi}P_l^m(\cos\theta)$. Here, P_l^m is the associated Legendre polynomials. If we assume that the solution must be single-valued under rotation, then it will follow that m must be an integer. If we further assume that it must be nonsingular at $\theta = 0$ and $\theta = \pi$, then from the standard theory of the Legendre equation it will follow that l must be a nonnegative integer in the range $l \geq |m|$. The normalized solutions that result from these assumptions are the well-known spherical harmonics

$$Y_l^m(\theta,\phi) = (-1)^{(m+|m|)/2}\left[\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}\right]^{1/2} e^{im\phi}P_l^{|m|}(\cos\theta)$$

Spin

A particular species of particle is characterized by a set of quantum numbers that includes the value of its spin s, it is often sufficient to treat the spin operators $\boldsymbol{S}$ as acting on the space of

dimension $2s+1$ that is spanned by the eigenvectors of for a fixed value of s.
If $s=\frac{1}{2}$, we have

$$S_x=\frac{1}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\quad S_y=\frac{1}{2}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\quad S_z=\frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

The spin operator in direction $\boldsymbol{n}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ is

$$\boldsymbol{S}_n=\frac{1}{2}\begin{bmatrix}\cos\theta & e^{-i\phi}\sin\theta\\ e^{i\phi}\sin\theta & -\cos\theta\end{bmatrix}$$

The eigenvectors are

$$\begin{bmatrix}e^{-i\phi/2}\cos\frac{\theta}{2}\\ e^{i\phi/2}\sin\frac{\theta}{2}\end{bmatrix}\quad\begin{bmatrix}-e^{-i\phi/2}\sin\frac{\theta}{2}\\ e^{i\phi/2}\cos\frac{\theta}{2}\end{bmatrix}$$

corresponding to eigenvalues $\frac{1}{2}$ and $-\frac{1}{2}$.
If $s=1$, we have

$$S_x=\sqrt{\frac{1}{2}}\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}\quad S_y=\sqrt{\frac{1}{2}}\begin{bmatrix}0 & -i & 0\\ i & 0 & -i\\ 0 & i & 0\end{bmatrix}\quad S_z=\sqrt{\frac{1}{2}}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{bmatrix}$$

The spin operator in direction $\boldsymbol{n}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ is

$$\boldsymbol{S}_n=\begin{bmatrix}\cos\theta & \sin\theta e^{-i\phi}\sqrt{\frac{1}{2}} & 0\\ \sin\theta e^{i\phi}\sqrt{\frac{1}{2}} & 0 & \sin\theta e^{-i\phi}\sqrt{\frac{1}{2}}\\ 0 & \sin\theta e^{i\phi}\sqrt{\frac{1}{2}} & -\cos\theta\end{bmatrix}$$

The eigenvectors are

$$\begin{bmatrix}\frac{1}{2}(1+\cos\theta)e^{-i\phi}\\ \sqrt{\frac{1}{2}}\sin\theta\\ \frac{1}{2}(1-\cos\theta)e^{i\phi}\end{bmatrix}\quad\begin{bmatrix}-\sqrt{\frac{1}{2}}\sin\theta e^{-i\phi}\\ \cos\theta\\ \sqrt{\frac{1}{2}}\sin\theta e^{i\phi}\end{bmatrix}\quad\begin{bmatrix}\frac{1}{2}(1-\cos\theta)e^{-i\phi}\\ -\sqrt{\frac{1}{2}}\sin\theta\\ \frac{1}{2}(1+\cos\theta)e^{i\phi}\end{bmatrix}$$

corresponding to eigenvalues 1, 0 and -1.

17.3 Rotation operator

Three parameters are required to describe an arbitrary rotation. A common parameterization is by the Euler angles. From the fixed system of axes $Oxyz$, a new rotated set of axes $Ox'y'z'$ is produced in three steps:

- Rotate through angle α about Oz, carrying Oy into Ou
- Rotate through angle β about Ou, carrying Oz into Oz'
- Rotate through angle γ about Oz' , carrying Ou into Oy'

Figure 17.1: Euler angles

The net rotation is

$$\boldsymbol{R}(\alpha,\beta,\gamma) = \boldsymbol{R}_{z'}(\gamma)\boldsymbol{R}_u(\beta)\boldsymbol{R}_z(\alpha) = e^{-i\gamma J_{z'}}e^{-i\beta J_u}e^{-i\alpha J_z}$$

Since $J_u = \boldsymbol{R}_z(\alpha)J_y\boldsymbol{R}_z(-\alpha)$, we have $\boldsymbol{R}_u(\beta) = \boldsymbol{R}_z(\alpha)\boldsymbol{R}_y(\beta)\boldsymbol{R}_z(-\alpha)$. Similarly, we can obtain $\boldsymbol{R}_{z'}(\gamma) = \boldsymbol{R}_u(\beta)\boldsymbol{R}_z(\gamma)\boldsymbol{R}_u(-\beta)$. So, the rotation operator is

$$\boldsymbol{R}(\alpha,\beta,\gamma) = \boldsymbol{R}_z(\alpha)\boldsymbol{R}_y(\beta)\boldsymbol{R}_z(\gamma) = e^{-i\alpha J_z}e^{-i\beta J_y}e^{-i\gamma J_z}$$

The matrix representation of the rotation operator in the basis $|j,m\rangle$

$$\langle j',m'|\boldsymbol{R}(\alpha,\beta,\gamma)|j,m\rangle = \delta_{jj'}D^{(j)}_{m'm}(\alpha,\beta,\gamma)$$

gives rise to the rotation matrices,

$$D^{(j)}_{m'm}(\alpha,\beta,\gamma) = \langle j',m'|e^{-i\alpha J_z}e^{-i\beta J_y}e^{-i\gamma J_z}|j,m\rangle = e^{-i(\alpha m'+\gamma m)}d^{(j)}_{mm'}(\beta)$$

where

$$d^{(j)}_{mm'}(\beta) = \langle j',m'|e^{-i\beta J_y}|j,m\rangle$$

For the case of $j=\frac{1}{2}$, we have $J_y = \frac{1}{2}\sigma_y$ and $\sigma_y^2 = I$. We can obtain

$$d^{(1/2)}(\beta) = \begin{bmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \cos\frac{\beta}{2} & \sin\frac{\beta}{2} \end{bmatrix}$$

Notice that this matrix is periodic in β with period 4π, but it changes sign when 2π is added to β. This double-valuedness under rotation by 2π is a characteristic of the full rotation matrix whenever j is a half odd-integer. The matrix is single-valued under rotation by 2π whenever j is an integer.
Rotation of angular momentum eigenvectors now can be written as

$$\boldsymbol{R}(\alpha,\beta,\gamma)|j,m\rangle = \sum_{m'} D^{(j)}_{m'm}(\alpha,\beta,\gamma)|j,m'\rangle$$

When it comes to spherical harmonics, we have

$$Y_l^m(\theta',\phi') = \boldsymbol{R}^{-1}((\alpha,\beta,\gamma))Y_l^m(\theta,\phi) = \sum_{m'} Y_l^{m'}(\theta,\phi)[D^{(j)}_{mm'}((\alpha,\beta,\gamma))]^*$$

By putting $\beta = \gamma = 0$ we obtain

$$Y_l^m(\theta, \phi + \alpha) = \sum_{m'} Y_l^{m'}(\theta, \phi)[D^{(j)}_{mm'}((\alpha, 0, 0))]^* = e^{i\alpha m} Y_l^m(\theta, \phi)$$

Setting $\phi = 0$ then yields

$$Y_l^m(\theta, \alpha) = e^{i\alpha m} Y_l^m(\theta, 0)$$

Since the direction $\theta = 0$ is the polar axis, continuity of the spherical harmonic requires that $Y_l^m(0, \alpha)$ be independent of α. Therefore we must have $Y_l^m(0, 0) = 0$ for $m \neq 0$, and so we can write

$$Y_l^m(\theta, 0) = c_l \delta m0$$

The we have

$$Y_l^m(\theta, \phi) = \sum_{m'} Y_l^{m'}(0, 0)[D^{(j)}_{mm'}((\phi, \theta, \gamma))]^* = c_l [D^{(j)}_{m0}((\phi, \theta, \gamma))]^*$$

for arbitrary γ, thus obtaining a simple relation between the spherical harmonics and the rotation matrices. Conventional normalization is obtained if we put

$$c_l = \left(\frac{2l + 1}{4\pi}\right)^{1/2}$$

The operator for a rotation through 2π about an axis along the unit vector $\boldsymbol{n}$ is $\boldsymbol{R}_n(2\pi) = e^{-2\pi i \boldsymbol{n}\cdot\boldsymbol{J}}$. Its effect on the standard angular momentum eigenvectors is

$$\boldsymbol{R}_n(2\pi) = (-1)^{2j}|j, m\rangle$$

We assume a rotation through 2π as a trivial operation that leaves everything unchanged, i.e. all dynamical variables are invariant under 2π rotation:

$$\boldsymbol{R}(2\pi) A \boldsymbol{R}^{-1}(2\pi) = A$$

where A may represent any physical observable.
the operator $\boldsymbol{R}_{2\pi}$ divides the vector space into two subspaces. A typical vector in the first subspace, denoted as $|+\rangle$, has the property $\boldsymbol{R}(2\pi)|+\rangle = |+\rangle$, whereas a typical vector in the second subspace, denoted as $|-\rangle$, has the property $\boldsymbol{R}(2\pi)|-\rangle = -|-\rangle$. Now, if A represents any physical observable, we have $\langle +|\boldsymbol{R}(2\pi)A|-\rangle = \langle +|A\boldsymbol{R}(2\pi)|-\rangle$, leading to

$$\langle +|A|-\rangle = 0$$

No physical observable can have nonvanishing matrix elements between states with integer angular momentum and states with half odd-integer angular momentum. This fact forms the basis of a superselection rule: There is no observable distinction among the state vectors of the form

$$|\Psi_\omega\rangle = |+\rangle + e^{i\omega}|-\rangle$$

for different values of the phase ω.

17.4 Addition of angular momentum

Let us consider a two-component system, each component of which has angular momentum degrees of freedom. Basis vectors for the composite system can be formed from the basis vectors of the components by taking all binary products of a vector from each set

$$|j_1, j_2, m_1, m_2\rangle = |j_1, m_1\rangle^{(1)}|j_2, m_2\rangle^{(2)}$$

These vectors are common eigenvectors of the four commutative operators $\boldsymbol{J}^{(1)} \cdot \boldsymbol{J}^{(1)}$, $\boldsymbol{J}^{(2)} \cdot \boldsymbol{J}^{(2)}$, $\boldsymbol{J}_z^{(1)}$, and $\boldsymbol{J}_z^{(2)}$. It is often desirable to form eigenvectors of the total angular momentum operators, $\boldsymbol{J} \cdot \boldsymbol{J}$ and $\boldsymbol{J}_z$, where the total angular momentum vector operator is

$$\boldsymbol{J} = \boldsymbol{J}^{(1)} \otimes \mathbf{1} + \mathbf{1} \otimes \boldsymbol{J}^{(2)}$$

This is useful when the system is invariant under rotation as a whole, but not under rotation of the two components separately. The eigenvectors of $\boldsymbol{J} \cdot \boldsymbol{J}$ and $\boldsymbol{J}_z$ may be denoted as $|\alpha, J, M\rangle$. It is easy to verify that the four operators $\boldsymbol{J}^{(1)} \cdot \boldsymbol{J}^{(1)}$, $\boldsymbol{J}^{(2)} \cdot \boldsymbol{J}^{(2)}$, $\boldsymbol{J} \cdot \boldsymbol{J}$ and $\boldsymbol{J}_z$ are mutually commutative, and hence they possess a complete set of common eigenvectors. Since the set of product vectors and the new set of total angular momentum eigenvectors are both eigenvectors of $\boldsymbol{J}^{(1)} \cdot \boldsymbol{J}^{(1)}$ and $\boldsymbol{J}^{(2)} \cdot \boldsymbol{J}^{(2)}$, the eigenvalues j_1 and j_2 will be constant in both sets. Therefore we may confine our attention to the vector space of dimension $(2j_1 + 1)(2j_2 + 1)$ that is spanned by product vectors with fixed values of j_1 and j_2.

Now the $2J + 1$ vectors $|\alpha, J, M\rangle$, with M in the range $-J \leq M \leq J$, span an irreducible subspace. Therefore if the vector $|\alpha, J, M\rangle$, for a particular value of M, can be constructed in the space under consideration, then so can the entire set of $2J + 1$ such vectors with M in the range $-J \leq M \leq J$. For a particular value of J, it might be possible to construct one such set of vectors, two or more linearly independent sets, or none at all.

Let $N(J)$ denote the number of independent sets that can be constructed. Let $n(M)$ be the degree of degeneracy, in this space, of the eigenvalue M . The relation between these two quantities is

$$n(M) = \sum_{J \geq |M|} N(J)$$

and hence

$$N(J) = n(J) - n(J + 1)$$

The product vectors $|j_1, m_1\rangle|j_2, m_2\rangle$ are eigenvectors of the operator $\boldsymbol{J}_z$, with eigenvalue $m_1 + m_2$, and the degree of degeneracy $n(M)$ is equal to the number of pairs (m_1, m_2) such that $M = m_1 + m_2$. Therefore,

$$n(M) = \begin{cases} 0 & |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & |j_1 - j_2| \leq M \leq |j_1 + j_2| \\ 2j_{min} + 1 & 0 \leq |M| \leq |j_1 - j_2| \end{cases}$$

It then follows that

$$N(J) = \begin{cases} 1 & |j_1 - j_2| \leq J \leq |j_1 + j_2| \\ 0 & \text{otherwise} \end{cases}$$

Figure 17.2: Possible values of $M = m_1 + m_2$, illustrated for $j_1 = 3$, $j_2 = 2$

It has turned out that $N(J)$ is never greater that 1, and so the vectors $|\alpha, J, M\rangle$ can be uniquely labelled by the eigenvalues of the four operators $\boldsymbol{J}^{(1)} \cdot \boldsymbol{J}^{(1)}$, $\boldsymbol{J}^{(2)} \cdot \boldsymbol{J}^{(2)}$, $\boldsymbol{J} \cdot \boldsymbol{J}$ and $\boldsymbol{J}_z$. Henceforth these total angular momentum eigenvectors will be denoted as $|j_1, j_2, J, M\rangle$. And we have the unitarity transformation

$$|j_1, j_2, J, M\rangle = \sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle\langle j_1, j_2, m_1, m_2|j_1, j_2, J, M\rangle$$

The coefficients of this transformation are called the Clebsch−Gordan coefficients, denoted as $(j_1, j_2, m_1, m_2|J, M)$. The phases of the CG coefficients are not yet defined because of the indeterminacy of the relative phases of the vectors $|j_1, j_2, J, M\rangle$. For different values of M but fixed J we adopt the usual phase convention that led to

$$J_+|j_1, j_2, J, M\rangle = \sqrt{(J + M + 1)(J - M)}|j_1, j_2, J, M + 1\rangle$$

This leaves one arbitrary phase for each J value, which we fix by requiring that $(j_1, j_2, j_1, J - j_1|J, J)$ be real and positive. It can be shown that all of the CG coefficients are now real.
We can also prove that the CG coefficient vanishes unless the following conditions are satisfied:

- $m_1 + m_2 = M$
- $|j_1 - j_2| \leq J \leq |j_1 + j_2|$
- $j_1 + j_2 + J =$ an integer

It is possible to calculate the values of the CG coefficients by successive application of the raising or lowering operator to

$$|j_1, j_2, J, M\rangle = \sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle(j_1, j_2, m_1, m_2|J, M)$$

The details of the calculation can be found in section 7.7 of *Quantum mechanics - a modern development(Leslie E. Ballentine)*.And we have Table of CG coefficients and Calculator of CG coefficients on the internet. A special case of angular momentum addition is spin−orbit coupling of spin $\frac{1}{2}$ particles, and we list the corresponding CG coefficients $(l, \frac{1}{2}, M - m_s, m_s|J, M)$ in the table 1.1.
Now let us consider the relation between CG coefficients and rotation matrices. On the one hand, we have

$$\langle j_1, j_2, m_1, m_2|\boldsymbol{R}|j_1, j_2, m_1', m_2'\rangle = D^{(j_1)}_{m_1 m_1'}(R) D^{(j_2)}_{m_2 m_2'}(R)$$

	$J=l+\frac{1}{2}$	$J=l-\frac{1}{2}$
$m_s=\frac{1}{2}$	$\left[\frac{l+M+\frac{1}{2}}{2l+1}\right]^{\frac{1}{2}}$	$-\left[\frac{l-M+\frac{1}{2}}{2l+1}\right]^{\frac{1}{2}}$
$m_s=-\frac{1}{2}$	$\left[\frac{l-M+\frac{1}{2}}{2l+1}\right]^{\frac{1}{2}}$	$\left[\frac{l+M+\frac{1}{2}}{2l+1}\right]^{\frac{1}{2}}$

Table 17.1: Spin-Orbit coupling

On the other hand, we have

$$\begin{aligned}
&\langle j_1, j_2, m_1, m_2|\boldsymbol{R}|j_1, j_2, m_1', m_2'\rangle \\
=& \sum_{J,M,J',M'} (j_1, j_2, m_1, m_2|J, M)(j_1, j_2, m_1', m_2'|J', M')\langle j_1, j_2, J, M|\boldsymbol{R}|j_1, j_2, J', M'\rangle \\
=& \sum_{J,M,M'} (j_1, j_2, m_1, m_2|J, M)(j_1, j_2, m_1', m_2'|J, M')D^{(J)}_{MM'}(R)
\end{aligned}$$

So, we can get

$$D^{(j_1)}_{m_1m_1'}(R)D^{(j_2)}_{m_2m_2'}(R) = \sum_{J,M,M'} (j_1, j_2, m_1, m_2|J, M)(j_1, j_2, m_1', m_2'|J, M')D^{(J)}_{MM'}(R)$$

It is called Clebsch-Gordan series.
Recall that

$$Y_l^m(\theta,\phi) = \left(\frac{2l+1}{4\pi}\right)^{1/2} [D^{(j)}_{m0}((\phi,\theta,0))]^*$$

so, we have

$$Y_{l_1}^{m_1}(\theta,\phi)Y_{l_2}^{m_2}(\theta,\phi) = \sum_{l,m}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}(l_1, l_2, m_1, m_2|l, m)(l_1, l_2, 0, 0|l, 0)Y_l^m(\theta,\phi)$$

The orthogonal relation of spherical harmonics then would imply that

$$\int d\Omega Y_l^{m*}(\theta,\phi)Y_{l_1}^{m_1}(\theta,\phi)Y_{l_2}^{m_2}(\theta,\phi) = \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}(l_1, l_2, m_1, m_2|l, m)(l_1, l_2, 0, 0|l, 0)$$

17.5 Tensor operators

Suppose the state of the system is $|\psi\rangle$, then the state after rotation R is $U(R)|\psi\rangle$, denoted as $|\psi'\rangle$. An operator K is called scalar operator if and only if

$$\langle\psi'|K|\psi'\rangle = \langle\psi|K|\psi\rangle$$

i.e.

$$U^{-1}(R)KU(R) = K$$

Taking the case of infinitesimal rotation, we can derive that

$$[\boldsymbol{J}, K] = 0$$

A group of operators $\boldsymbol{V}$ is called vector operator if and only if

$$\langle\psi'|V_i|\psi'\rangle = R_{ii'}\langle\psi|V_{i'}|\psi\rangle$$

i.e.

$$U^{-1}(R)V_iU(R) = \sum_{i'} R_{ii'}V_{i'}$$

Taking the case of infinitesimal rotation, we can derive that

$$[J_i, V_j] = i\epsilon_{ijk}V_k$$

If $\boldsymbol{V}$ and $\boldsymbol{W}$ are vector operators, we can prove that $\boldsymbol{V}\cdot\boldsymbol{W}$ is scalar operator and $\boldsymbol{V}\times\boldsymbol{W}$ is vector operator.
Similarly, tensor operators are defined as

$$U^{-1}(R)T_{ij\cdots k}U(R) = \sum_{i'\cdots} R_{ii'}R_{jj'}\cdots R_{kk'}T_{i'j'\cdots k'}$$

Such a tensor is known as a Cartesian tensor.
The trouble with a Cartesian tensor is that it is reducible, i.e. it can be decomposed into objects that transform independently under rotations. For example, the trace of a tensor transform like a scalar under rotations. So we now define spherical tensor operators which are irreducible under rotations. We define a spherical tensor operator of rank k with $(2k+1)$ components as

$$U^{-1}(R)T_q^{(k)}U(R) = \sum_{q'=-k}^{k}[D_{qq'}^{(k)}(R)]^*T_{q'}^{(k)}$$

or equivalently

$$U(R)T_q^{(k)}U^{-1}(R) = \sum_{q'=-k}^{k}D_{q'q}^{(k)}(R)T_{q'}^{(k)}$$

where $D_{qq'}^{(k)}$ is the rotation matrix.
Taking the case of infinitesimal rotation, we can derive that

$$[J_\pm, T_q^{(k)}] = \sqrt{(k\mp q)(k\pm q+1)}T_{q\pm1}^{(k)}$$

$$[J_z, T_q^{(k)}] = qT_q^{(k)}$$

For example, spherical components of a vector operator $\boldsymbol{V}$,

$$V_{-1} = \frac{V_x - iV_y}{\sqrt{2}} \quad V_0 = V_z \quad V_1 = -\frac{V_x + iV_y}{\sqrt{2}}$$

satisfy the commutation relation above, so they are spherical tensor of rank 1. Generally, if $\boldsymbol{V}$ is a vector operator, then $Y_l^m(\boldsymbol{V})$ is a spherical tensor of ranks l.
Spherical tensors can be formed as products of other spherical tensors, we have following theorem.

Theorem 17.1

Let $X_{q_1}^{(k_1)}$ and $Z_{q_2}^{(k_2)}$ be irreducible spherical tensors of rank k_1 and k_2. Then

$$T_q^{(k)} = \sum_{q_1 . q_2} (k_1, k_2, q_1, q_2 | k, q) X_{q_1}^{(k_1)} Z_{q_2}^{(k_2)}$$

is a irreducible spherical tensor of rank k.

The proof can be found in section 3.10 of *Modern Quantum Mechanics(J.J.Sakurai)*.
For example, suppose $\boldsymbol{V}$ and $\boldsymbol{U}$ are spherical tensor

$$T_0^{(0)} = \sqrt{\frac{1}{3}}(U_{-1}V_1 + U_1V_{-1} - U_0U_0) = -\sqrt{\frac{1}{3}}(U_x^2 + U_y^2 + U_z^2)$$

is tensor of rank 1, then is a spherical tensor of rank 0. Another important theorem is Wigner-Eckart theorem.

Theorem 17.2 Wigner-Eckart theorem

The matrix elements of tensor operators with respect to angular-momentum eigenstates satisfy

$$\langle \tau', j', m' | T_q^{(k)} | \tau, j, m \rangle = (j, k, m, q | j', m') \frac{\langle \tau', j' || T^{(k)} || \tau, j \rangle}{\sqrt{2j+1}}$$

where the double-bar matrix element is independent of m and m' and q.

The proof can be found in section 3.10 of *Modern Quantum Mechanics(J.J.Sakurai)*.
So, for scaler operator K, we have

$$\langle \tau', j', m' | S | \tau, j, m \rangle = \delta_{jj'} \delta_{mm'} \frac{\langle \tau', j' || S || \tau, j \rangle}{\sqrt{2j+1}}$$

For spherical tensor of rank 1, we have

$$\langle \tau', j', m' | V_q | \tau, j, m \rangle = (j, 1, m, q | j', m') \frac{\langle \tau', j' || V_q || \tau, j \rangle}{\sqrt{2j+1}}$$

It would vanish unless

$$m' - m = q \quad j' - j = 0, 1, -1 \quad j \text{ and } j' \text{ are not both } 0$$

For $j = j'$, Wigner-Eckart theorem - when applied to the vector operator- takes a particularly simple form. We can derive that

$$\langle \tau', j, m' | V_q | \tau, j, m \rangle = \frac{\langle \tau', j, m | \boldsymbol{J} \cdot \boldsymbol{V} | \tau, j, m \rangle}{j(j+1)} \langle j, m' | J_q | j, m \rangle$$

For example, the magnetic moment operator for an atom has the form

$$\boldsymbol{\mu} = \frac{-e}{2m_e}(g_L \boldsymbol{L} + g_s \boldsymbol{S})$$

The parameters g_L and g_S have approximately the values $g_L = 1$ and $g_S = 2$. The former is an generalization of the magnetic moment we calculated in classical electrodynamics for a system of charged particles. The latter will be discussed in quantum field theory. We define the effective Lande factor as

$$\langle \tau, J, M' | \boldsymbol{\mu} | \tau, J, M \rangle = \frac{-e}{2m_e} g_{eff} \langle J, M' | \boldsymbol{J} | J, M \rangle$$

Then, we have

$$g_{eff} = \frac{\langle \tau, J, M | g_L \boldsymbol{L} \cdot \boldsymbol{J} + g_s \boldsymbol{S} \cdot \boldsymbol{J} | \tau, J, M \rangle}{J(J+1)} = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}$$

17.6 Spherical potential well

The stationary states of a particle in the spherical potential well are determined by

$$-\frac{1}{2m}\nabla^2 \Psi + W(r)\Psi = E\Psi$$

In spherical coordinates,

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 \frac{\partial}{\partial r}\right] + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left[\sin\theta \frac{\partial}{\partial \theta}\right] + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}$$

So, the eigenvalue equation becomes

$$-\frac{1}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 \frac{\partial \Psi}{\partial r}\right] + \frac{L^2}{2Mr^2}\Psi + W(r)\Psi = E\Psi$$

Suppose the eigenfunctions have the factored form

$$\Psi(r, \theta, \phi) = Y_l^m(\theta, \phi)\frac{u(r)}{r}$$

The radial function then satisfies the equation

$$-\frac{1}{2m}\frac{d^2 u(r)}{dr^2} + \left[\frac{l(l+1)}{2mr^2} + W(r)\right] u(r) = Eu(r)$$

The radial function must satisfy the boundary condition $u(0) = 0$ since $\Psi(r, \theta, \phi)$ would otherwise have an r^{-1} singularity at the origin. The normalization $\langle \Psi | \Psi \rangle = 1$ implies that

$$\int_0^\infty = |u(r)|^2 dr = 1$$

The hydrogen atom

The hydrogen atom is a two-particle system consisting of an electron and a proton. The Hamiltonian is

$$H = \frac{P_e^2}{2m_e} + \frac{P_p^2}{2m_p} - \frac{e^2}{4\pi|\boldsymbol{Q}_e - \boldsymbol{Q}_p|}$$

We take as independent variables the center of mass and relative coordinates of the particles

$$\boldsymbol{Q}_c = \frac{m_e\boldsymbol{Q}_e + m_p\boldsymbol{Q}_p}{m_e + m_p} \quad \boldsymbol{Q}_r = \boldsymbol{Q}_e - \boldsymbol{Q}_p$$

The corresponding momentum operators are

$$\boldsymbol{P}_c = \boldsymbol{P}_e + \boldsymbol{P}_p \quad \boldsymbol{P}_r = \frac{m_p\boldsymbol{P}_e - m_e\boldsymbol{P}_p}{m_e + m_p}$$

We can verify that

$$[Q_{c\alpha}, P_{c\beta}] = [Q_{r\alpha}, P_{r\beta}] = i\delta_{\alpha\beta} \quad [Q_{c\alpha}, P_{r\beta}] = [Q_{r\alpha}, P_{c\beta}] = 0$$

The Hamiltonian becomes

$$H = \frac{P_c^2}{2(m_e + m_p)} + \frac{p_r^2}{2\mu} - \frac{e^2}{4\pi|\boldsymbol{Q}_r|}$$

where μ is called the reduced mass, and is defined by $\mu \equiv \frac{m_e m_p}{m_e + m_p}$. The center of mass behaves as a free particle, and its motion is not coupled to the relative coordinate. We shall confine our attention to the internal degrees of freedom described by the relative coordinate $\boldsymbol{Q}_r$. The energy eigenvalue equation in coordinate representation is

$$-\frac{1}{2\mu}\nabla^2\Psi(\boldsymbol{r}) - \frac{e^2}{4\pi r}\Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r})$$

Suppose $\Psi(r, \theta, \phi) = Y_l^m(\theta, \phi)\frac{u(r)}{r}$, we have

$$-\frac{1}{2\mu}\frac{d^2u(r)}{dr^2} + \left[\frac{l(l+1)}{2\mu r^2} - \frac{e^2}{4\pi r}\right]u(r) = Eu(r)$$

Define

$$\rho \equiv \alpha r \quad \alpha \equiv \sqrt{8\mu|E|} \quad \lambda \equiv \frac{e^2}{4\pi}\sqrt{\frac{\mu}{2|E|}}$$

we have

$$\frac{d^2u}{d\rho^2} + \left[-\frac{1}{4} + \frac{\lambda}{\rho} - \frac{l(l+1)}{\rho^2}\right]u = 0$$

As $\rho \to \infty$, we have $u \sim e^{-\rho/2}$. And as $\rho \to 0$, we have $u \sim \rho^{l+1}$. So, we can suppose

$$u(\rho) = \rho^{l+1}e^{-\rho/2}v(\rho)$$

And we can get

$$\rho\frac{d^2v}{d\rho^2} + (2l + 2 - \rho)\frac{dv}{d\rho} + (\lambda - l - 1)v = 0$$

It is the so-called Confluent Hypergeometric Differential Equation. When $\lambda - 1 - l = n_r$, we have regular solutions. Solutions are Associated Laguerre Polynomial, and will be denoted as $L^{2l+1}_{n-l-1}(\rho)(n = n_r + l + 1)$. The energy levels are

$$E_n = -\frac{\mu e^4}{32\pi^2 n^2}$$

The degeneracy of an eigenvalue E_n is

$$\sum_{l=0}^{n-1}(2l+1) = n^2$$

Note: *The degeneracy of an energy level of a hydrogen atom is greater than this by a factor of* 4*, which arises from the two-fold orientational degeneracies of the electron and proton spin states. This four-fold degeneracy is modified by the hyperfine interaction between the magnetic moments of the electron and the proton.*

The orthonormal energy eigenfunctions for the hydrogen atom are

$$\Psi_{nlm}(r,\theta,\phi) = \left[\frac{4(n-l-1)!}{(na_0)^3 n[(n+l)!]^3}\right]^{\frac{1}{2}} \rho^l L^{2l+1}_{n-l-1}(\rho) e^{-\rho/2} Y^m_l(\theta,\phi)$$

where $\rho = \alpha r = \frac{2r}{na_0}$, and $a_0 \equiv \frac{4\pi}{\mu e^2}$ is a characteristic length for the atom, known as the Bohr radius.The ground state wave function is

$$\Psi_{000} = (\pi a_0^3)^{-\frac{1}{2}} e^{-\frac{r}{a_0}}$$

A measure of the spatial extent of the bound states of hydrogen is given by the averages of various powers of the distance r.

$$\begin{aligned}
\langle r\rangle &= n^2 a_0 \left\{1 + \frac{1}{2}\left[1 - \frac{l(l+1)}{n^2}\right]\right\} \\
\langle r^2\rangle &= n^4 a_0^2 \left\{1 + \frac{3}{2}\left[1 - \frac{l(l+1) - 1/3}{n^2}\right]\right\} \\
\langle \frac{1}{r}\rangle &= \frac{1}{n^2 a_0}
\end{aligned}$$

Chapter 18
Discrete Symmetries

18.1 Space inversion

The space inversion transformation is $\boldsymbol{x} \to -\boldsymbol{x}$. The corresponding operator on state vector space is usually called the parity operator. It will be denoted by P. By definition, the parity operator reverses the signs of the position operator and the momentum operator

$$P^{-1}\boldsymbol{X}P = -\boldsymbol{X} \quad P^{-1}\boldsymbol{P}P = -\boldsymbol{P}$$

It follows that the orbital angular momentum, $\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P}$, is unchanged by the parity transformation. This property is extended, by definition, to any angular momentum operator,

$$P^{-1}\boldsymbol{J}P = \boldsymbol{J}$$

We can verify that P must be linear by applying space inversion to the commutation relation $[X_i, P_i] = i$. Therefore the parity operator is a unitary operator rather than an anti-unitary operator. Since two consecutive space inversions produce no change at all, it follows that the states described by $|\psi\rangle$ and by $P^2|\psi\rangle$ must be the same. Thus the operator P^2 can differ from the identity operator by at most a phase factor. This phase factor is left arbitrary. It is most convenient to choose that phase factor to be unity, and hence we have

$$P = P^{-1} = P^\dagger$$

Further more, we can derive that

$$P|\boldsymbol{x}\rangle = |-\boldsymbol{x}\rangle$$

So the effect of P on a wave function is

$$P\psi(\boldsymbol{x}) \equiv \langle \boldsymbol{x}|P|\psi\rangle = \langle -\boldsymbol{x}|\psi\rangle = \psi(-\boldsymbol{x})$$

From the fact the $P^2 = 1$, it follows that P has eigenvalues ± 1. Any even function, $\psi_e(\boldsymbol{x}) = \psi_e(-\boldsymbol{x})$, is an eigenfunction on P with eigenvalue 1, and any odd function, $\psi_o(\boldsymbol{x}) = -\psi_o(-\boldsymbol{x})$, is an eigenfunction of P with eigenvalue -1. A function corresponding to parity $+1$ is also said to be of even parity, and a function corresponding to parity -1 is said to be of odd parity. If the parity of operator K is p, i.e.

$$PKP = pK$$

the parity of the state $|\psi_1\rangle$ and $|\psi_2\rangle$ are p_1 and p_2 respectively. Then we can prove that

$$\langle\psi|K|\psi\rangle$$

vanishes unless $p = p_1 p_2$.

Example
Under space inversion, $\boldsymbol{x} \to -\boldsymbol{x}$, the spherical harmonic undergoes the transformation

$$Y_l^m(\theta, \phi) \to Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi)$$

Hence the single particle orbital angular momentum eigenvector $|l, m\rangle$ is also an eigenvector of parity, with parity $(-1)^l$.
A total orbital angular momentum eigenvector for a two- electron atom is of the form

$$|l_1, l_2, L, M\rangle = \sum_{m_1, m_2} \langle l_1, l_2, m_1, m_2 | l_1, l_2, L, M\rangle |l_1, m_1\rangle \otimes |l_2, m_2\rangle$$

It is apparent that

$$P|l_1, l_2, L, M\rangle = (-1)^{l_1 + l_2} |l_1, l_2, L, M\rangle$$

and that $(-1)^{l_1+l_2} \neq (-1)^L$. Thus we see that, in general, the parity of an angular momentum state is not determined by its total angular momentum.

If the parity operator P commutes with the Hamiltonian H, then parity eigenvalue ± 1 is a conserved quantity. In that case an even parity state can never acquire an odd parity component, and an odd parity state can never acquire an even parity component.
If $|\psi(t)\rangle$ is a physical process of the system with Hamiltonian H, then we have Schrödinger equation

$$H|\psi\rangle = i\frac{\partial |\psi\rangle}{\partial t}$$

If $PH = HP$, we can verify that

$$HP|\psi\rangle = i\frac{\partial P|\psi\rangle}{\partial t}$$

So, the space inversion of $|\psi(t)\rangle$, $P|\psi(t)\rangle$, can also be a possible physical process of the system.
Experiments have shown that parity in β decay is not conserved.

18.2 Time reversal

The effect of the time reversal operator T is to reverse the linear and angular momentum while leaving the position unchanged. Thus we require, by definition,

$$T^{-1}\boldsymbol{X}T = \boldsymbol{X} \quad T^{-1}\boldsymbol{P}T = -\boldsymbol{P} \quad T^{-1}\boldsymbol{J}T = -\boldsymbol{J}$$

We can verify that T must be anti-linear by applying space inversion to the commutation relation $[X_i, P_i] = i$. Therefore the parity operator is a anti-unitary operator.
The time evolution of a system satisfied Schrödinger equation

$$H|\psi(t)\rangle = i\frac{\partial |\psi(t)\rangle}{\partial t}$$

Suppose that $TH = HT$, we can derive that

$$HT|\psi(t)\rangle = -i\frac{\partial T|\psi(t)\rangle}{\partial t}$$

i.e.

$$HT|\psi(-t)\rangle = i\frac{\partial T|\psi(-t)\rangle}{\partial t}$$

$T|\psi(-t)\rangle$ is also a possible physical process of the system.

In coordinate representation the Schrödinger equation takes the form

$$\left[-\frac{1}{2m}\nabla^2 + W(\boldsymbol{x})\right]\psi(\boldsymbol{x},t) = i\frac{\partial\psi(\boldsymbol{x},t)}{\partial t}$$

Its complex conjugate is

$$\left[-\frac{1}{2m}\nabla^2 + W^*(\boldsymbol{x})\right]\psi^*(\boldsymbol{x},t) = -i\frac{\partial\psi^*(\boldsymbol{x},t)}{\partial t}$$

The condition for the Hamiltonian to be invariant under complex conjugation is that the potential be real: $W = W^*$. In that case it is apparent that if $\psi(\boldsymbol{x},t)$ is a solution then so is $\psi^*(\boldsymbol{x},t)$. This suggests that we may identify the time reversal operator with the complex conjugation operator in this representation,

$$T = K_0$$

where, by definition, $K_0\psi(\boldsymbol{x},t) = \psi^*(\boldsymbol{x},t)$. In this case T is its own inverse. The formal expression for an arbitrary vector in coordinate representation is $|\psi\rangle = \int\psi(\boldsymbol{x})|\boldsymbol{x}\rangle d^3\boldsymbol{x}$, where the basis vector $|\boldsymbol{x}\rangle$ is an eigenvector of the position operator. Since T is equal to the complex conjugation operator, its effect is simply $T|\psi\rangle = \int\psi^*(\boldsymbol{x})|\boldsymbol{x}\rangle d^3\boldsymbol{x}$, with $T|\boldsymbol{x}\rangle = |\boldsymbol{x}\rangle$.

In momentum representation, an arbitrary vector can be written as

$$|\psi\rangle = \int\psi(\boldsymbol{p})|\boldsymbol{p}\rangle d^3\boldsymbol{p}$$

Since

$$T|\boldsymbol{p}\rangle = \int\langle\boldsymbol{x}|\boldsymbol{p}\rangle^*|\boldsymbol{x}\rangle d^3\boldsymbol{x} = |-\boldsymbol{p}\rangle$$

so we have

$$T|\psi\rangle = \int\psi^*(\boldsymbol{p})|-\boldsymbol{p}\rangle d^3\boldsymbol{p} = \int\psi^*(-\boldsymbol{p})|\boldsymbol{p}\rangle d^3\boldsymbol{p}$$

The time reversal operator must reverse the angular momentum. For spin operator, we have

$$T^{-1}\boldsymbol{S}T = -\boldsymbol{S}$$

In the standard representation of the spin operators, S_x and S_z are real, while S_y is imaginary. The time reversal operator T cannot be equal to the complex conjugation operator K_0 in this representation, since the effect of the latter is

$$K_0S_xK_0 = S_x \quad K_0S_yK_0 = -S_y \quad K_0S_zK_0 = S_z$$

Let us write the time reversal operator as $T = YK_0$, where Y is a linear operator. Y must have the following properties:

$$Y^{-1}S_xY = -S_x \quad Y^{-1}S_yY = S_y \quad Y^{-1}S_zY = -S_z$$

And Y must operate only on the spin degrees of freedom. A reasonable choice is that $Y = e^{-i\pi S_y}$, whose effect is to rotate spin (and only spin) through the angle π about the y axis. Therefore the explicit form of the time reversal in this representation is

$$T = e^{-i\pi S_y}K_0$$

Two successive applications of the time reversal transformation, must leave the physical situation unchanged. Therefore

$$T^2|\Psi\rangle = c|\Psi\rangle$$

where $|c| = 1$. And we have

$$T^2(T|\Psi\rangle) = T(T^2|\Psi\rangle) = T(c|\Psi\rangle) = c^*T|\Psi\rangle$$

so

$$T^2(|\Psi\rangle + T|\Psi\rangle) = c|\Psi\rangle + c^*T|\Psi\rangle = c'(|\Psi\rangle + T|\Psi\rangle)$$

And we can determine that $c' = c^* = c$. Thus we must have $c = \pm 1$, i.e.

$$T^2|\Psi\rangle = \pm|\Psi\rangle$$

In the particularly representation, we have

$$T^2 = e^{-i\pi S_y}K_0e^{-i\pi S_y}K_0 = e^{-i2\pi S_y}$$

This may equivalently be written as

$$T^2 = e^{-i2\pi J_y}$$

since $e^{-i2\pi L_y} = I$. So we have an identity

$$T^2 = R(2\pi)$$

Kramer's theorem

Let us consider the energy eigenvalue equation, $H|\Psi\rangle = E|\Psi\rangle$, for a time-reversal-invariant Hamiltonian, $TH = HT$. Then $HT|\Psi\rangle = TH|\Psi\rangle = ET|\Psi\rangle$, and so both $|\Psi\rangle$ and $T|\Psi\rangle$ are eigenvectors with energy eigenvalue E. There are two possibilities: (a) $|\Psi\rangle$and $T|\Psi\rangle$ are linearly dependent, and so describe the same state. (b) $|\Psi\rangle$and $T|\Psi\rangle$ are linearly independent, and so describe two degenerate states.
Suppose that (a) is true, in which case we must have $T|\Psi\rangle = a|\Psi\rangle$ with $|a| = 1$. A second application of T yields $T^2|\Psi\rangle = |\Psi\rangle$. So for those states that satisfy $T^2|\Psi\rangle = -|\Psi\rangle$ it is necessarily true that $|\Psi\rangle$and $T|\Psi\rangle$ are linearly independent, degenerate states. This result is known as Kramer's theorem: any system for which $T^2|\Psi\rangle = -|\Psi\rangle$ has only degenerate energy levels.

Chapter 19
Approximation method

19.1 Time independent perturbation theory

19.1.1 Brillouin-Wigner perturbation theory

We consider an unperturbed Hamiltonian H_0 with eigenvalues ϵ_k and eigenstates $|k\alpha\rangle$, where α is an index introduced to resolve degeneracies, so that

$$H_0|k\alpha\rangle = \epsilon_k|k\alpha\rangle$$

We pick one of these levels ϵ_n for study, so the index n will be fixed for the following discussion. We denote the eigenspace of the unperturbed system corresponding to eigenvalue ϵ_n by $\mathcal{H}$, so that the unperturbed eigenkets $\{|n\alpha\rangle, \alpha = 1, 2, \cdots\}$ form a basis in this space.
We take the perturbed Hamiltonian to be $H = H_0 + \lambda H_1$, where λ is a formal expansion parameter that we allow to vary between 0 and 1 to interpolate between the unperturbed and perturbed system. When the perturbation is turned on, the unperturbed energy level ϵ_n may split and shift. We denote one of the exact energy levels that grows out of ϵ_n by E. We let $|\psi\rangle$ be an exact energy eigenket corresponding to energy E, so that

$$H|\psi\rangle = (H_0 + \lambda H_1)|\psi\rangle = E|\psi\rangle$$

Both E and $|\psi\rangle$ are understood to be functions of λ; as $\lambda \to 0$, E approaches ϵ_n and $|\psi\rangle$ approaches some state lying in $\mathcal{H}_n$. We break the Hilbert space into the subspace $\mathcal{H}_n$ and its orthogonal complement which we denote by $\mathcal{H}_n^\perp$. The components of $|\psi\rangle$ parallel and perpendicular to $\mathcal{H}_n$ are conveniently expressed in terms of the projector P onto the subspace $\mathcal{H}_n$ and the orthogonal projector Q, defined by

$$P = \sum_\alpha |n\alpha\rangle\langle n\alpha| \quad Q = \sum_{k\neq n,\alpha} |k\alpha\rangle\langle k\alpha|$$

These projectors satisfy

$$P^2 = P \quad Q^2 = Q \quad PQ = QP = 0 \quad P + Q = I \quad [P, H_0] = [Q, H_0] = 0$$

The component $P|\psi\rangle$ is a linear combination of the known unperturbed eigenstates $\{|n\alpha\rangle, \alpha = 1, 2, \cdots\}$, and is easily characterized. The orthogonal component $Q|\psi\rangle$ is harder to find. It turns out it is possible to write a neat power series expansion for this solution. Firstly, we have

$$(E - H_0)|\psi\rangle = \lambda H_1|\psi\rangle$$

Now we define a new operator R

$$R \equiv \sum_{k\neq n,\alpha} \frac{|k\alpha\rangle\langle k\alpha|}{E-\epsilon_k}$$

Note: *If there are other unperturbed energy levels ϵ_k lying close to ϵ_n, then the perturbation could push the exact energy E near to or past some of these other levels, and then other small denominators would make R ill defined. This will certainly happen if the perturbation is large enough. For the time being we will assume this does not happen, so that R is free of small denominators. When this is not the case we shall refer to "nearly degenerate perturbation theory", which is discussed later.*

The operator R satisfies

$$PR = RP = 0 \quad QR = RQ = R \quad R(E-H_0) = (E-H_0)R = Q$$

Then we have

$$R(E-H_0)|\psi\rangle = Q|\psi\rangle = \lambda R H_1|\psi\rangle$$

and

$$|\psi\rangle = P|\psi\rangle + \lambda R H_1|\psi\rangle$$

$|\psi\rangle$ can be solved as a series of $P|\psi\rangle$:

$$|\psi\rangle = \frac{1}{1-\lambda R H_1}P|\psi\rangle = P|\psi\rangle + \lambda R H_1 P|\psi\rangle + \lambda^2 R H_1 R H_1 P|\psi\rangle + \cdots$$

19.1.2 Nondegenerate perturbation theory

In nondegenerate perturbation theory the level ϵ_n of H_0 is nondegenerate. Then the index α is not needed for the level ϵ_n, and we can write simply $|n\rangle$ for the corresponding eigenstate. We assume that $P|\psi\rangle$ is normalized rather than $\psi\rangle$ so that

$$P|\psi\rangle = |n\rangle$$

With this normalization convention, we have

$$\langle n|\psi\rangle = 1$$

Now the series becomes

$$|\psi\rangle = |n\rangle + \lambda \sum_{k\neq n,\alpha} |k\alpha\rangle \frac{\langle k\alpha|H_1|n\rangle}{E-\epsilon_k} + \lambda^2 \sum_{k\neq n,\alpha}\sum_{k'\neq n,\alpha'} |k\alpha\rangle \frac{\langle k\alpha|H_1|k'\alpha'\rangle\langle k'\alpha'|H_1|n\rangle}{(E-\epsilon_k)(E-\epsilon_{k'})}$$

To find an equation for E, we have

$$\langle n|E-H_0|\psi\rangle = E-\epsilon_n = \lambda\langle n|H_1|\psi\rangle$$

then we can get

$$\begin{aligned} E &= \epsilon_n + \lambda\langle n|H_1|n\rangle + \lambda^2\langle n|H_1 R H_1|n\rangle + \lambda^3\langle n|H_1 R H_1 R H_1|n\rangle + \cdots \\ &= \epsilon_n + \lambda\langle n|H_1|n\rangle + \lambda^2 \sum_{k\neq n,\alpha} \frac{\lambda\langle n|H_1|k\alpha\rangle\langle k\alpha|H_1|n\rangle}{E-\epsilon_k} \\ &+ \lambda^3 \sum_{k\neq n,\alpha}\sum_{k'\neq n,\alpha'} \frac{\langle n|H_1|k\alpha\rangle\langle k\alpha|H_1|k'\alpha'\rangle\langle k'\alpha'|H_1|n\rangle}{(E-\epsilon_k)(E-\epsilon_{k'})} + \cdots \end{aligned}$$

It is easy to get E up to $O(\lambda^3)$,

$$E = \epsilon_n + \lambda\langle n|H_1|n\rangle + \lambda^2 \sum_{k\neq n,\alpha} \frac{\lambda\langle n|H_1|k\alpha\rangle\langle k\alpha|H_1|n\rangle}{\epsilon_n - \epsilon_k} + O(\lambda^3)$$

and $|\psi\rangle$ up to $O(\lambda^2)$,

$$|\psi\rangle = |n\rangle + \lambda \sum_{k\neq n,\alpha} |k\alpha\rangle \frac{\langle k\alpha|H_1|n\rangle}{\epsilon_n - \epsilon_k} + O(\lambda^2)$$

Higher corrections can be found on the internet.

19.1.3 Degenerate perturbation theory

In the case that the unperturbed energy level ϵ_n is degenerate, we have

$$P|\psi\rangle = \sum_\alpha |n\alpha\rangle c_\alpha$$

and

$$\langle n\alpha|P|\psi\rangle = \langle n\alpha|\psi\rangle = c_\alpha$$

Then we can obtain an equation for the c_α,

$$\langle n\alpha|E - H_0|\psi\rangle = c_\alpha(E - \epsilon_n) = \lambda\langle n\alpha|H_1|\psi\rangle$$

then we can get

$$\begin{aligned}(E - \epsilon_n)c_\alpha &= \lambda \sum_\beta \langle n\alpha|H_1|n\beta\rangle c_\beta + \lambda^2 \sum_\beta \langle n\alpha|H_1 R H_1|n\beta\rangle c_\beta + \cdots \\ &= \lambda \sum_\beta \langle n\alpha|H_1|n\beta\rangle c_\beta + \lambda^2 \sum_\beta \sum_{k\neq n,\gamma} \frac{\lambda\langle n\alpha|H_1|k\gamma\rangle\langle k\gamma|H_1|n\beta\rangle}{E - \epsilon_k} c_\beta + \cdots\end{aligned} \tag{19.1}$$

This equation must be solved simultaneously for the eigenvalues E and the unknown expansion coefficients c_α.
If we truncate the series at first order, we see that the corrections $E - \epsilon_n$ to the energies are determined as the eigenvalues of the matrix $\langle n\alpha|H_1|n\beta\rangle$, and the coefficients c_α are the corresponding eigenvectors. This determines the energies to first order, but the coefficients c_α only to zeroth order. Then $P|\psi\rangle$ becomes known to zeroth order and $Q|\psi\rangle$ to first order.
The first order matrix may or may not have degeneracies itself. If it does not, then all degeneracies are lifted at first order; if it does, the remaining degeneracies may be lifted at a higher order, or may persist to all orders. Degeneracies that persist to all orders are almost always due to some symmetry of the system, which can usually be recognized at the outset.
The higher order corrections can be calculated step by step, which will not be listed here.

Now let us consider the case in which the unperturbed levels of H_0 , while not technically degenerate, are close to one another. Suppose to be specific that two levels, say, ϵ_n and ϵ_m, are close enough to one another that first order perturbations will push the exact level E close to

or onto the unperturbed level ϵ_m.
In this case we choose some energy, call it $\bar{\epsilon}$, which is close to ϵ_n and ϵ_m. Then let us take the original unperturbed Hamiltonian and perturbation and rearrange them in the form,

$$H = H_0 + H_1 = H_0' + H_1'$$

where

$$\begin{aligned} H_0 &= \sum_{k\alpha} \epsilon_k |k\alpha\rangle\langle k\alpha| \\ H_0' &= \sum_{k\neq m,n;\alpha} \epsilon_k |k\alpha\rangle\langle k\alpha| + \sum_{k=m,n;\alpha} \bar{\epsilon}|k\alpha\rangle\langle k\alpha| \\ H_1' &= H_1 + \sum_{k=m,n;\alpha} (\epsilon_k - \bar{\epsilon})|k\alpha\rangle\langle k\alpha| \end{aligned}$$

Then standard degenerate perturbation theory may be applied. We will call this procedure "nearly degenerate perturbation theory."

19.2 Application of time independent perturbation theory in hydrogen atom

19.2.1 Stark effect

The Stark effect concerns the behaviour of atoms in external electric fields. We choose hydrogen atom because it is single-electron atoms. The hydrogen atom will be modelled with the central force Hamiltonian

$$H_0 = \frac{p^2}{2m} - \frac{e^2}{4\pi r}$$

In this Hamiltonian we ignore spin and other small effects such as relativistic corrections, hyperfine effects and the Lamb shift. These effects cause a splitting and shifting of the energy levels of our simplified model, as well as the introduction of new quantum numbers and new degrees of freedom. But these effects are all small, and if the applied electric field is strong enough, it will overwhelm them and the physical consequences will be much as we shall describe them with our simplified model.
The unperturbed energy levels in hydrogen are given by

$$E_n = -\frac{1}{2n^2}\frac{e^2}{4\pi a_0}$$

where a_0 is the Bohr radius. These levels are n^2 degenerate.
As for the perturbation, let us write $\boldsymbol{F}$ for the external electric field , and let us take it to lie in the z-direction. Thus, the perturbing potential has the form

$$V_1 = -(-e)\boldsymbol{F}\cdot\boldsymbol{x} = eFz$$

For small z, the attractive Coulomb field dominates the total potential and we have the usual Coulomb well that supports atomic bound states. However, for large negative z, the unperturbed potential goes to zero, while the perturbing potential becomes large and negative. At

intermediate values of negative z, the competition between the two potentials gives a maximum in the total potential. The electric force on the electron is zero at the maximum of the potential. Given the relative weakness of the applied field, the maximum must occur at a distance from the nucleus that is large in comparison to the Bohr radius a_0. Atomic states with small principal quantum numbers n lie well inside this radius. The perturbation analysis we shall perform applies to these states.
The bound states of the unperturbed system are able to tunnel through the potential barrier. When an external electric field is turned on, the bound states of the atom cease to be bound in the strict sense, and become resonances. Electrons that tunnel through the barrier and emerge into the classically allowed region at large negative z will accelerate in the external field, leaving behind an ion. This is the phenomenon of field ionization. This effect can be neglected if the external field is weak enough and the lifetime of the "bound state" is long enough.

In the case of hydrogen, the ground state is $|100\rangle$. The first order shift in the ground state energy level is given by

$$\Delta E^{(1)}_{gnd} = \langle 100|eFz|100\rangle = 0$$

which vanishes because the parity of z is odd, but $\langle 100|$ and $|100\rangle$ have the same parity.
For the excited states of hydrogen, according to first order degenerate perturbation theory, the shifts in the energy levels E_n are given by the eigenvalues of the $n^2 \times n^2$ matrix,

$$\langle nlm|eFz|nl'm'\rangle$$

According to the Wigner-Eckart theorem and parity, the matrix elements vanish unless $l - l' = \pm 1$ and $m = m'$. Consider, for example, the case $n = 2$. The four degenerate states are $|2,0,0\rangle$, $|2,1,-1\rangle$, $|2,1,0\rangle$ and $|2,1,1\rangle$. Only the states $|2,0,0\rangle$ and $|2,1,0\rangle$ are connected by the perturbation. Therefore of the 16 matrix elements, the only nonvanishing ones are

$$\langle 2,0,0|eFz|2,1,0\rangle = -W = -3eFa_0$$

and its complex conjugate. The matrix connecting the two states $|2,0,0\rangle$ and $|2,1,0\rangle$ is

$$\begin{pmatrix} 0 & -W \\ -W & 0 \end{pmatrix}$$

and its eigenvalues are the first order energy shifts in the $n = 2$ level,

$$\Delta E^{(1)}_2 = \pm W$$

In addition, the two states $|2,1,-1\rangle$ and $|2,1,1\rangle$ do not shift their energies at first order. The perturbed eigenfunctions are

$$|+W\rangle = \frac{|2,0,0\rangle - |2,1,0\rangle}{\sqrt{2}} \quad |-W\rangle = \frac{|2,0,0\rangle + |2,1,0\rangle}{\sqrt{2}}$$

This is zeroth order part of the exact eigenstates.
Now let us look at the exact symmetries of the full, perturbed Hamiltonian $H = H_0 + H_1$, without doing perturbation theory at all. Since $[H, L_z]$ the exact eigenstates of H can be chosen

to be eigenstates of L_z as well. Denote these by $|\gamma m\rangle$, where γ is an additional index needed to specify an energy eigenstate. Thus, we have

$$L_z|\gamma m\rangle = m|\gamma m\rangle \quad H|\gamma m\rangle = E_{\gamma m}|\gamma m\rangle$$

where $E_{\gamma m}$ is allowed to depend on m since the full rotational symmetry is broken.
As for time reversal, the state $T|\gamma m\rangle$ must be an eigenstate of energy with eigenvalue $E_{\gamma m}$ since $TH = HT$. But because $T^{-1}L_zT = -L_z$, it also follows that $T|\gamma m\rangle$ is an eigenstate of L_z with eigenvalue $-m$. If $m \neq 0$, we must have a degeneracy of at least two. The only energy levels that can be nondegenerate are those with $m = 0$. In the example above, even higher order corrections cannot separate $|2, 1, -1\rangle$ and $|2, 1, 1\rangle$.

19.2.2 Fine structure

Fine structure of atoms concerns the effects of relativity and spin on the dynamics of the electron. Both these effects are of the same order of magnitude, and must be treated together in any realistic treatment of the atomic structure.
The fine structure terms account for relativistic effects through order v^2 , and have the effect of enlarging the Hilbert space by the inclusion of the spin degrees of freedom, introducing new quantum numbers, and shifting and splitting the energy levels of the electrostatic model. The splitting in particular means that spectral lines that appear a singlets under low resolution become closely spaced multiplets under higher resolution.
Derivation of the exact form of relativistic corrections of Hamiltonian in quantum mechanics can be very rigorous and needs some reasonable guess. The details of derivation can be found in lecture notes on fine structure by Robert G. Littlejohn. Here we just list the result.

$$H_{FS} = H_{RKE} + H_D + H_{SO}$$

The term H_{RKE} is due to the second order term of the expansion series of $E = \sqrt{p^2 + m^2}$. (The first order term is just the kinetic energy in non relativistic quantum mechanics). We have

$$H_{RKE} = -\frac{p^4}{8m^3}$$

The term H_D comes out as a result of virtual process $e^- \to e^- + e^- + e^+$ in the region whose is scale is smaller than the Compton length $\lambda_C = \frac{1}{m} = \alpha a_0$ of electrons. Such virtual states appear in perturbation theory when one sums over intermediate states, which derive ultimately from a resolution of the identity. The effect is to smear out the position of the atomic electron over a distance of order λ_C. We have

$$H_D = \frac{1}{8m^2}\nabla^2 V$$

The term H_{SO} arises because the electric field of nuclei generates a magnetic field in the rest frame of electron. We have

$$H_{SO} = \frac{1}{2m^2}\frac{1}{r}\frac{dV}{dr}\boldsymbol{L}\cdot\boldsymbol{S}$$

The unperturbed energy levels in hydrogen are given by

$$E_n = -\frac{1}{2n^2}\frac{e^2}{4\pi a_0}$$

When spin of electron is taken into account, these levels are $2n^2$ degenerate. One choice of base is $|nlm_ls\rangle$. It is the eigenvector of operator L^2, L_z and S_z. However, L_z and S_z do not commute with H_{SO}. A better choice of base is $|nljm_j\rangle$. It is the eigenvector of operator L^2, J^2 and J_z. H_{SO}, H_{RKE} and H_{SO} are all commute with L^2, J^2 and J_z. So

$$\langle nl'j'm_j'|H|nljm_j\rangle$$

vanishes unless $l' = l$, $j = j'$ and $m_j' = m_j$. The final results are

$$\langle nljm_j|H_{RKE}|nljm_j\rangle = -\alpha^2 E_n \frac{1}{n^2}\left(\frac{3}{4} - \frac{n}{l+\frac{1}{2}}\right)$$

$$\langle nljm_j|H_D|nljm_j\rangle = -\alpha^2 E_n \frac{1}{n}\delta_{l0}$$

$$\langle nljm_j|H_{SO}|nljm_j\rangle = -\alpha^2 E_n \frac{1}{2n}\frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+\frac{1}{2})(l+1)}$$

When we add them up to get the total energy shift due to the fine structure we find

$$\Delta E_{FS} = -\alpha^2 E_n \frac{1}{n^2}\left(\frac{3}{4} - \frac{n}{j+\frac{1}{2}}\right)$$

It is independent of the orbital angular momentum quantum number l, although each of the individual terms does depend on l. However, the total energy shift does depend on j in addition to the principal quantum number n, so when we take into account the fine structure corrections, the energy levels of hydrogen atom have the form E_{nj}.

Besides fine structure effect, the remaining important effects causing energy shift are hyperfine effects and the Lamb shift.
The Lamb shift is a shift in the energy levels due to the interaction of the electron with the vacuum fluctuations of the quantized electromagnetic field. It has small effects on the s-states ($l = 0$) of hydrogen, thereby introducing a dependence of the energy levels on l. Thus, including the Lamb shift, the energy levels in hydrogen have the form E_{nlj} , and the only degeneracy is that due to rotational invariance. It will be further discussed in quantum electrodynamics.
Hyperfine effects are caused by the interaction between electro spin and nuclei spin, and will be discussed later.

19.2.3 Zeeman effect

The Zeeman effect concerns the interaction of atomic systems with external magnetic fields. The Hamiltonian for the electron in hydrogen atom is

$$H = \frac{(\boldsymbol{P} + e\boldsymbol{A})^2}{2m} - \frac{e^2}{4\pi r} + H_{FS} + g_e\mu_B \boldsymbol{S}\cdot\boldsymbol{B}$$

where $g_e \approx 2$ and $\mu_B \equiv \frac{e}{2m}$. We assume a uniform magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$. We take the gauge

$$\boldsymbol{A} = \frac{1}{2}\boldsymbol{B}\times\boldsymbol{r}$$

which is Coulomb gauge so that $\boldsymbol{\nabla} \times \boldsymbol{A} = 0$. This implies

$$\boldsymbol{P} \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{P}$$

so the cross terms in the expansion of the kinetic energy can be written in either order. We also notice that

$$\boldsymbol{P} \cdot \boldsymbol{A} = \frac{1}{2}\boldsymbol{P} \cdot \boldsymbol{B} \times \boldsymbol{r} = \frac{1}{2}\boldsymbol{B} \cdot \boldsymbol{L}$$

At last, we have

$$H = H_a + H_Z + H_B + H_{FS}$$

where

$$H_a = \frac{p^2}{2m} - \frac{e^2}{4\pi r} \quad H_Z = \frac{e}{2m}(L_z + 2S_z)B \quad H_B = \frac{e^2}{8m}B^2(x^2 + y^2)$$

Suppose the typical energy of the term H_i is E_i, then we have $E_a \sim \frac{me^4}{32n^2\pi^2\hbar^2\epsilon_0^2}$, $E_z \sim \frac{ne\hbar B}{2m}$ and $E_B \sim \frac{e^2}{8m}n^4a_0^2 = \frac{2n^4\pi^2\epsilon_0^2\hbar^4B^2}{m^3e^2}$. So

$$\frac{E_Z}{E_a} \sim \frac{16\pi^2n^3\hbar^3\epsilon_0^2}{m^2e^3}B \sim \frac{n^3B}{2 \times 10^5T}$$

$$\frac{E_B}{E_a} \sim \left[\frac{8\pi^2n^3\hbar^3\epsilon_0^2}{m^2e^3}B\right]^2 \sim \left[\frac{n^3B}{4 \times 10^5T}\right]^2$$

In the previous section, we have derived that

$$\frac{E_{FS}}{E_a} \sim \frac{3\alpha^2}{4n^2} \sim \frac{1}{2.5 \times 10^4n^2}$$

So,in the usual experimental condition, we have

$$E_B \ll E_z \ll E_a$$

So, in the following discussion, we will neglect E_B term. But whether $\frac{E_{FS}}{E_Z}$ is much larger than 1, much smaller than 1 or close to 1 depends on the B and n.

If H_{FS} is much smaller than H_Z and can be neglected, then we have

$$H = H_a + \frac{e}{2m}(L_z + 2S_z)B$$

The eigenvector of H is $|nlm_lm_s\rangle$ with eigenvalue $E = E_n + \mu_B B(m_l + 2m_s)$.
If H_{FS} is much smaller than H_Z and but cannot be neglected, we may treat it as a perturbation. For simplicity, we only take H_{SO} into account. Up to the first order, we consider the matrix element

$$\langle nlm_lm_s|f(r)\boldsymbol{L} \cdot \boldsymbol{S}|nl'm_l'm_s'\rangle$$

Since $[H_{SO}, L^2] = 0$, the term above vanishes unless $l = l'$. So we focus on the matrix in the subspace $l = l'$. For the $2p$ orbit of hydrogen, there is a 2-fold degeneracy between $|2, 1, -1, \frac{1}{2}\rangle$ and $|2, 1, 1, -\frac{1}{2}\rangle$. This makes one 2×2 matrix. Let us look at the off-diagonal element,

$$\langle 2, 1, -1, \frac{1}{2}|f(r)\boldsymbol{L} \cdot \boldsymbol{S}|2, 1, 1, -\frac{1}{2}\rangle$$

Figure 19.1: Zeeman effect for $n = 2$ in Hydrogen atom

in which we use the identity

$$\boldsymbol{L} \cdot \boldsymbol{S} = \frac{1}{2}(L_+S_- + L_-S_+) + L_zS_z$$

To be non-vanishing, the operator in the middle of the matrix element must connect states with $\Delta m_l = 2$, but in fact that operator permits only $\Delta_m = 0, \pm 1$. Therefore the off-diagonal matrix element vanishes and the energy shift is determined by diagonal elements. For $2p$ orbit, we can get

$$\Delta E = \langle nlm_lm_s|f(r)\boldsymbol{L} \cdot \boldsymbol{S}|nlm_lm_s\rangle \propto m_lm_s$$

As for the $2s$ levels, for them $\boldsymbol{L} = 0$ (that is, the operator $\boldsymbol{L}$ vanishes on the 2s-subspace), so $\Delta E = 0$.

The final case we shall examine is the weak field limit, in which $H_z \ll H_{FS}$ and we will treat H_z as perturbation.

Note: *In the case of hydrogen, one should also consider the Lamb shift for a realistic treatment. For example, in the $n = 2$ levels of hydrogen, the Lamb shift is about 10 times smaller than the fine structure energy shifts, indicating that we really should question how the Lamb shift compares to the Zeeman term which is also (by our assumptions) much smaller than the fine structure term.*

The eigenvector of $H_a + H_{FS}$ are $|nljm_j\rangle$ with eigenvalue E_{nj}. Up to the first order,the matrix elements we need have the form

$$\langle nl'jm_j'|H_z|nljm_j\rangle$$

Since $[H_Z, L^2] = 0$ and $[H_Z, J_z] = 0$, off-diagonal matrix element vanishes automatically.The energy shift is

$$\Delta E = \mu_B B\langle nljm_j|L_z + 2S_z|nljm_j\rangle = g_{eff}\mu_B Bm_j$$

where

$$g_{eff} = 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}$$

19.2.4 Hyperfine structure

The nucleus of an atom contains localized charge and current distributions, which produce electric and magnetic fields that can be decomposed into multipole fields much as in classical

electrostatics or magnetostatics. The first of the multipole moments, the electric monopole, is of course the Coulomb electrostatic field that holds the electrons in their orbits and produces the gross structure of the atom. The higher order multipole moments produce small corrections to the atomic structure that are known generally as hyperfine effects.
Multipole moments for a system of charges has been discussed in the part of classical electrodynamics. In quantum mechanics, we recall that the intrinsic magnetic moment operator of an electron is defined in the space of electron spin. So we may infer that multipole moments of nuclei is defined in the space of nuclear spin $\boldsymbol{I}$. We let the quantum number of the operator I^2 be i, so that I^2 has eigenvalues $i(i+1)$. The nuclear Hilbert space is a $(2i+1)$ dimensional space in which the standard basis is $|im_i\rangle$ with $-i \leq m \leq i$.
Not all the multipole fields that occur classically are allowed in the case of a nuclei. There are two rules governing the allowed multipole moments of the nucleus.
The first is that electric multipoles of odd k and magnetic mutlipoles of even k are forbidden. For example, if the nucleus had an electric dipole moment, the perturbing Hamiltonian would be

$$H_1 = -e\frac{\boldsymbol{d}\cdot\boldsymbol{r}}{r^3}$$

And just like $\boldsymbol{\mu}$,$\boldsymbol{d}$, must be proportional to the spin, say $\boldsymbol{d} = k\boldsymbol{I}$, because all vector operators on a single irreducible subspace are proportional (Wigner-Eckart theorem). Thus,

$$H_1 = -ke\frac{\boldsymbol{I}\cdot\boldsymbol{r}}{r^3}$$

We find that H_1 violates time reversal and parity

$$TH_1T^\dagger = -H_1 \quad PH_1P^\dagger = -H_1$$

The weak interactions do violate parity, and we do know that time reversal (more precisely, CP) is violated at a very small level in certain decay processes, so it is possible that the terms forbidden by this rule actually exist at a small level. For example, the neutron or the electron may have an electric dipole moment, but if such moments exist, they are certainly very small and can be neglected in our discussion.
The second rule states that a 2^k-pole can occur only if $k \leq 2i$. For example, the proton with $i = \frac{1}{2}$ can possess an electric monopole moment and a magnetic dipole moment, but not an electric quadrupole moment. Lying behind this rule is the fact that the operator representing the 2^k -pole on the nuclear Hilbert space is, in fact, an order k irreducible tensor operator. But the maximum order of an irreducible tensor operator on the nuclear Hilbert space with spin i is $k = 2i$.
So for hydrogen atom, whose nuclear spin is $i = \frac{1}{2}$, the only term we have to concern is magnetic moment. A point magnetic dipole of moment $\boldsymbol{\mu}$ is

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\boldsymbol{\mu}\times\boldsymbol{r}}{r^3}$$

To avoid the sigularuty at origin, we modify $\boldsymbol{A}(\boldsymbol{r})$ as

$$\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{\mu}\times\boldsymbol{r}\begin{cases}\frac{1}{a^3} & r < a\\ \frac{1}{r^3} & r > a\end{cases}$$

taking into account of the finite size of nuclei. By taking the curl we compute the magnetic field

$$\boldsymbol{B}(\boldsymbol{r}) = \begin{cases} \frac{2\boldsymbol{\mu}}{a^3} & r < a \\ \boldsymbol{\mu} \cdot \hat{T} & r > a \end{cases}$$

where $T_{ij} = \frac{3x_i x_j - r^2 \delta_{ij}}{r^5}$ We define

$$\Delta(r) \equiv \begin{cases} \frac{1}{a^3} & r < a \\ 0 & r > a \end{cases}$$

and

$$f(r) \equiv \begin{cases} 0 & r < a \\ 1 & r > a \end{cases}$$

Then we can write

$$\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{\mu} \times \boldsymbol{r} \left[\Delta(r) + \frac{f(r)}{r^3} \right]$$

and

$$\boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{\mu} \cdot \left[2\Delta(r)\hat{I} + f(r)\hat{T} \right]$$

In the limit $a \to 0$, we have

$$\lim_{a \to 0} \Delta(r) = \frac{4\pi}{3}\delta(\boldsymbol{r}) \quad \lim_{a \to 0} f(r) = 1$$

The Hamiltonian of the system is

$$H = \frac{(\boldsymbol{p} + e\boldsymbol{A})^2}{2m} - \frac{e^2}{4\pi r} + H_{FS} + H_{Lamb} + \frac{e}{m}\boldsymbol{S} \cdot B$$

The expressions for $\boldsymbol{A}$ and $\boldsymbol{B}$ are are given above and $\boldsymbol{\mu}$ in the expression are given by

$$\boldsymbol{\mu} = g_p \mu_p \boldsymbol{I}$$

where g_p is g-factor of proton and $\mu_p \equiv \frac{e}{2m_p}$. Thus, the Hamiltonian must be interpreted as an operator acting the total Hilbert space

$$\mathcal{H} = \mathcal{H}_{elec} \otimes \mathcal{H}_{nucl}$$

For $\mathcal{H}_{elec}$ the obvious basis is $|nljm_j\rangle$ with energies E_{nlj} when there is no hyperfine terms. In hydrogen energies depend on l because of the Lamb shift. The obvious basis in $\mathcal{H}_{nucl}$ is $|im_i\rangle$. Thus we define the basis states in $\mathcal{H}$ as $|nljm_jm_i\rangle$ (we suppress the index i since it is a constant).It is called uncoupled basis. Now we expand the Hamiltonian and neglect the term A^2, writing the result as $H = H_0 + H_1$, where

$$H_0 = \frac{p^2}{2m} - \frac{e^2}{4\pi r} + H_{FS} + H_{Lamb}$$

and

$$H_1 = 2\mu_B(\boldsymbol{p} \cdot A + \boldsymbol{S} \cdot \boldsymbol{B})$$

Since

$$\boldsymbol{p}\cdot(\boldsymbol{I}\cdot r)=\boldsymbol{I}\cdot(\boldsymbol{r}\cdot P)=\boldsymbol{I}\cdot\boldsymbol{L}$$

we have

$$H_{1,orbi}\equiv 2\mu_B(\boldsymbol{p}\cdot\boldsymbol{A})=k(\boldsymbol{I}\cdot\boldsymbol{L})\left[\Delta(r)+\frac{f(r)}{r^3}\right]$$

where $k\equiv= g_e g_p \mu_B \mu_p$. On the other hand, we have

$$H_{1,spin}\equiv 2\mu_B\boldsymbol{S}\cdot\boldsymbol{B}=k\left[2\Delta(r)\boldsymbol{I}\cdot\boldsymbol{S}+f(r)\boldsymbol{I}\cdot\hat{T}\cdot\boldsymbol{S}\right]$$

In the limit $a\to 0$, we have

$$H_{1,orbi}=k(\boldsymbol{I}\cdot\boldsymbol{L})\left[\frac{4\pi}{3}\delta(\boldsymbol{r})+\frac{1}{r^3}\right]$$

and

$$H_{1,spin}=k\left[\frac{8\pi}{3}\delta(\boldsymbol{r})\boldsymbol{I}\cdot\boldsymbol{S}+\boldsymbol{I}\cdot\hat{T}\cdot\boldsymbol{S}\right]$$

Since the coupling term (e.g.$\boldsymbol{I}\cdot S$) are not invariant under either electronic rotations alone or under nuclear rotations alone. The uncoupled basis is not the best one for carrying out the perturbation calculation. The total angular momentum of the system are defined by

$$\boldsymbol{F}\equiv\boldsymbol{I}+\boldsymbol{J}=\boldsymbol{I}+\boldsymbol{L}+\boldsymbol{S}$$

The new coupled basis is $|nljfm_f\rangle$. Since $[\boldsymbol{F},H]=0$, $|nljfm_f\rangle$ is the eigenvector of H. The energy shift caused by H_1 is

$$\Delta E=\langle nljfm_f|H_1|nljfm_f\rangle$$

After a lengthy calculation, we can get

$$\Delta E=\frac{g_e g_p\mu_B\mu_p}{4\pi a_0^3}\frac{1}{n^3}\frac{f(f+1)-j(j+1)-i(i+1)}{j(j+1)(2l+1)}$$

The energy levels were E_{nlj} before the hyperfine interactions were turned on, but since ΔE depends on f, they now have the form E_{nljf} . The energy eigenstates are $|nljfm_f\rangle$, and are $(2f+1)$-fold degenerate, causing the fine structure levels of hydrogen to split, giving rise to hyperfine multiplets.
For example, the ground state $|1,0,\frac{1}{2}\rangle$ splits into two levels $f=0$ and $f=1$. This $f=0$ level is the true ground state of hydrogen. It is nondegenerate. The $f=1$ level is 3-fold degenerate, and lies above the ground state by an energy of approximately $1.42GHz$ in frequency units, or $21cm$ in wave length units.

19.3 Time dependent perturbation theory

Dyson series

Time-dependent perturbation theory applies to Hamiltonians of the form

$$H=H_0+H_1(t)$$

where H_0 is solvable and H_1 is treated as a perturbation.In time-dependent perturbation theory, we are usually interested in time-dependent transitions between eigenstates of the unperturbed system induced by the perturbation H_1. Time-dependent transitions are usually described by the transition amplitude, defined as the quantity

$$\langle f|U(t)|i\rangle$$

where $U(t)$ is the exact time evolution operator for the Hamiltonian, and where $|i\rangle$ and $|f\rangle$ are two eigenstates of the unperturbed Hamiltonian H_0. Let us denote the unperturbed time-evolution operator by $U_0(t)$ and the exact one by $U(t)$. These operators satisfy the evolution equations

$$i\frac{\partial U_0(t)}{\partial t} = H_0 U_0(t) \quad i\frac{\partial U(t)}{\partial t} = HU(t)$$

Since H_0 is independent of time, we have $U_0 = e^{-iH_0 t}$.

Suppose the state in Schrödinger picture is $|\psi_S(t)\rangle$, then we define the state in interaction picture as

$$|\psi_I(t)\rangle \equiv U_0^\dagger(t)|\psi_S(t)\rangle$$

Similarly, we define the operator in interaction picture as

$$A_I(t) \equiv U_0^\dagger(t) A_S(t) U_0(t)$$

Let us define $W(t)$ as the operator that evolves kets in the interaction picture forward from time 0 to final time t, i.e.

$$|\psi_I(t)\rangle = W(t)|\psi_I(0)\rangle$$

We can verify that

$$W(t) = U_0^\dagger U(t)$$

And we can get the time evolution of $W(t)$

$$i\frac{\partial W}{\partial t} = H_{1I}(t)W(t)$$

The formal solution is

$$W(t) = I + (-i)^n \sum_{n=1}^{\infty} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n H_{1I}(t_1)H_{1I}(t_2)\cdots H_{1I}(t_n)$$

Let us assume for simplicity that H_0 has a discrete spectrum $H_0|n\rangle = E_n|n\rangle$. We assume that the system is initially in an eigenstate of the unperturbed system, what we will call the "initial" state $|i\rangle$ with energy E_i. Then

$$|\psi_I(t)\rangle = W(t)|i\rangle$$

Let us expand the exact solution of the Schrödinger equation in the interaction picture in the unperturbed eigenstates

$$|\psi_I(t)\rangle = \sum_n c_n(t)|n\rangle$$

We then have

$$c_n(t) = \langle n|W(t)|i\rangle = \langle n|U_0^\dagger U(t)|i\rangle = e^{iE_n t}\langle n|U(t)|i\rangle$$

So the transition amplitudes in the interaction picture and those in the Schrödinger picture are related by a simple phase factor.The transition probabilities are the squares of the amplitudes and are the same in either case. The perturbation expansion of the transition amplitude $c_n(t)$ is

$$c_n(t) = \delta_{ni} + c_n^{(1)}(t) + \cdots$$

where

$$c_n^{(1)}(t) = \frac{1}{i}\int_0^t dt'\langle n|H_{1I}(t')|i\rangle = \frac{1}{i}\int_0^t dt' e^{i(E_n-E_i)t'}\langle n|H_1(t')|i\rangle$$

Constant and harmonic perturbation

If H_1 is time-independent, we have

$$c_n^{(1)}(t) = \frac{2}{i}e^{i\omega_{ni}t/2}\left(\frac{\sin\omega_{ni}t/2}{\omega_{ni}}\right)\langle n|H_1|i\rangle$$

Up to the first order, the transition probability is

$$P_n(t) = 4\left(\frac{\sin^2\omega_{ni}t/2}{\omega_{ni}^2}\right)|\langle n|H_1|i\rangle|^2 \quad (n \neq i)$$

Another case that is important in practice is when H_1 has a periodic time dependence of the form

$$H_1(t) = Ke^{-i\omega_0 t} + K^\dagger e^{i\omega_0 t}$$

We can get

$$c_n^{(1)}(t) = \frac{2}{i}\left[e^{i(\omega_{ni}-\omega 0)t/2}\left(\frac{\sin(\omega_{ni}-\omega_0)t/2}{\omega_{ni}-\omega_0}\right)\langle n|K|i\rangle + e^{i(\omega_{ni}+\omega_0)t/2}\left(\frac{\sin(\omega_{ni}+\omega_0)t/2}{\omega_{ni}+\omega_0}\right)\langle n|K^\dagger|i\rangle\right]$$

Often, we are most interested in those final states to which most of the probability goes, which are the states for which one or the other of the two denominators is small. For these states we have

$$E_n \approx E_i \pm \omega_0$$

We call these two cases absorption and stimulated emission, respectively. Taking the case of absorption, and looking only at final states that are near resonance, we can write the transition probability to first order of perturbation theory as

$$P_n(t) = 4\left(\frac{\sin^2(\omega_{ni}-\omega_0)t/2}{(\omega_{ni}-\omega_0)^2}\right)|\langle n|K|i\rangle|^2$$

Transition probability

Let us fix the final state and examine how the probability develops as a function of time in first order time-dependent perturbation theory. Obviously $P_n(0) = 0$. At later times we see that $P_n(t)$ oscillates at frequency ω_{ni} between 0 and a maximum proportional to $\frac{1}{\omega_{ni}}$. The frequency ω_{ni} measures how far the final state is off resonance, that is, how much it fails to conserve energy. If this frequency is large, the probability oscillates rapidly between zero and a small maximum. But as we move the state closer to the initial state in energy, ω_{ni} gets smaller, the period of oscillations becomes longer, and the amplitude grows.
If there is a final state degenerate in energy with the initial state, then $\omega_{ni} = 0$ and the time-dependent factor takes on its limiting value $\frac{t^2}{4}$. In this case, first order perturbation theory predicts that the probability grows without bound. This is an indication of the fact that at sufficiently long times first order perturbation theory breaks down and we must take into account higher order terms in the perturbation expansion. But at short times it is correct that P_n for a state on resonance grows as t^2.

Now let us fix the time t and examine how the expression for $P_n(t)$ in first order perturbation theory depends on the energy of the final state. To do this we focus on $\frac{\sin^2(\omega t/2)}{\omega^2}$ as a function of ω.
The curve of the function consists of oscillations under the envelope $\frac{1}{\omega^2}$, with zeroes at $\omega = \frac{2n\pi}{t}$. The central lobe has height $\frac{t^2}{4}$ and width that is proportional to $\frac{1}{t}$, so the area of the central lobe is proportional to t. We can derive that

$$\lim_{t\to\infty} \frac{1}{t}\frac{\sin^2 \omega t/2}{\omega^2} = \frac{\pi}{2}\delta(\omega)$$

The δ-function enforces energy conservation in the limit $t \to \infty$. But at finite times, transitions take place to states in a range of energies about the initial energy. This width is of order $\frac{1}{t}$. This is an example of the energy-time uncertainty relation, $\Delta t \Delta E \sim 1$, indicating that a system that is isolated (not subjected to a measurement) over a time interval Δt has an energy that is uncertain by an amount $\Delta E \sim \frac{1}{\Delta t}$.
It is conventional to define the transition rate as the transition probability per unit time,i.e.

$$\Gamma(i \to f) = \lim_{t\to\infty} P(i \to f)t$$

Up to the first order, we have

$$\Gamma(i \to f) = 2\pi\delta(E_i - E_f)|V_{fi}|^2$$

where $V_{fi} = \langle f|H_1|i\rangle$.

19.4 Atomic Radiation

We apply time-dependent perturbation theory to the interaction of atomic electron with the classical radiation field. The basic Hamiltonian, with A^2 omitted, is

$$H = \frac{p^2}{2m} - e\phi(\boldsymbol{x}) + \frac{e}{m}\boldsymbol{A}\cdot\boldsymbol{p}$$

Specifically, we work with a linear polarized monochromatic field of the plane wave for

$$\boldsymbol{A} = 2A_0\boldsymbol{\epsilon}\cos(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t) = A_0\boldsymbol{\epsilon}[e^{i\boldsymbol{k}\cdot\boldsymbol{x}}e^{-i\omega t} + e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}e^{i\omega t}]$$

So, $\frac{e}{m}\boldsymbol{A}\cdot\boldsymbol{p}$ is a Harmonic perturbation with $K = \frac{eA_0}{m}e^{i\boldsymbol{k}\cdot\boldsymbol{x}}(\boldsymbol{\epsilon}\cdot\boldsymbol{p})$. For example, take the case of absorption, the transition rate is

$$\Gamma(i\to f) = 2\pi\frac{e^2|A_0|^2}{m^2}|\langle f|e^{i\boldsymbol{k}\cdot\boldsymbol{x}}(\boldsymbol{\epsilon}\cdot\boldsymbol{p})|i\rangle|^2\delta(E_f - E_i - \omega)$$

The cross section can be calculated by the transition rate of absorption.

The electrical dipole approximation is based on the fact that the wavelength of radiation field is far longer than the atomic dimension, so that the series

$$e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = 1 + ii\boldsymbol{k}\cdot\boldsymbol{x} + \cdots$$

can be approximated by its leading term, 1. Suppose the direction of the EM is wave is z, we now have

$$\langle f|e^{i\boldsymbol{k}\cdot\boldsymbol{x}}(\boldsymbol{\epsilon}\cdot\boldsymbol{p})|i\rangle = \langle f|p_z|i\rangle = \frac{m}{i}\langle f|[z, H_0]|i\rangle = im(E_n - E_i)\langle f|n|i\rangle$$

where $H_0 = \frac{p^2}{2m} - e\phi(\boldsymbol{x})$.
Recall the Wigner-Eckart theorem, we have the selection rules: The matrix element would vanish unless $\Delta m = 0$ and $\Delta l = \pm 1$ (Δl can be 0 because parity consideration. Further more, if the EM wave is circular polarized, we would have $\Delta m = \pm 1$. If the fine structure is taken into account, we would have $\Delta j = 0, \pm 1$ and $\Delta l = \pm 1$. If the hyperfine structure is taken into account, we would have $\Delta f = 0, \pm 1$, $\Delta j = 0, \pm 1$ and $\Delta l = \pm 1$.

19.5 Variational Method

Let H be a Hamiltonian which is assumed to have some bound states. Let the discrete (bound state) energy eigenvalues be ordered $E_0 < E_1 < \cdots$. The eigenvalues are allowed to be degenerate. There may also be a continuous spectrum above some energy, as often happens in practice.
Let $|\psi\rangle$ be any normalizable state. Then the theorem states that

$$\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} > E_0$$

The state $|\psi\rangle$ is chosen to be an approximation to the ground state, based on physical intuition or other criteria, and the upper bound on E_0 that is obtained is often quite useful.
In practice we often choose a continuous family of trial wave functions. Let λ be a continuous parameter, and let us write the family as $|\psi(\lambda)\rangle$. Then we define a function of λ, really an energy function,

$$F(\lambda) = \frac{\langle\psi(\lambda)|H|\psi(\lambda)\rangle}{\langle\psi(\lambda)|\psi(\lambda)\rangle}$$

We minimize this by finding the root λ_0 of

$$\frac{\partial F}{\partial \lambda} = 0$$

so that the best estimate to the ground state wave function out of the family is $|\psi(\lambda)\rangle$ and the best estimate for the ground state energy is $F(\lambda_0)$. Of course we must check that the root is actually a minimum.

In practice the normalization denominators are often inconvenient. One possibility is simply to normalize each member of the set of trial wave functions so those denominators are not present. But often it is easier to enforce normalization by using Lagrange multipliers. We introduce the function,

$$F(\lambda, \beta) = \langle \psi(\lambda)|H|\psi(\lambda)\rangle - \beta(\langle \psi(\lambda)|\psi(\lambda)\rangle - 1)$$

and then require

$$\frac{\partial F}{\partial \lambda} = 0 \quad \frac{\partial F}{\partial \beta} = 0$$

Chapter 20
Many body problem

20.1 Identical particles

If two particles are identical, their exchange must not change physical quantities, so we have

$$|\cdots\psi_j\cdots\psi_i\cdots\rangle = E_{ij}|\cdots\psi_i\cdots\psi_j\cdots\rangle = e^{i\theta}|\cdots\psi_i\cdots\psi_j\cdots\rangle$$

where E_{ij} is the operator to exchange particle i and j and psi_i is the quantum number to describe the state of particle i. In coordinate representation, it can be expressed as

$$\Psi(\cdots\boldsymbol{x}_j\cdots\boldsymbol{x}_i\cdots) = e^{i\theta}\Psi(\cdots\boldsymbol{x}_i\cdots\boldsymbol{x}_j\cdots)$$

In three dimension space, the value of $e^{i\theta}$ can only be ± 1. If the spin of the particle is integer, then the phase factor must be 1 and the particle is called boson. If the spin of the particle is half-integer, then the phase factor must be -1 and the particle is called fermion. This is called spin-statistics theorem and can only be proved by relativistic quantum field theory. Here, we will take it for granted.

Note: *In two dimension space, the phase $e^{i\theta}$ can be anything, and the particles that obey quantum statistics of this sort are called anyons. There is a brief introduction in chapter 12.1 and 12.2 of the book Quantum Field Theory and the Standard Model (Matthew D. Schwartz).*

So not all vectors in the space $\mathcal{H}_1\otimes\cdots\otimes\mathcal{H}_n$ are physical state. The physical states must be the eigenvectors of all E_{ij} with eigenvalue 1 (-1) for bosons (fermions). The space composed of physical states is the so-called Fock space. For example, suppose there are three particles with different state a, b, c. $|abc\rangle$ is not a physical state. The physical state for bosons is

$$\frac{1}{\sqrt{3!}}\left(|abc\rangle + |acb\rangle + |bca\rangle + |bac\rangle + |cab\rangle + |cba\rangle\right)$$

The physical state for fermions is

$$\frac{1}{\sqrt{3!}}\left(|abc\rangle - |acb\rangle + |bca\rangle - |bac\rangle + |cab\rangle - |cba\rangle\right)$$

In general, for N particles filling N distinct states, there are $N!$ states to start with, but there is only one totally symmetric state and one totally anti-symmetric state, and the rest of $N!-2$ states are thrown out. Therefore quantum statistics reduces the size of the Hilbert space quite dramatically.

Further more, not all Hermitian operators are physical observables. A nonphysical operator is one that takes a state in the physical subspace of the Hilbert space (one that satisfies the right

symmetry under exchange), and maps it into a nonphysical state (one that does not have the right symmetry). An example is the operator $\boldsymbol{X}_1$. We might call this the operator corresponding to the measurement of the position of particle 1. The problem with this operator from a physical standpoint is that you cannot measure the position of particle 1. You can select a region of space, and ask whether there is a particle in that region. But if you find one, you cannot say whether it is particle 1 or particle 2, since they are identical. If O is a physical observables, we must have

$$[E_{ij}, O] = 0$$

for all ij.

Figure 20.1: The path integral for three identical fermions

In path integral formulation of quantum mechanics, we calculate the transition amplitudes

$$\langle \boldsymbol{x}_f, t_f | \boldsymbol{x}_i, t_i \rangle = \int \mathcal{D}\boldsymbol{x}(t) e^{i\int_{t_i}^{t_f} dt L(t)}$$

The generalization to the N -particle case is

$$\langle \boldsymbol{x}_{1f} \cdots \boldsymbol{x}_{Nf}, t_f | \boldsymbol{x}_{1i} \cdots \boldsymbol{x}_{Ni}, t_i \rangle = \int \mathcal{D}\boldsymbol{x}_1(t) \cdots \boldsymbol{x}_N(t) e^{i\int_{t_i}^{t_f} dt L(t)}$$

Here, the particle 1 at the initial position $\boldsymbol{x}_{1i}$ moves to the final position $\boldsymbol{x}_{1f}$, the particle 2 at the initial position $\boldsymbol{x}_{2i}$ to $\boldsymbol{x}_{2f}$, etc, and you sum over all possible paths. When the particles are identical, however, we need to introduce proper (anti-)symmetry of the state. For fermions (The case for bosons can be obtained easily by dropping all minus signs), we introduce the anti-symmetrized position bra

$$\langle [\boldsymbol{x}_1 \cdots \boldsymbol{x}_N] | = \frac{1}{\sqrt{N!}} \sum_\sigma (-1)^\sigma \langle \boldsymbol{x}_{\sigma(1)} \cdots \boldsymbol{x}_{\sigma(N)} |$$

We now consider path integral representation of the transition amplitudes

$$\langle [\boldsymbol{x}_{1f} \cdots \boldsymbol{x}_{Nf}], t_f | [\boldsymbol{x}_{1i} \cdots \boldsymbol{x}_{Ni}], t_i \rangle$$

On the other hand, the Lagrangian for identical particles must be invariant under the exchange of particles, we can prove that

$$\langle [\boldsymbol{x}_{1f} \cdots \boldsymbol{x}_{Nf}], t_f | [\boldsymbol{x}_{1i} \cdots \boldsymbol{x}_{Ni}], t_i \rangle = \sum_\sigma \langle \boldsymbol{x}_{\sigma(1)f} \cdots \boldsymbol{x}_{\sigma(N)f}, t_f | \boldsymbol{x}_{1i} \cdots \boldsymbol{x}_{Ni}, t_i \rangle$$

In other words, the path integral sums over all possible paths allowing the positions at the final time slice are interchanged in all possible ways starting from the positions at the initial time slice.

20.2 Non-relativistic quantum field theory

20.2.1 Motivation and formulation of quantum field theory

There are some limitations of of multi-body Schrödinger wave function. Firstly, when the number of particles is large, multi-body Schrödinger wave function would be cumbersome. Secondly, it is incapable of describing processes where the number of particles changes.
The aim of the quantum field theory is to come up with a formalism which is completely equivalent to multi-body Schrödinger equations but just better: it allows you to consider a variable number of particles all within the same framework and can even describe the change in the number of particles. It also gives totally symmetric or anti-symmetric multi-body wave function automatically. It also allows a systematic way of organizing perturbation theory in terms of Feynman diagrams. It is particularly suited to multi-body problems.
In quantum mechanics, you start with classical particle Hamiltonian mechanics, with no concept of wave or interference. After quantizing it, we introduce Schrödinger wave function and there emerges concepts of wave and its interference. In quantum field theory, you start with classical wave equation, with no concept of particle. After quantizing it, we find particle interpretation of excitations in the system.

Let us consider a classical field equation

$$\left(i\frac{\partial}{\partial t}+\frac{\nabla^2}{2m}\right)\psi(\boldsymbol{x},t)=0$$

A solution to this field equation is that of a plane wave

$$\psi(\boldsymbol{x},t)=e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}$$

where $\omega=\frac{k^2}{2m}$. This classical field equation can be derived from the action

$$S=\int dtd\boldsymbol{x}\mathcal{L}\quad \mathcal{L}=\psi^*\left(i\frac{\partial}{\partial t}+\frac{\nabla^2}{2m}\right)\psi$$

We can even add a non-linear term in the action, for instance

$$\mathcal{L}=\psi^*\left(i\frac{\partial}{\partial t}+\frac{\nabla^2}{2m}\right)\psi-\frac{1}{2}\lambda\psi^{*2}\psi^2$$

Variation method gives a non-linear field equation

$$\left(i\frac{\partial}{\partial t}+\frac{\nabla^2}{2m}-\lambda\psi^*\psi\right)\psi(\boldsymbol{x},t)=0$$

Now we quantize the Schrödinger field by canonical quantization method. A more formal discussion on the motivation and validity of canonical quantization will be discussed in relativistic quantum field theory.
The canonically conjugate momenta is

$$\pi(\boldsymbol{x})=\frac{\partial\mathcal{L}}{\partial\dot{\psi}(\boldsymbol{x})}=i\psi^\dagger(\boldsymbol{x})$$

We now introduce canonical commutation relation

$$[\psi(\boldsymbol{x}), \pi(\boldsymbol{y})] = i\delta(\boldsymbol{x} - \boldsymbol{y})$$

so

$$[\psi(\boldsymbol{x}), \psi^\dagger(\boldsymbol{y})] = \delta(\boldsymbol{x} - \boldsymbol{y})$$

We now regard $\psi(\boldsymbol{x})$ as annihilation operator and $\psi(\boldsymbol{x})$ creation operator of a particle at position $\boldsymbol{x}$.
The Hamiltonian of the system is

$$H = \int d\boldsymbol{x} \left(-\psi^\dagger \frac{\nabla^2}{2m} \psi + \frac{1}{2} \lambda \psi^{\dagger 2} \psi^2 \right)$$

We have

$$[H, \psi^\dagger] = -\frac{\nabla^2}{2m} \psi^\dagger + \lambda \psi^{\dagger 2} \psi$$

20.2.2 Particles in quantum field theory

We define the "vacuum" $|0\rangle$ which is annihilated by the annihilation operator

$$\psi(\boldsymbol{x})|0\rangle = 0$$

and construct the Fock space by

$$|\boldsymbol{x}_1 \cdots \boldsymbol{x}_N\rangle = \frac{1}{\sqrt{N!}} \psi^\dagger(\boldsymbol{x}_1) \cdots \psi^\dagger(\boldsymbol{x}_N)|0\rangle$$

The state $|\boldsymbol{x}_1 \cdots \boldsymbol{x}_N\rangle$ is an n-particle state of identical bosons in the position eigenstate at $\boldsymbol{x}_1 \cdots \boldsymbol{x}_N$. We will verify this interpretation below explicitly.

Let us look at the one-particle state

$$|x\rangle = \psi^\dagger(\boldsymbol{x})|0\rangle$$

We can derive that

$$\langle \boldsymbol{y}|\boldsymbol{x}\rangle = \delta(\boldsymbol{x} - \boldsymbol{y})$$

Therefore, this state is normalized in the same way as the one-particle position eigenstate in quantum mechanics. We define a general one-particle state in the quantum field theory as

$$|\Psi(t)\rangle \equiv \int d\boldsymbol{x} \Psi(\boldsymbol{x}, t) \psi^\dagger(\boldsymbol{x})|0\rangle$$

$\psi(\boldsymbol{x})$ corresponds to the Schrödinger wave function in the particle quantum mechanics. The Schrödinger equation in quantum field theory is

$$i\frac{\partial|\Psi(t)\rangle}{\partial t} = H|\psi(t)\rangle$$

Since we can derive that

$$\begin{aligned} H|\Psi(t)\rangle &= \int d\boldsymbol{x}\Psi(\boldsymbol{x},t)H\psi^\dagger(\boldsymbol{x})|0\rangle \\ &= \int d\boldsymbol{x}\Psi(\boldsymbol{x},t)[H,\psi^\dagger](\boldsymbol{x})|0\rangle \\ &= \int d\boldsymbol{x}\left(-\frac{\nabla^2}{2m}\Psi(\boldsymbol{x},t)\right)|\boldsymbol{x}\rangle \end{aligned}$$

So we have

$$i\frac{\partial\Psi(\boldsymbol{x},t)}{\partial t} = -\frac{\nabla^2\Psi(\boldsymbol{x},t)}{2m}$$

This is exactly the Schrödinger equation in coordinate representation of quantum mechanics for one isolate particle.

Let us next study the two-particle state

$$|\boldsymbol{x}_1\boldsymbol{x}_2\rangle = \frac{1}{\sqrt{2}}\psi^\dagger(\boldsymbol{x}_1)\psi^\dagger(\boldsymbol{x}_2)|0\rangle$$

We can derive that

$$\langle\boldsymbol{x}_1\boldsymbol{x}_2|\boldsymbol{y}_1\boldsymbol{y}_2\rangle = \frac{1}{2}\left(\delta(\boldsymbol{x}_1-\boldsymbol{y}_1)\delta(\boldsymbol{x}_2-\boldsymbol{y}_2)+\delta(\boldsymbol{x}_1-\boldsymbol{y}_2)\delta(\boldsymbol{x}_2-\boldsymbol{y}_1)\right)$$

This normalization suggests that we are dealing with a two-particle state of identical particles, because the norm is non-vanishing when $\boldsymbol{x}_1=\boldsymbol{y}_1$ and $\boldsymbol{x}_2=\boldsymbol{y}_2$, but also when $\boldsymbol{x}_1=\boldsymbol{y}_2$ and $\boldsymbol{x}_2=\boldsymbol{y}_1$, i.e., two particles interchanged. A general two-particle state is given by

$$|\Psi(t)\rangle \equiv \int d\boldsymbol{x}_1 d\boldsymbol{x}_2\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)\psi^\dagger(\boldsymbol{x}_1)\psi^\dagger(\boldsymbol{x}_2)|0\rangle$$

Because $[\psi^\dagger(\boldsymbol{x}_1),\psi^\dagger(\boldsymbol{x}_2)]$, the integration over $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is symmetric under the interchange of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and hence $\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)=\Psi(\boldsymbol{x}_2,\boldsymbol{x}_1,t)$. The symmetry under the exchange suggests that we are dealing with identical bosons.
Since

$$\begin{aligned} H|\Psi(t)\rangle &= \frac{1}{\sqrt{2}}\int d\boldsymbol{x}_1 d\boldsymbol{x}_2\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)H\psi^\dagger(\boldsymbol{x}_1)\psi^\dagger(\boldsymbol{x}_2)|0\rangle \\ &= \frac{1}{\sqrt{2}}\int d\boldsymbol{x}\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)([H,\psi^\dagger(\boldsymbol{x}_1)]\psi^\dagger(\boldsymbol{x}_2)+\psi^\dagger(\boldsymbol{x}_1)[H,\psi^\dagger(\boldsymbol{x}_2)])|0\rangle \\ &= \int d\boldsymbol{x}_1 d\boldsymbol{x}_2\left(-\frac{\nabla_1^2}{2m}-\frac{\nabla_2^2}{2m}+\lambda\delta(\boldsymbol{x}_1-\boldsymbol{x}_2)\right)\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)|\boldsymbol{x}_1\boldsymbol{x}_2\rangle \end{aligned}$$

So we have

$$i\frac{\partial\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)}{\partial t} = \left(-\frac{\nabla_1^2}{2m}-\frac{\nabla_2^2}{2m}+\lambda\delta(\boldsymbol{x}_1-\boldsymbol{x}_2)\right)\Psi(\boldsymbol{x}_1,\boldsymbol{x}_2,t)$$

It is a Schrödinger equation for two-particle wave function with a delta function potential as an interaction between them. Therefore, the Fock space with two creation operators correctly

describes the two-particle quantum mechanics.
If we want a general interaction potential between them, the action must be modified to

$$S = \int dt \left[\int d\boldsymbol{x} \psi^\dagger(\boldsymbol{x}) \left(i\frac{\partial}{\partial t} + \frac{\nabla^2}{2m} \right) \psi(\boldsymbol{x}) - \frac{1}{2} \int d\boldsymbol{x} d\boldsymbol{y} \psi^\dagger(\boldsymbol{x}) \psi^\dagger(\boldsymbol{y}) V(\boldsymbol{x}-\boldsymbol{y}) \psi(\boldsymbol{x}) \psi(\boldsymbol{y}) \right]$$

The corresponding Hamiltonian is

$$H = \int d\boldsymbol{x} \psi^\dagger \frac{-\nabla^2}{2m} \psi + \frac{1}{2} \int d\boldsymbol{x} d\boldsymbol{y} \psi^\dagger(\boldsymbol{x}) \psi^\dagger(\boldsymbol{y}) V(\boldsymbol{x}-\boldsymbol{y}) \psi(\boldsymbol{x}) \psi(\boldsymbol{y})$$

Similarly, we can derive the equation

$$i\frac{\partial \Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)}{\partial t} = \left(-\frac{\nabla_1^2}{2m} - \frac{\nabla_2^2}{2m} + V(\boldsymbol{x}_1 - \boldsymbol{x}_2) \right) \Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$$

Generally, for the n-particle state, we have

$$i\frac{\partial \Psi(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n, t)}{\partial t} = \left(\sum_i -\frac{\nabla_i^2}{2m} + \sum_{i<j} V(\boldsymbol{x}_i - \boldsymbol{x}_j) \right) \Psi(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n, t)$$

So it is reasonable that an n-particle state can be constructed as

$$|\Psi(t)\rangle = \frac{1}{\sqrt{n!}} \int d\boldsymbol{x}_1 \cdots d\boldsymbol{x}_n \Psi(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n) \psi^\dagger(\boldsymbol{x}_1) \cdots \psi^\dagger(\boldsymbol{x}_n) |0\rangle$$

If we are interested in a system in a background potential. A good example is the electrons in an atom, where all of them are moving in the background Coulomb potential due to the nucleus. In this case, the correct field-theory Hamiltonian

$$H = \int d\boldsymbol{x} \psi^\dagger(\boldsymbol{x}) \left(\frac{-\nabla^2}{2m} - \frac{Ze^2}{4\pi|\boldsymbol{x}|} \right) \psi(\boldsymbol{x}) + \frac{1}{2} \int d\boldsymbol{x} d\boldsymbol{y} \psi^\dagger(\boldsymbol{x}) \psi^\dagger(\boldsymbol{y}) \frac{e^2}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} \psi(\boldsymbol{x}) \psi(\boldsymbol{y})$$

The total number of particles is an eigenvalue of the operator

$$N \equiv \int d\boldsymbol{x} \psi^\dagger(\boldsymbol{x}) \psi(\boldsymbol{x})$$

We can get the commutation relation

$$[N, \psi] = -\psi \quad [N, \psi^\dagger] = \psi^\dagger$$

So we can derive

$$N|\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\rangle = n|\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\rangle$$

20.2.3 Momentum space

We define creation and annihilation operators in the momentum space as

$$a(\boldsymbol{p}) \equiv (2\pi)^{-3/2} \int d\boldsymbol{x}\psi(\boldsymbol{x})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} \quad a^\dagger(\boldsymbol{p}) \equiv (2\pi)^{-3/2} \int d\boldsymbol{x}\psi^\dagger(\boldsymbol{x})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}$$

So we can derive the commutation relation

$$[a(\boldsymbol{p}), a^\dagger(\boldsymbol{q})] = \delta(\boldsymbol{p}-\boldsymbol{q}) \quad [a(\boldsymbol{p}), a(\boldsymbol{q})] = [a^\dagger(\boldsymbol{p}), a^\dagger(\boldsymbol{q})] = 0$$

We can rewrite the Hamiltonian in the momentum space. The free part of the Hamiltonian is

$$H_0 = \int d\boldsymbol{x}\psi^\dagger \frac{-\nabla^2}{2m}\psi = \int d\boldsymbol{p}\frac{p^2}{2m}a^\dagger(\boldsymbol{p})a(\boldsymbol{p})$$

The free Hamltonian simply countes the number of particles in a given momentum state and assigns the energy $\frac{p^2}{2m}$ accordingly.
The interaction Hamltonian is

$$\begin{aligned}\Delta H &= \frac{1}{2}\int d\boldsymbol{x}d\boldsymbol{y}\psi^\dagger(\boldsymbol{x})\psi^\dagger(\boldsymbol{y})V(\boldsymbol{x}-\boldsymbol{y})\psi(\boldsymbol{x})\psi(\boldsymbol{y}) \\ &= \frac{1}{2}\int d\boldsymbol{p}d\boldsymbol{q}d\boldsymbol{p}'d\boldsymbol{q}'V(\boldsymbol{p}-\boldsymbol{q})a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{p}')a(\boldsymbol{q})a(\boldsymbol{q}')\delta(\boldsymbol{p}+\boldsymbol{p}'-\boldsymbol{q}-\boldsymbol{q}')\end{aligned}$$

where

$$V(\boldsymbol{p}-\boldsymbol{q}) = \frac{1}{(2\pi)^3}\int d\boldsymbol{x}V(\boldsymbol{x})e^{-i(\boldsymbol{p}-\boldsymbol{q})\cdot\boldsymbol{x}}$$

The delta function represents the momentum conservation in the scattering process due to the potential V. The potential term of Hamiltonian causes scattering, by annihilating two particles in momentum states $\boldsymbol{q}$, $\boldsymbol{q}'$ and create them in different momentum states $\boldsymbol{p}$, $\boldsymbol{p}'$ with the amplitude $V(\boldsymbol{p}-\boldsymbol{q})$.

20.2.4 Fermions

We have seen that the quantized Schrödinger field gives multi-body states of identical bosons. For fermions, we should use anti-commutation relations rather than commutation relations.

$$\{\psi(\boldsymbol{x}), \psi^\dagger(\boldsymbol{y})\} = \delta(\boldsymbol{x}-\boldsymbol{y}) \quad \{\psi(\boldsymbol{x}), \psi(\boldsymbol{y})\} = \{\psi^\dagger(\boldsymbol{x}), \psi^\dagger(\boldsymbol{y})\} = 0$$

One noteworthy point is that $\psi^\dagger(\boldsymbol{x}) = \frac{1}{2}\{\psi^\dagger(\boldsymbol{x}), \psi^\dagger(\boldsymbol{x})\} = 0$. What this means is that one cannot create two particles at the same position, an expression of Pauli's exclusion principle for fermions.
Here are a few useful identities. Similarly to the identity of commutators $[A, BC] = A[B, C] + [A, B]C$, we find

$$[A, BC] = \{A, B\}C - B\{A, C\} \quad [AB, C] = A\{B, C\} - \{A, C\}B$$

So we can get

$$[H, \psi^\dagger(\boldsymbol{x})] = -\frac{\nabla^2}{2m}\psi^\dagger(\boldsymbol{x}) + \int d\boldsymbol{y}\psi^\dagger(\boldsymbol{x})\psi^\dagger(\boldsymbol{y})V(\boldsymbol{x}-\boldsymbol{y})\psi(\boldsymbol{y})$$

It is the same result as in the case of bosons! Consider a two-particle state

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \int d\boldsymbol{x}_1 d\boldsymbol{x}_2 \Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, t) \psi^\dagger(\boldsymbol{x}_1) \psi^\dagger(\boldsymbol{x}_2) |0\rangle$$

From the anti-commutation relations, we have

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = -\Psi(\boldsymbol{x}_2, \boldsymbol{x}_1)$$

Such a state indeed describes identical fermions. And we can get the same Schrödinger equation as in the case for bosons.

Chapter 21
Scattering theory

21.1 Lippmann–Schwinger equation

Imagine a particle coming in and getting scattered by a short-ranged potential $V(x)$ located around the origin $x \sim 0$. The time-independent Schrödinger equation is simply

$$(H_0 + V)|\psi\rangle = E|\psi\rangle$$

Here, $H_0 = \frac{p^2}{2m}$ is the free-particle Hamiltonian operator. We can write the solution as

$$|\psi^{(\pm)}\rangle = \frac{1}{E - H_0 \pm i\epsilon} V|\psi^{(\pm)}\rangle + |\phi\rangle$$

Here, $H_0|\phi\rangle = E|\phi\rangle$. In coordinate representation,

$$\psi^{(\pm)}(\mathbf{x}) = \phi(\mathbf{x}) + \int d^3x' \langle \boldsymbol{x}| \frac{1}{E - H_0 \pm i\epsilon} |\boldsymbol{x}'\rangle V(\boldsymbol{x}')\psi^{(\pm)}(\boldsymbol{x}')$$

Here, $\phi(\boldsymbol{x}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{(2\pi)^{\frac{3}{2}}}$. Define the Green function as

$$G_\pm(\boldsymbol{x}, \boldsymbol{x}') \equiv \frac{1}{2m} \langle \boldsymbol{x}| \frac{1}{E - H_0 \pm i\epsilon} |\boldsymbol{x}'\rangle$$

We can derive that

$$G_\pm(\boldsymbol{x}, \boldsymbol{x}') = -\frac{1}{4\pi} \frac{e^{\pm ik|\boldsymbol{x}-\boldsymbol{x}'|}}{|\boldsymbol{x} - \boldsymbol{x}'|}$$

where $k = \sqrt{2mE}$. And it is easy to show that

$$(\nabla^2 + k^2) G_\pm(\boldsymbol{x}, \boldsymbol{x}') = \delta(\boldsymbol{x} - \boldsymbol{x}')$$

So, we have

$$\psi^{(\pm)}(\boldsymbol{x}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{(2\pi)^{\frac{3}{2}}} - 2m \int d^3x' \frac{1}{4\pi} \frac{e^{\pm ik|\boldsymbol{x}-\boldsymbol{x}'|}}{|\boldsymbol{x} - \boldsymbol{x}'|} V(\boldsymbol{x}')\psi^{(\pm)}(\boldsymbol{x}')$$

We now can interpret $\psi^+(\boldsymbol{x})$ as a superposition of incident plane wave and scattered wave which propagate from scatterer to outside region. From now on, we will denote it as $\psi(\boldsymbol{x})$. The experiment is done typically by placing the detector far away from the scatterer $|\boldsymbol{x}| \ll a$ where a is the "size" of the scatterer. The integration over $\boldsymbol{x}'$, on the other hand, is limited

within the "size" of the scatterer because of the $V(\boldsymbol{x}')$ factor. Therefore, we are in the situation $|\boldsymbol{x}| \ll |\boldsymbol{x}'|$, and hence can use the approximation

$$|\boldsymbol{x} - \boldsymbol{x}'| \approx |\boldsymbol{x}| - \frac{\boldsymbol{x}' \cdot \boldsymbol{x}}{|\boldsymbol{x}|}$$

Under this limit,

$$\psi(\boldsymbol{x}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{(2\pi)^{\frac{3}{2}}} - 2m\frac{e^{ikr}}{4\pi r}\int d^3x' e^{-i\boldsymbol{k}'\cdot\boldsymbol{x}'} V(\boldsymbol{x}')\psi(\boldsymbol{x}')$$

Here, $r = |\boldsymbol{x}|$ and $\boldsymbol{k}' = k\frac{\boldsymbol{x}}{r}$. It is customary to write this equation in the form

$$\psi(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\left(e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + f(\boldsymbol{k},\boldsymbol{k}')\frac{e^{ikr}}{r}\right)$$

Here,

$$f(\boldsymbol{k},\boldsymbol{k}') \equiv -\frac{m}{2\pi}(2\pi)^3\langle\boldsymbol{k}'|V|\psi\rangle$$

Recall the definition of cross section

$$\sigma \equiv \frac{\text{Number of Events}}{\text{Time} \times \text{Incident Flux}}$$

So, the differential cross section for particles being scattered into the solid angle is

$$d\sigma = \frac{|\boldsymbol{j}_{\text{scatt}}|r^2 d\Omega}{|\boldsymbol{j}_{\text{inc}}|} = |f(\boldsymbol{k},\boldsymbol{k}')|^2 d\Omega$$

In a more realistic situation, we should use wave packets to describe the scattering process. The basic picture is a free wave packet approaches the scattering center. After a long time, we have both the original wave packet moving in the original direction plus a spherical wave front that moves outward. The details can be found in the section 3 of the lecture notes *Scattering Theory I (Hitoshi Murayama)*.

Furthermore, if we require that the normalization of the wave function should always satisfy $\int dx^3|\psi(\boldsymbol{x})|^2 = 1$ for any t, as guaranteed by the unitarity of time evolution operator. This requirement leads to a special requirement on the scattered wave, and hence $f(\boldsymbol{k},\boldsymbol{k}')$, from witch we can derive the optical theorem.

Theorem 21.1 Optical theorem

$$\text{Im} f(\theta = 0) = \frac{k\sigma_{\text{tot}}}{4\pi}$$

where

$$f(\theta = 0) \equiv f(\boldsymbol{k},\boldsymbol{k}),$$

the setting of $\boldsymbol{k} \equiv \boldsymbol{k}'$ imposes scattering in the forward direction, and

$$\sigma_{\text{tot}} = \int \frac{d\sigma}{d\Omega} d\Omega$$

The meaning of this theorem is clear. Because the scattered wave takes the probability away to different directions, the total probability for the particle to go to the forward direction (unscattered) should decrease. This decrease is caused by the interference between the unscattered and scattered waves and hence is proportional to $f(0)$. On the other hand, the amount of decrease in the forward direction should equal the total probability at other directions, which is proportional to the total cross section. The proof can be found in the section 4 of the lecture notes *Scattering Theory I (Hitoshi Murayama)*.

21.2 Born approximation

If $|\psi\rangle = |\phi\rangle + O(V)$ is close to $|\phi\rangle$, we can solve the Lippmanmn-Schwinger equation by perturbation theory. The lowest order approximation in V is

$$|\psi\rangle = \frac{1}{E - H_0 + i\epsilon} V|\phi\rangle + |\phi\rangle$$

This is called Born approximation. In coordinate representation,

$$f^{(1)}(\boldsymbol{k}, \boldsymbol{k}') = -\frac{m}{2\pi} \int d^3x V(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}}$$

Here, $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}'$. If the potential is central, we can derive that

$$f^{(1)}(\boldsymbol{k}, \boldsymbol{k}') = -\frac{2m}{q} \int_0^\infty dr\, rV(r) \sin(qr)$$

Yukawa potential

$$V = \frac{\alpha}{r} e^{-\mu r}$$

So, we can derive

$$f(\theta) = -\frac{2m\alpha}{q^2 + \mu^2}$$

Different cross section is therefore given by

$$\frac{d\sigma}{d\Omega} = (2m\alpha)^2 \frac{1}{[2k^2(1 - \cos\theta) + \mu^2]^2}$$

The total cross section is obtained by integrating over $d\Omega$,

$$\sigma = (2m\alpha)^2 \frac{4\pi}{4k^2\mu^2 + \mu^4}$$

Coulomb potential

$$V = \frac{\alpha}{r}$$

Take the limit $\mu \to 0$, we can get

$$f(\theta) = -\frac{2m\alpha}{q^2}$$

Different cross section is given by

$$\frac{d\sigma}{d\Omega} = (\frac{\alpha}{4E})^2 \frac{1}{\sin^4 \frac{\theta}{2}}$$

The total cross section diverges. The divergence is in the $\cos\theta$ integral when $\theta \to 0$. In other words, the divergence occurs for the small momentum transfer $q \to 0$, which corresponds to large distances. The reason why the total cross section diverges is because the Coulomb potential is actually a long-range force. No matter how far the incident particles are from the charge, there is always an effect on the motion of the particles and they get scattered.

Form factor

If the source of Coulomb potential has an distribution $\rho_N(\boldsymbol{x})$, then

$$V(\boldsymbol{x}) = \int d^3x \frac{\alpha}{|\boldsymbol{x} - \boldsymbol{x}'|} \rho(\boldsymbol{x}')$$

Note that the potential is mathematically a convolution of the Coulomb potential and the probability density. Since the first Born amplitude is nothing but the Fourier transform of the potential, the convolution becomes a product of Fourier transforms, one for the Coulomb potential and the other for the probability density. So

$$f(\theta) = f(\theta)_{\text{pointlike}} F(q)$$

Here,

$$F(q) \equiv \int d^3x \rho_N(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}},$$

being called form factor.

Born expansion

Define T-matrix by

$$V|\psi\rangle = T|\phi\rangle$$

Using the definition of the T-matrix, we find

$$f(\boldsymbol{k}, \boldsymbol{k}') = -\frac{m}{2\pi}(2\pi)^3 \langle \boldsymbol{k}'|T|\boldsymbol{k}\rangle$$

Using the Lippmann–Schwinger equation and multiplying the both sides by V from left, we find

$$T|\phi\rangle = V \frac{1}{E - H_0 + i\epsilon} T|\phi\rangle + V|\phi\rangle$$

A formal solution to the T-matrix is

$$T = \frac{1}{1 - V\frac{1}{E - H_0 + i\epsilon}} V$$

By Taylor expanding this operator in geometric series, we find

$$T = V + V\frac{1}{E - H_0 + i\epsilon} V + V\frac{1}{E - H_0 + i\epsilon} V \frac{1}{E - H_0 + i\epsilon} V + \cdots$$

So,

$$|\psi\rangle = \left(1 + \frac{1}{E - H_0 + i\epsilon}V + \frac{1}{E - H_0 + i\epsilon}V\frac{1}{E - H_0 + i\epsilon}V + \cdots\right)|\phi\rangle$$

The first term is the wave which did not get scattered. The second term is the wave that gets scattered at a point in the potential and then propagates outwards by the propagator. In the third term, the wave gets scattered at a point in the potential, propagates for a while, and gets scattered again at another point in the potential, and propagates outwards. In the $n+1$-th term, there are n times scattering of the wave before it propagates outwards.

21.3 Partial wave analysis

Partial wave expansion

When the potential is **central**, angular momentum is conserved due to Noether's theorem. Therefore, we can expand the wave function in the eigenstates of the angular momentum. Obtained waves with definite angular momenta are called partial waves. We can solve the scattering problem for each partial wave separately, and then in the end put them together to obtain the full scattering amplitude. The plane wave can be expanded as follows.

$$e^{ikz} = \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta)$$

Here, $j_l(kr)$ is spherical Bessel functions of first kind. The asymptotic behaviour of $j_l(kr)$ at large r can be written as

$$j_l(kr) \sim \frac{\sin(kr - \frac{l\pi}{2})}{kr}$$

so,

$$e^{ikz} \sim \frac{1}{2ikr}\sum_{l=0}^{\infty}(2l+1)(e^{ikr} - (-1)^l e^{-ikr})P_l(\cos\theta)$$

Meanwhile, the f factor can be expanded as

$$f(\theta) = \sum_{l=0}^{\infty} f_l(2l+1)P_l(\cos\theta)$$

Optical theorem constraint

The cross section can be represented by expansion coefficient of f factor as

$$\sigma = 4\pi\sum_l(2l+1)|f_l|^2$$

On the other hand,

$$\mathrm{Im}f(0) = \sum_l(2l+1)\mathrm{Im}f_l$$

From optical theorem we can derive that

$$|f_l|^2 = \frac{1}{k}\mathrm{Im}f_l$$

This constraint can be rewritten as

$$|1 + 2ikf_l|^2 = 1$$

So we can define a phase δ_l as

$$1 + 2ikf_l = e^{2i\delta_l}$$

or equivalently,

$$f_l = \frac{1}{k} e^{i\delta_l} \sin(\delta_l)$$

Phase shifts

We can derive the asymptotic behaviour of the wave function as

$$\psi(\boldsymbol{x}) \sim \frac{1}{2ikr} \sum_l (2l+1) P_l(\cos\theta) [e^{ikr} e^{2i\delta_l} - (-1)^l e^{-ikr}]$$

Compare it to the case of the plane wave without scattering. What this equation says is that the wave converging on the scatterer has the well-defined phase factor $-(-1)^l$, the same as in the case without scattering. On the other hand, the wave that emerges from the scatterer has an additional phase factor $e^{2i\delta_l}$. All what scattering did is to shift the phase of the emerging wave by $2\delta_l$. The reason why this is merely a phase factor is the conservation of probability. What converged to the origin must come out with the same strength. But this shift in the phase causes the interference among all partial waves different from the case without the phase shifts, and the result is not a plane wave but contains the scattered wave.
In terms of the phase shifts, the cross section is given by

$$\sigma = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l$$

Actual calculation of phase shifts is basically to solve the Schrödinger equation for each partial waves,

$$\left[-\frac{1}{r} \frac{d^2}{dr^2} r + \frac{l(l+1)}{r^2} + 2mV(r) \right] R_l(r) = k^2 R_l(r)$$

After solving the equation, we take the asymptotic limit $r \to \infty$, and write $R_l(r)$ as a linear combination of $j_l(kr) \cos\delta_l - n_l(kr) \sin\delta_l$. The relative coefficients of j_l and n_l determines the phase shift δ_l, and hence the cross section.

Hard sphere scattering

The potential for hard sphere scattering is

$$V = \begin{cases} 0 & (r > a) \\ \infty & (r < a) \end{cases}$$

The infinite potential corresponds to the boundary condition $R_l(a) = 0$. We first analyze the S-wave ($l = 0$). The Schrödinger equation is simply

$$-\frac{d^2 r R_0(r)}{dr^2} = k^2 r R_0(r)$$

The solution is

$$rR_0(r) = c\sin(k(r-a)) = \frac{ce^{ika}}{2i}\left[e^{i(kr-2ka)} - e^{-ikr}\right]$$

So we have $\delta_0 = -ka$. The reason behind the phase shift is that the wave cannot penetrate into $r < a$, the wave is shifted outwards, which is the shift in the phase $-ka$. The cross section from the S-wave scattering is

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2 ka$$

Let us generalize the discussion to the case of a little bit penetrable potential

$$V = \begin{cases} 0 & (r > a) \\ V_0 & (r < a) \end{cases}$$

Define $K^2 \equiv 2mV_0$. If $k > K$, we have

$$rR_0 = \begin{cases} \sin\sqrt{k^2 - K^2}r & (r < a) \\ \sin(kr + \delta_0) & (r > a) \end{cases}$$

By matching the logarithmic derivatives of the wave function at $r = a$, we find

$$\delta_0 = \tan^{-1}\left[\frac{k}{\sqrt{k^2 - K^2}}\tan\sqrt{k^2 - K^2}a\right] - ka$$

For $k \gg K$, one can neglect K and the phase shift vanishes. The energy is too large to care the slight potential and there is no scattering any more.
If $k < K$, we have

$$rR_0 = \begin{cases} \sinh\sqrt{K^2 - k^2}r & (r < a) \\ \sin(kr + \delta_0) & (r > a) \end{cases}$$

The phase shift is obtained as

$$\delta_0 = \tan^{-1}\left[\frac{k}{\sqrt{K^2 - k^2}}\tanh\sqrt{K^2 - k^2}a\right] - ka$$

For $k \ll K$, we have

$$\delta_0 \sim ka\left[\frac{1}{Ka}\tanh Ka - 1\right]$$

The phase shift δ_0 always starts linearly with k at small momentum, and the slope is negative. This is a completely general result for a repulsive potential, and a convenient quantity

$$a_0 = \lim_{k\to 0} -\frac{d\delta_0}{dk}$$

is called the scattering length, as it has the dimension of the length. This quantity basically measures how big the scatterer is. The cross section at $k \to 0$ limit is then given by $4\pi a_0^2$. For the hard sphere potential, the scattering length is indeed the size of the sphere.

For the hard sphere problem, the phase shifts for higher partial waves can be worked out similarly. We have

$$\tan\delta_l = \frac{j_l(ka)}{n_l(ka)}$$

The cross section is then given by

$$\sigma = \frac{4\pi}{k^2}\sum_l (2l+1)\sin^2\delta_l = \sum_{l=0}^{\infty}\frac{4\pi(2l+1)}{k^2}\frac{[j_l(ka)]^2}{[j_l(ka)]^2+[n_l(ka)]^2}$$

For small momenta $k \ll a^{-1}$, we can use the power expansion of the spherical Bessel function

$$j_l(r) \sim \frac{r^l}{(2l+1)!!} \quad n_l(r) \sim -\frac{(2l-1)!!}{r^{l+1}}$$

and get that

$$\delta_l \sim (ka)^{2l+1}$$

So phase shift (and so cross section) is smaller for higher partial waves. This is easy to understand. When k is small, the centrifugal barrier does not allow the particle to reach $r = a$ classically. Therefore the effect of the potential is extremely suppressed.
On the other hand at high monenta, $\sin^2\delta_l$ oscillates between 0 and 1 as a function of l up to $l \sim ka$. Above this value, the phase shift drops rapidly to zero. This makes sense from the classical physics intuition. When $l > ka$, the impact parameter is larger than the size of the target and there should not be any scattering.

21.4 Resonance

21.4.1 Attractive potential

As for attractive potential scattering

$$V = \begin{cases} 0 & (r > a) \\ -V_0 & (r < a) \end{cases}$$

the phase shift for S-wave is

$$\delta_0 = \tan^{-1}\left[\frac{k}{\sqrt{k^2+K^2}}\tan\sqrt{k^2+K^2}a\right] - ka$$

The scattering length is

$$a_0 = a\left[1 - \frac{\tan Ka}{Ka}\right]$$

For small K, the scattering length is negative. This is easy to understand because the wave is pulled into the potential rather pushed out unlike the repulsive case. However, once we make the potential more attractive (larger K), the scattering length grows and becomes even infinite

at $K = \frac{\pi}{2}$.
Let us study the analytic structure of the scattering amplitude more carefully. We have

$$e^{2i\delta_0} = e^{-2ika} \frac{1 + i\frac{k}{\sqrt{k^2+K^2}} \tan\sqrt{k^2+K^2}a}{1 - i\frac{k}{\sqrt{k^2+K^2}} \tan\sqrt{k^2+K^2}a}$$

It can have a pole if

$$1 - i\frac{k}{\sqrt{k^2+K^2}} \tan\sqrt{k^2+K^2}a = 0$$

This equation appears impossible to satisfy, but it can be on the complex plane of k. For a pure imaginary $k = i\kappa$, the equation becomes

$$\kappa = -\frac{\sqrt{K^2-\kappa^2}}{\tan\sqrt{K^2-\kappa^2}a}$$

This is the condition for bound states (the scattering wave e^{ike} becomes $e^{-\kappa r}$, which is trapped by scatter). By decreasing K from a sufficiently large value with bound states, the bound state energies $E = -\frac{\kappa^2}{2m}$ move up. When $Ka = (n + \frac{1}{2})\pi$, $\tan Ka \to \infty$, and we find a bound state approaching $\kappa = k = 0$. This is when the scattering length diverges. In other words, the infinite scattering cross section at $k = 0$ happens because there is a bound state exactly at $k = 0$.
This can also be seen on the complex k plane in the following manner. The lower half plane is unphysical as it corresponds to an exponentially growing wave function at the infinity for the scattered wave. When there are bound states, we see poles along the positive imaginary axis. By decreasing K, the poles along the positive imaginary axis go down, and a pole reaches the origin. By further decreasing K, the pole goes below the origin into the unphysical region. However, the existence of a pole just below the origin makes the scattering amplitude at $k \to 0$ large and results in an anomalously large cross section.

21.4.2 Delta-shell potential

As for delta shell scattering

$$V = \gamma\delta(r-a)$$

the phase shift for S-wave is

$$e^{2i\delta_0} = e^{-2ika} \frac{\sin ka + \frac{k}{2m\gamma}e^{ika}}{\sin ka + \frac{k}{2m\gamma}e^{-ika}}$$

The poles satisfies that

$$e^{2ika} = 1 - 2i\frac{k}{2m\gamma}$$

When γ is large, we can get

$$k \approx \frac{n\pi}{a + \frac{1}{2m\gamma}} - i\left(\frac{1}{2m\gamma}\right)^2 \frac{n^2\pi^2}{a^3} + O(\gamma^{-2})$$

The poles are in the unphysical lower half plane. But when γ is large, the poles are very close to the real axis, and the scattering amplitude receives a large enhancement due to these poles. In the limit of $\gamma \to \infty$, or in other words in the limit of no coupling between the regions inside and outside the shell, they become poles along the real axis. They are the discrete states inside the shell in this limit. By making γ finite, we introduce coupling between the discrete states inside the shell to the continuum states outside the shell.

It is instructive to solve Schrödinger equation for the values of k which correspond to the location of poles. Because the outgoing wave e^{ikr} is enhanced relative to the incoming wave e^{-ikr} by an infinite amount due to the pole, the boundary condition is that the solution is "purely outgoing", i.e.

$$rR_0 = \begin{cases} \sin kr & (r < a) \\ \sin ka e^{ik(r-a)} & (r > a) \end{cases} \qquad \mathrm{Re}(k) > 0$$

Because the factor $e^{ik(r-a)}$ grows exponentially at large r due to the negative imaginary part in k, the solution is not a regular normalizable solution. In the large γ limit, $\sin ka \sim O(\gamma^{-1})$ is small. Therefore the wave function almost vanishes at the shell. Outside the shell, the wave function oscillates at the small amplitude $\sin(ka)$, which however starts growing again due to the $e^{ik(r-a)}$ factor exponentially.
We now put the time dependence in. The energy eigenvalue is $E = \frac{k^2}{2m}$, where k is at the pole. If the pole is at

$$k = k_0 - i\kappa$$

the energy eigenvalue is at

$$E = E_0 - i\frac{\Gamma}{2} = \frac{k_0^2}{2m} - i\frac{k_0\kappa}{m} + O(\kappa^2)$$

For instance in the large γ limit

$$E \sim \frac{1}{2m}\left(\frac{n\pi}{a + \frac{1}{2m\gamma}}\right)^2 - i\left(\frac{n\pi}{2ma}\right)^3 \frac{2}{\gamma^2 a} + O(\gamma^{-3})$$

The time dependence of the wave function is simply

$$rR_0(r,t) = rR_0(r)e^{-iE_0t}e^{-\frac{\Gamma t}{2}} = \begin{cases} \sin kr\, e^{-iE_0t}e^{-\frac{\Gamma t}{2}} & (r < a) \\ \sin ka\, e^{ik(r-a)}e^{-iE_0t}e^{-\frac{\Gamma t}{2}} & (r > a) \end{cases}$$

Inside the shell, it shows an exponentially decaying probability density uniformly over space. Outside the shell, the probability density is $|rR_0|^2 \propto e^{2\kappa r - \Gamma t}$, which shows the probability flowing out to infinity with speed $\Gamma/2\kappa = \frac{k_0}{m}$, nothing but the velocity of the particle itself. In other words, the wave function describes a "bound state" inside shell decaying into a continuum state outside the shell moving away at the expected velocity. The resonances can be viewed as quasi-bound states which decay into continuum states. The lifetime of the quasi-bound states is $\tau = \frac{1}{\Gamma}$. A more rigorous treatment of resonance using wave packets can be found in the section 5 of the lecture notes *Scattering Theory III (Hitoshi Murayama)*.

21.4.3 General description of resonances

In general, once we know that there is a pole just below the real axis, we can approximate the phase shift by the contribution from the pole only, ignoring a continuum:

$$e^{2i\delta_l} \sim \frac{g(E)}{E - E_0 + i\Gamma/2}$$

Because of the unitarity $|e^{2i\delta_l}|^2 = 1$, we immediately conclude

$$e^{2i\delta_l} \sim e^{2i\theta} \frac{E - E_0 - i\Gamma/2}{E - E_0 + i\Gamma/2}$$

Ignoring the continuum contribution $e^{2i\theta}$,

$$\sin^2 \delta_l = \frac{\Gamma^2/4}{(E - E_0)^2 + \Gamma^2/4}$$

Figure 21.1: Resonance curve

We then can establish the relationship between the life time of the quasi-bound state and the FWHM of the resonance

$$\tau \Delta E \sim 1$$

21.5 Two-to-two scattering

Similar to the discussion in hydrogen atom, we take as independent variables the center of mass and relative coordinates of the particles

$$\boldsymbol{X} = \frac{m_1 \boldsymbol{x}_1 + m_2 \boldsymbol{x}_2}{m_1 + m_2} \qquad \boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$$

The corresponding momentum operators are

$$\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2 \quad \boldsymbol{p} = \frac{m_1 \boldsymbol{p}_1 - m_2 \boldsymbol{p}_2}{m_1 + m_2}$$

Then the Hamiltonian becomes

$$H = \frac{P^2}{2M} + \frac{p^2}{2\mu} + V(|\boldsymbol{x}|)$$

Then the problem reduces to the potential scattering problem for a particle of mass $\mu = \frac{m_1 m_2}{m_1+m_2}$.

If two particles that scatter are identical particles, such as electron-electron scattering or scattering of two identical atoms, symmetry of the wave function needs to be considered. Under the interchange of two particles, the center of mass motion is not affected, but the relative coordinates change their signs. If they have spins, their spins need to be interchanged at the same time.
If two particles are identical spinless bosons, there is no spin degrees of freedom and the interchange of particles is simply $\boldsymbol{x} \to -\boldsymbol{x}$ in the wave function. Because they are bosons, the wave function should not change under the interchange of particles, and hence the wave function must be an even function of $\boldsymbol{x}$. Therefore the asymptotic form of the wave function must be changed to

$$\psi(\boldsymbol{x}) \to e^{ikz} + e^{-ikz} + [f(\theta) + f(\pi - \theta)]\frac{e^{ikr}}{r}$$

The differential cross section is then found to be

$$\frac{d\sigma}{d\Omega} = |f(\theta) + f(\pi - \theta)^2|$$

Note that one should not integrate over the entire solid angle to obtain the total cross section because (θ, ϕ) and $(\pi - \theta, \phi + \pi)$ correspond to an identical state.
For two spin $\frac{1}{2}$ fermions, there are two possible spin wave functions, symmetric $S = 1$ and anti-symmetric $S = 0$. Therefore depending on the spin wave function, we either have a anti-symmetric or symmetric spatial wave function, respectively. In particular, the differential cross section is the same as the spinless bosons for the anti-symmetric spin wave function $S = 0$ while it is

$$\frac{d\sigma}{d\Omega} = |f(\theta) - f(\pi - \theta)^2|$$

for the symmetric spin wave function $S = 1$.

21.6 Time-dependent formulation

Recall the time-dependent perturbation theory, the rate of the initial state $|i\rangle$ to transform to the final state $|f\rangle$ to the first order in perturbation V is

$$\Gamma(i \to f) = 2\pi\delta(E_i - E_f)|V_{fi}|^2$$

When applied to the scattering problem, an additional issue is to define how we sum over the final states. In particular, we would like to sum over the continuum plane-wave states, and we must make the sum well-defined. To define the sum over the continuum states, it is useful to consider the system in a cube of size L. We impose the periodic boundary condition. The plane-wave solutions in this box are given by

$$\psi_{n_x,n_y,n_z}(\boldsymbol{x}) = \frac{1}{L^{3/2}} e^{2\pi i(n_x x + n_y y + n_z z)/L}$$

In the limit $L \to \infty$, we have

$$\sum_{n_x,n_y,n_z} = \left(\frac{L}{2\pi}\right)^3 \sum_{n_x,n_y,n_z} \left(\frac{2\pi}{L}\right)^3 \to \left(\frac{L}{2\pi}\right)^3 \int d\boldsymbol{p} = \int \frac{d\boldsymbol{x}d\boldsymbol{p}}{(2\pi)^3}$$

Coming back to the scattering problem, we sum over the final states to define the probability of the outgoing particle to go into various momentum states

$$\sum_f \Gamma(i \to f) = \int \frac{L^3 d\boldsymbol{p}}{(2\pi)^3} 2\pi\delta(E_i - E_f)|V_{fi}|^2$$

Note that

$$V_{fi} = \int d\boldsymbol{x} \frac{e^{-i\boldsymbol{p}_f\cdot\boldsymbol{x}}}{L^{3/2}} V(\boldsymbol{x}) \frac{e^{i\boldsymbol{p}_i\cdot\boldsymbol{x}}}{L^{3/2}} = \frac{1}{L^3} \int d\boldsymbol{x} V(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}}$$

where $\boldsymbol{q} = \boldsymbol{p}_i - \boldsymbol{p}_f$. Since the incident flux of the incoming particle is $\frac{v}{L^3}$, the cross section is

$$\sigma = \frac{L^3}{v} \sum_f \Gamma(i \to f) = \frac{m}{p_i} \int \frac{d\boldsymbol{p}}{(2\pi)^3} 2\pi\delta(E_i - E_f) \left| \int d\boldsymbol{x} V(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}} \right|^2$$

Note that $E = \frac{p^2}{2m}$ and $\delta(E_i - E_f) = \frac{m}{p_i}\delta(p_f - p_i)$, finally we can get

$$\sigma = \int d\Omega \left| \frac{m}{2\pi} \int d\boldsymbol{x} V(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}} \right|^2$$

This is nothing but the Born approximation for the cross section.

Part V

Quantum Field Theory

Chapter 22
From classical field to quantum field

22.1 Motivation

The state of the field is described by an element $|\psi\rangle$ in Hilbert space. The measurement of the field is described by an operator field $\phi_a(\boldsymbol{x})$. The mean value of the measurement of the field is described by Erenfest's theorem

$$\frac{d\langle\psi|\phi_a(\boldsymbol{x})|\psi\rangle}{dt} = -i\langle\psi|[\phi_a(\boldsymbol{x}), H_S]|\psi\rangle$$

If $[\phi_a(\boldsymbol{x}), H_S]_Q = i[\phi_a(\boldsymbol{x}), H_S]_C$, we can reproduce the classical field theory as an average effect of quantum field theory. We also note that the commutation relation here $[A, B] = AB - BA$ obeys the same operation laws as the Poisson bracket in classical field theory. So, what we need here is

1. the canonical quantization

$$[\phi_a(\boldsymbol{x}), \phi_b(\boldsymbol{y})] = 0 \quad [\pi^a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] = 0 \quad [\phi_a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] = i\delta_a^b\delta(\boldsymbol{x} - \boldsymbol{y})$$

2. the same definition of $\mathcal{L}$,π^a and H as those in corresponding classical theory.

In Heisenberg picture, we define

$$A_H \equiv U^{-1}(t, t_0)A_S U(t, t_0)$$

Especially,

$$\phi_a(x) \equiv U^{-1}(t, t_0)\phi_a(\boldsymbol{x})U(t, t_0)$$
$$\pi^a(x) \equiv U^{-1}(t, t_0)\pi^a(\boldsymbol{x})U(t, t_0)$$

If $A = f(\phi_a(\boldsymbol{x}), \pi^a(\boldsymbol{x}), t)$, then we have $A_H = f(\phi_a(x), \pi^a(x), t)$.
The canonical quantization can be generalized to the field operator in any time

$$[\phi_a(\boldsymbol{x}, t), \phi_b(\boldsymbol{y}, t)] = 0 \quad [\pi^a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)] = 0 \quad [\phi_a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)] = i\delta_a^b\delta(\boldsymbol{x} - \boldsymbol{y})$$

The dynamics of the field can be describe by Heisenberg equation

$$\frac{d\phi_a(x)}{dt} = -i[\phi_a(x), H_H]$$
$$\frac{d\pi^a(x)}{dt} = -i[\pi^a(x), H_H]$$

whose form will be equivalent to the classical field equation.

22.2 Lorentz invariance in quantum field theory

$$|\bar{\psi}\rangle = U(\Lambda)|\psi\rangle$$

Scalar fields:

$$\langle\bar{\psi}|\phi(x)|\bar{\psi}\rangle = \langle\psi|\phi(\Lambda^{-1}x)|\psi\rangle$$

$$U^{-1}(\Lambda)\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x)$$

Vector fields:

$$\langle\bar{\psi}|A^\mu(x)|\bar{\psi}\rangle = \langle\psi|\Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)|\psi\rangle$$

$$U^{-1}(\Lambda)A^\mu(x)U(\Lambda) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$$

Lorentz invariance means Lagrangian must be a scalar, or more loosely, action must be invariant under Lorentz transformation.

22.3 Momentum

The definition of momentum is the same as that in classical theory:

$$T^{\mu\nu} \equiv -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial^\nu\phi_a + \eta^{\mu\nu}\mathcal{L} \quad \partial_\mu T^{\mu\nu} = 0$$

and

$$P^\mu \equiv \int T^{0\mu}d^3x \quad \frac{dP^\mu}{dt} = 0$$

$$P^0 = H, \quad P^i = \int -\pi^a\partial^i\phi_a d^3x$$

We can get the commutation relation

$$\begin{aligned}
[\phi_a, P^\mu] &= -i\partial^\mu\phi_a \\
[\pi^a, P^\mu] &= -i\partial^\mu\pi^a \\
[P^\mu, P^\nu] &= 0
\end{aligned}$$

We define the translation operator $T(s)$ by

$$T^{-1}(s)\phi_a(x)T(s) = \phi_a(x-s)$$

then we can derive that

$$T(\epsilon) = 1 - i\epsilon_\mu P^\mu \quad T(s) = e^{-iP^\mu s_\mu}$$

22.4 Angular Momentum

The definition of angular momentum is the same as that in classical theory:

$$M^{\mu\nu\rho} \equiv x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} (\Sigma^{\nu\rho})_a{}^b \phi_b$$

and

$$M^{\nu\rho} \equiv \int M^{0\nu\rho} d^3x \quad \frac{dM^{\nu\rho}}{dt} = 0$$

$$M^{\mu\nu} = \int (x^\mu T^{0\nu} - x^\nu T^{0\mu} - \pi^a (\Sigma^{\mu\nu})_a{}^b \phi_b) d^3x$$

We also define that

$$M_L^{\mu\nu} \equiv \int (x^\mu T^{0\nu} - x^\nu T^{0\mu}) d^3x \quad M_S^{\mu\nu} \equiv \int (-\pi^a (\Sigma^{\mu\nu})_a{}^b \phi_b) d^3x$$

$$(L^{\mu\nu})_a{}^b \equiv -i(x^\mu \partial^\nu - x^\nu \partial^\mu)\delta_a{}^b \quad (S^{\mu\nu})_a{}^b \equiv -i(\Sigma^{\mu\nu})_a{}^b$$

We can get the commutation relation

$$M^{\mu\nu} = M_L^{\mu\nu} + M_S^{\mu\nu}$$

$$[\phi_a, M_L^{\mu\nu}] = (L^{\mu\nu})_a{}^b \phi_b \quad [\phi_a, M_S^{\mu\nu}] = (S^{\mu\nu})_a{}^b \phi_b$$

$$[\pi^a, M_L^{\mu\nu}] = (L^{\mu\nu})_b{}^a \pi^b \quad [\pi^a, M_S^{\mu\nu}] = -(S^{\mu\nu})_b{}^a \pi^b$$

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(-\eta^{\nu\rho} M^{\mu\sigma} + \eta^{\sigma\mu} M^{\rho\nu} + \eta^{\mu\rho} M^{\nu\sigma} - \eta^{\sigma\nu} M^{\rho\mu})$$

We again define $J_i \equiv \frac{1}{2}\epsilon_{ijk} M^{jk}$ and $K_i \equiv M^{i0}$, the communication relation can be written as

$$\begin{aligned} [J_i, J_j] &= i\epsilon_{ijk} J_k \\ [J_i, K_j] &= i\epsilon_{ijk} K_k \\ [K_i, K_j] &= -i\epsilon_{ijk} J_k \end{aligned}$$

Further more,

$$[P^\mu, M^{\rho\sigma}] = i(\eta^{\mu\sigma} P^\mu - \eta^{\mu\rho} P^\sigma)$$

$$\begin{aligned} [J_i, H] &= 0 \\ [J_i, P_j] &= i\epsilon_{ijk} P_k \\ [K_i, H] &= iP_i \\ [K_i, P_j] &= i\delta_{ij} H \end{aligned}$$

At last, we also define $L_i \equiv \frac{1}{2}\epsilon_{ijk} M_L^{jk}$ and $S_i \equiv \frac{1}{2}\epsilon_{ijk} M_S^{jk}$. So

$$\begin{aligned} [L_i, S_j] &= 0 \\ [S_i, P_j] &= 0 \\ [L_i, P_j] &= i\epsilon_{ijk} P_k \end{aligned}$$

We define the rotation operator $U(\Lambda)$ by

$$U^{-1}(\Lambda)\phi_a(x)U(\Lambda) = \mathcal{S}_a{}^b\phi_b(\Lambda^{-1}x)$$

where

$$\mathcal{S}_a{}^b = \delta_a{}^b + \frac{i}{2}\delta\omega_{\alpha\beta}(S^{\alpha\beta})_a{}^b$$

We can derive that

$$U(1+\delta\omega) = 1 + \frac{i}{2}\delta\omega_{\mu\nu}M^{\mu\nu} \quad U(\Lambda) = e^{\frac{i}{2}\theta_{\mu\nu}M^{\mu\nu}}$$

$$U^{-1}(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu{}_\nu P^\nu$$

$$U^{-1}(\Lambda)M^{\mu\nu}U(\Lambda) = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma M^{\rho\sigma}$$

22.5 Anticommutation relation

We define the anticommutation relation of operators as $\{A, B\} \equiv AB + BA$. Suppose that the field operator and its canonical momentum operator has the following anticommutation relation

$$\{\phi_a(\boldsymbol{x},t), \phi_b(\boldsymbol{y},t)\} = 0 \quad \{\pi^a(\boldsymbol{x},t), \pi^b(\boldsymbol{y},t)\} = 0 \quad \{\phi_a(\boldsymbol{x},t), \pi^b(\boldsymbol{y},t)\} = i\delta_a^b\delta(\boldsymbol{x}-\boldsymbol{y})$$

If the operator A composes of the terms like $\pi^a\mathcal{E}_a{}^b\phi_b$, we can show that the value of $[\phi_a, A]$ and $[\pi^a, A]$ is the same as those in the theory quantized with commutation relation. It is easy to verify that the operator P^i and $M_S^{\mu\nu}$ has the required form. The form of H is determined by the specific theory. As we can see later, the Hamiltonian of Dirac field has the required form. When it is quantized with anticommutation relation, the commutation relation between field operator and momentum, angular momentum discussed in previous section will be hold automatically.

Chapter 23
Spin 0 Field

23.1 Klein-Gordon field

Lagrangian

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 + \Omega_0$$

Field equation

$$(\partial^\mu\partial_\mu - m^2)\phi = 0$$

Hamiltonian

$$\pi = \dot{\phi}$$
$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 - \Omega_0$$
$$H = \int \mathcal{H} d^3x$$

Momentum and angular momentum

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}(\frac{1}{2}\partial^\mu\phi\partial_\mu\phi + \frac{1}{2}m^2\phi^2 - \Omega_0)$$
$$P^0 = H \quad P^i = \int -\pi\nabla^i\phi d^3x$$
$$J_k = \int -\pi\epsilon_{ijk}x^j\nabla^k\phi d^3x$$

23.2 Canonical quantization Formulation

Canonical quantization

$$\begin{aligned}
[\phi(\boldsymbol{x},t),\phi(\boldsymbol{y},t)] &= 0 \\
[\pi(\boldsymbol{x},t),\pi(\boldsymbol{y},t)] &= 0 \\
[\phi(\boldsymbol{x},t),\pi(\boldsymbol{y},t)] &= i\delta(\boldsymbol{x}-\boldsymbol{y})
\end{aligned}$$

Fourier expansion

$$\phi(\boldsymbol{x},t) = \int \widetilde{dk}\left[a(\boldsymbol{k})e^{ikx} + a^\dagger(\boldsymbol{k})e^{-ikx}\right]$$

$$\pi(\boldsymbol{x},t) = -i\int \widetilde{dk}\omega\left[a(\boldsymbol{k})e^{ikx} - a^\dagger(\boldsymbol{k})e^{-ikx}\right]$$

Here, $k^2 = \boldsymbol{k}^2 - \omega^2 = -m^2$, $kx = \boldsymbol{k}\cdot\boldsymbol{x} - \omega t$, $\widetilde{dk} = \frac{d^3k}{(2\pi)^3 2\omega}$

$$a(\boldsymbol{k}) = \int d^3x e^{-ikx}(i\pi + \omega\phi)$$

$$a^\dagger(\boldsymbol{k}) = \int d^3x e^{ikx}(-i\pi + \omega\phi)$$

$$\begin{aligned}
[a(\boldsymbol{p}), a(\boldsymbol{q})] &= 0 \\
\left[a^\dagger(\boldsymbol{p}), a^\dagger(\boldsymbol{q})\right] &= 0 \\
\left[a(\boldsymbol{p}), a^\dagger(\boldsymbol{q})\right] &= (2\pi)^3 2\omega\delta(\boldsymbol{p}-\boldsymbol{q})
\end{aligned}$$

Operator represented by a and $a^\dagger$

$$H = \int \widetilde{dk}\,\omega\, a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + (\mathcal{E}_0 - \Omega_0)V \quad \mathcal{E}_0 = \frac{1}{2}(2\pi)^{-3}\int d^3k\,\omega$$

$$P^i = \int \widetilde{dk}\, k^i\, a^\dagger(\boldsymbol{k})a(\boldsymbol{k})$$

Particles

$$[H, a(\boldsymbol{k})] = -\omega a(\boldsymbol{k}) \quad [H, a^\dagger(\boldsymbol{k})] = \omega a^\dagger(\boldsymbol{k})$$

$$[P^i, a(\boldsymbol{k})] = -k^i a(\boldsymbol{k}) \quad [P^i, a^\dagger(\boldsymbol{k})] = k^i a^\dagger(\boldsymbol{k})$$

Let $|p\rangle = a^\dagger(\boldsymbol{p})|0\rangle$,so

$$H|p\rangle = \omega_p|p\rangle \quad P^i|p\rangle = p^i|p\rangle$$

So, we interpret the state $|\boldsymbol{p}\rangle$ as the momentum eigenstate of a single particle of mass m. We can also show that $J_i|\boldsymbol{p}=0\rangle = 0$, so the particle carries no internal angular momentum.

Causality

The amplitude for a particle to propagate from y to x is $\langle 0|\phi(x)\phi(y)|0\rangle$, denoted by $D(x-y)$.

$$D(x-y) = \int \widetilde{dk}e^{ik(x-y)}$$

$$[\phi(x), \phi(y)] = D(x-y) - D(y-x)$$

If $x-y$ is space-like, a continuous Lorentz transformation can take $(x-y)$ to $-(x-y)$. So $[\phi(x), \phi(y)] = 0$ for space-like $x-y$. A measurement performed at one point can not affect a measurement at another point whose separation is space-like.

The Klein-Gordon propagator

$$D_R(x-y) \equiv \theta(x^0-y^0)\langle 0|[\phi(x)\phi(y)]|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+m^2}e^{ip(x-y)}$$

Figure 23.1: Retarded Green Function

$$(\partial^2 - m^2)D_R(x-y) = i\delta(x-y)$$

$$D_F(x-y) \equiv \langle 0|T\phi(x)\phi(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+m^2-i\epsilon}e^{ip(x-y)}$$

Here, T stands for time ordering, placing all operators evaluated at later times to the left.

Figure 23.2: Feynman Green Function

23.3 Perturbation theory for canonical quantization

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4$$

$$H = H_0 + H_{int} \quad H_{int} = \int d^3x \frac{\lambda_0^4}{4!}\phi^4(\boldsymbol{x})$$

23.3.1 Perturbation expansion of correlation functions

Note: *The ground state of the interaction field theory is denoted by $|\Omega\rangle$, the ground state of the free field theory is denoted by $|0\rangle$. The zero of energy is defined by $H_0|0\rangle = 0$ and $E_0 = \langle\Omega|H|\Omega\rangle$.*

$$\phi(t_0,\boldsymbol{x}) = \int \frac{d^3p}{(2\pi)^3}(a(\boldsymbol{p})e^{i\boldsymbol{p}\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}})$$

$$\phi(t,\boldsymbol{x}) = e^{iH(t-t_0)}\phi(t_0,\boldsymbol{x})e^{-iH(t-t_0)}$$

$$\phi_I(t,\boldsymbol{x}) \equiv e^{iH_0(t-t_0)}\phi(t_0,\boldsymbol{x})e^{-iH_0(t-t_0)}$$

$$H_I(x) = \int d^3x \frac{\lambda_0^4}{4!}\phi_I^4$$

The perturbation expansion of correlation functions is

$$\langle\Omega|T\{\phi(x)\phi(y)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle}$$

The proof can be found in chapter 4.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

Theorem 23.1 Wick's theorem

$T\{\phi(x_1)\phi(x_2)\cdots\phi(x_n)\} = N\{\phi(x_1)\phi(x_2)\cdots\phi(x_n) +$ all possible contractions $\}$ ♣

Normal order : all the a's are to the right of all the $a^\dagger$.

The proof can be found in chapter 4.3 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

Example:

$$\begin{aligned}\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\}|0\rangle &= D_F(x_1-x_2)D_F(x_3-x_4)\\ &+ D_F(x_1-x_3)D_F(x_2-x_4)\\ &+ D_F(x_1-x_4)D_F(x_2-x_3)\end{aligned}$$

23.3.2 Feynman diagram

Expand $\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle$ to the first order of λ_0

$$\begin{aligned}&\langle 0|T\left\{\phi_I(x)\phi_I(y)\frac{-i\lambda_0}{4!}\int dz^4\phi_I(z)\phi_I(z)\phi_I(z)\phi_I(z)\right\}|0\rangle\\ =\ & 3\cdot(\frac{-i\lambda_0}{4!})D_F(x-y)\int d^4zD_F(z-z)D_F(z-z)\\ +\ & 12\cdot(\frac{-i\lambda_0}{4!})\int d^4zD_F(x-z)D_F(y-z)D_F(z-z)\end{aligned}$$

It can be represented by the so called Feynman diagram.
The symmetry factor of the first diagram is $S = \frac{4!}{3} = 8$. The symmetry factor of the second diagram is $S = \frac{4!}{12} = 2$. The Feynman rules for ϕ^4 theory are:

1. For each propagator, $P = D_F(x-y)$
2. For each vertex, $V = (-i\lambda_0)\int d^4z$

Figure 23.3: Feynman diagram representation of perturbation expansion

3. For each external point, $E = 1$
4. Divided by the symmetry factor

At last, we can prove that

$\langle\Omega|T\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\}|\Omega\rangle =$ sum of all E-connected diagrams with n external points

Here, the "E-disconnected" means disconnected from all external points", being called "vacuum bubbles". They vacuum bubbles in $\langle 0|T\left\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle$ are all cancelled by the $\langle 0|T\left\{\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle$.

23.4 Path integral formulation

23.4.1 Basic formulation

Recall the path integrals formulation in quantum mechanics, we have

$$\langle\phi_b(\boldsymbol{x})|e^{-iHT}|\phi_a(\boldsymbol{x})\rangle = \int \mathcal{D}\phi\mathcal{D}\pi\exp\left[i\int_0^T d^4x(\pi\dot{\phi} - \frac{1}{2}\pi^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi))\right]$$

Here, $\langle\phi_b(\boldsymbol{x})|$ is the eigenstate of $\phi_S(\boldsymbol{x}) = \phi_H(\boldsymbol{x},0)$ with eigenvalue $\phi_b(\boldsymbol{x})$ at time $t = T$,$|\phi_a(\boldsymbol{x})\rangle$ is the eigenstate of $\phi_S(\boldsymbol{x})$ with eigenvalue $\phi_a(\boldsymbol{x})$ at time $t = 0$.
Since the exponential is quadratic in π, we can complete the square and evaluate the $\mathcal{D}(\pi)$ integral to obtain

$$\langle\phi_b(\boldsymbol{x})|e^{-iHT}|\phi_a(\boldsymbol{x})\rangle = \int \mathcal{D}\phi\exp\left[i\int_0^T d^4x\mathcal{L}\right]$$

Now we can abandon the Hamiltonian formalism and take the the equation above to define the Hamiltonian dynamics.

Note: *We emphasize that in this section, ϕ_H denotes the Heisenberg picture of field whose value is operators,ϕ_S denotes the Schrödinger picture of field, $\phi(x)$ is classical field whose value is ordinary number.*

Correlation function

$$\langle\Omega|T\phi_H(x_1)\phi_H(x_2)|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\phi(x_1)\phi(x_2)\exp\left[i\int_T^T d^4x\mathcal{L}\right]}{\int \mathcal{D}\phi\exp\left[i\int_T^T d^4x\mathcal{L}\right]}$$

The proof can be found in chapter 9.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*.

Functional derivatives and the generating functional

We define the generating functional as

$$Z[J] \equiv \int \mathcal{D}\phi \exp\left[i\int d^4x\mathcal{L} + J(x)\phi(x)\right]$$

We can prove that

$$\langle\Omega|T\phi_H(x_1)\cdots\phi_H(x_n)|\Omega\rangle = \frac{1}{Z[0]}\left(-i\frac{\delta}{\delta J(x_1)}\right)\cdots\left(-i\frac{\delta}{\delta J(x_n)}\right)Z[J]\Big|_{J=0}$$

Here, $Z[0] \equiv Z[J=0]$.

23.4.2 Free field theory

In Klein-Gordon field theory,

$$\int d^4x[\mathcal{L}_0(\phi) + J\phi] = \int d^4x[\frac{1}{2}\phi(\partial^2 - m^2 + i\epsilon)\phi + J\phi]$$

Define

$$\phi'(x) \equiv \phi(x) + \int d^4y(-iD_F(x-y))J(y)$$

Recall that $(\partial^2 - m^2)D_F(x-y) = i\delta(x-y)$, we can derive that

$$\int d^4x[\mathcal{L}_0 + J\phi] = \int d^4x[\frac{1}{2}\phi'(\partial^2 - m^2 + i\epsilon)\phi'] - \int d^4xd^4y\frac{1}{2}J(x)[-iD_F(x-y)]J(y)$$

After integration, we can know that

$$Z[J] = Z[0]\exp[-\frac{1}{2}\int d^4xd^4yJ(x)D_F(x-y)J(y)]$$

So,

$$\langle 0|T\phi_H(x_1)\phi_H(x_2)|0\rangle = \frac{1}{Z[0]}\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\exp\left[-\frac{1}{2}\int d^4xd^4yJ(x)D_F(x-y)J(y)\right]\Big|_{J=0} = D_F(x_1-x_2$$

23.5 Perturbation theory for path integral quantization

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4 \\
\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_1 \quad \mathcal{L}_1 = -\frac{\lambda_0}{4!}\phi^4(\boldsymbol{x}) \\
Z[J] &= \int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+J\phi]} \\
&= e^{i\int d^4y\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(y)})}\int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+J\phi]} \\
&= Z_0[0]e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)})}\exp[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)] \\
&= Z_0[0]\sum_{V=0}^{\infty}\frac{1}{V!}[\frac{-i\lambda_0}{4!}\int d^4x(\frac{1}{i}\frac{\delta}{\delta J(x)})^4]^V \times \sum_{P=0}^{\infty}\frac{1}{P!}[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)]^P
\end{aligned}$$

If we focus on a term with particular values of V and P, then the number of surviving sources (after we take all the functional derivatives) is $E = 2P - 4V$. The $4V$ functional derivatives can act on the 2P sources in $\frac{(2P)!}{(2P-4V)!}$ different combinations. However, many of the resulting expressions are algebraically identical.

To organize them, we introduce Feynman diagrams similar to that in perturbation theory of canonical quantization. In these diagrams, a line segment stands for a propagator $D_F(x-y)$, a filled circle at one end of a line segment for a source $i\int d^4x J(x)$, and a vertex joining four line segments for $-i\lambda_0 \int d^4z$.

For each diagram, we can assign a symmetry factor S_P similar to that in perturbation theory for canonical quantization. Due to the fact that some external sources are identical here, usually $S_P \neq S_C$. But when calculating the correlation function, the exchange of the order of functional derivatives to identical sources can eliminate the difference.

We can demonstrate that

$$Z[J] = Z_0[0]\exp(\sum_I C_I)$$

Here, C_I stands for a particular connected diagram, including its symmetry factor. We can define $W[J]$ as

$$Z[J] \equiv Z[0]\exp(-iW[J])$$

As, $W[0] = 0$, we know

$$-iW[J] = \sum_{I\neq\{0\}} C_I$$

The notation $I \neq \{0\}$ means that the vacuum diagrams are omitted from the sum.
The detailed discussion can be found in chapter 9 of *Quantum field theory (M. Srednicki)*.

23.6 Symmetries in the functional formalism

Equations of motion

The correlation function of the field theory is given by

$$\langle\Omega|T\phi_H(x_1)\cdots\phi_H(x_n)|\Omega\rangle = Z[0]^{-1}\int \mathcal{D}\phi e^{iS}\phi(x_1)\cdots\phi(x_n)$$

The equation of motion of classical field theory will be give by

$$\frac{\delta S}{\delta\phi(x)} = 0$$

In quantum field theory, we derive the equation of motion by claim that the path integral will be invariant under the infinitesimal change of field, i.e. $\phi(x) \to \phi(x) + \epsilon(x)$. Define

$$Z[\phi(x_1),\cdots,\phi(x_n)] \equiv \int \mathcal{D}\phi e^{iS}\phi(x_1)\cdots\phi(x_n)$$

We know

$$\delta Z = \int \mathcal{D}\phi e^{iS} \left\{ \int d^4x \epsilon(x) \left[i\frac{\delta S}{\delta\phi(x)}\phi(x_1)\cdots\phi(x_n) + \delta(x-x_1)\cdots\phi(x_n) + \cdots + \phi(x_1)\cdots\delta(x-x_n) \right] \right\}$$

so

$$\langle \frac{\delta S}{\delta\phi(x)}\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n} \langle \phi(x_1)\cdots(i\delta(x-x_i))\cdots\phi(x_n)\rangle$$

Example: For Klein-Gordon field,

$$\mathcal{L} = -\frac{1}{2}\phi^\mu\phi\partial_\mu\phi - \frac{1}{2}m\phi^2$$

The variation of S gives

$$\frac{\delta S}{\delta\phi(x)} = (\partial^2 - m^2)\phi(x)$$

So, we can get

$$(\partial^2 - m^2)\langle 0|T\phi(x)\phi(x_1)|0\rangle = i\delta(x-x_1)$$

Conservation law

Consider a local field theory of a set of fields $\phi_a(x)$, governed by a Lagrangian $\mathcal{L}(\phi)$. An infinitesimal symmetry transformation on the fields ϕ_a will be of the general form

$$\phi_a(x) \to \phi_a(x) + \epsilon\Delta\phi_a(x)$$

We assume that when ϵ is a constant, the action is invariant under this transformation. Then the Lagrangian must be invariant up to a total divergence:

$$\mathcal{L}[\phi] \to \mathcal{L}[\phi] + \epsilon\partial_\mu K^\mu$$

If the symmetry parameter ϵ depend on x, the variation of Lagrangian will be

$$L[\phi] \to \mathcal{L}[\phi] + (\partial_\mu\epsilon)\Delta\phi_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + \epsilon\partial_\mu K^\mu$$

So,

$$\frac{\delta S}{\delta\epsilon(x)} = \partial_\mu j^\mu \quad j^\mu = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\Delta\phi_a + K^\mu$$

If the measure $\mathcal{D}\phi$ is invariant under the transformation, we can derive similarly that

$$\langle \partial_\mu j^\mu(x)\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n}\langle \phi(x_1)\cdots(i\Delta\phi(x_i)\delta(x-x_i))\cdots\phi(x_n)\rangle$$

23.7 Cross section and the S-matrix

For a relativistic collinear scattering process, we have

$$dN = \sigma|\boldsymbol{v}_1 - \boldsymbol{v}_2|n_1 dt$$

Consider a $2 \to n$ process:

$$p_1 + p_2 \to \{p_j\}$$

Suppose the volume of the space in which the scattering process takes place is V, the duration of the scattering process is T. So, the density of incident particle is

$$n_1 = \frac{1}{V}$$

and the number of events is

$$N = \frac{|\langle i|S|f\rangle|^2}{|\langle i|i\rangle||\langle f|f\rangle|} d\Pi.$$

Here, $d\Pi$ is the region of final state momenta at which we are looking

$$d\Pi = \prod_j \frac{V}{(2\pi)^3} d^3 p_j$$

Recall that

$$\delta^{(3)}(0) = \frac{V}{(2\pi)^3} \quad \delta^{(4)}(0) = \frac{VT}{(2\pi)^4} \quad \langle p|p\rangle = (2\pi)^3 2\omega\delta^{(3)}(0)$$

and write the scattering amplitude in the form of

$$\langle i|S|f\rangle = i\mathcal{M}(2\pi)^4\delta(\sum p_f - \sum p_i)$$

We can calculate that

$$|\langle i|S|f\rangle|^2 = |\mathcal{M}|^2 VT(2\pi)^4\delta(\sum p_j - p_1 - p_2) \quad \langle i|i\rangle = (2E_1V)(2E_2V) \quad \langle f|f\rangle = \prod_j (2E_jV)$$

At last, putting everything together, we have

$$d\sigma = \frac{1}{(2E_1)(2E_2)|\boldsymbol{v}_1 - \boldsymbol{v}_2|}|\mathcal{M}|^2 d\Pi_{\text{LIPS}}$$

where,

$$d\Pi_{\text{LIPS}} = \prod_j \frac{d^3 p_j}{(2\pi)^3}\frac{1}{2E_j}(2\pi)^4\delta(\sum p_j - p_1 - p_2)$$

is called the Lorentz-invariant phase space (LIPS).
All the factors of V and T have dropped out, so now it is trivial to take $V \to \infty$ and $T \to \infty$.

Decay rates

The definition of decay rates is

$$\Gamma \equiv \frac{\text{Number of Events}}{\text{Time}}$$

Consider a $1 \to n$ process:

$$p_1 \to \{p_j\}$$

We can derive the decay rates by the similar method,

$$d\Gamma = \frac{1}{2E_1}|\mathcal{M}|^2 d\Pi_{\text{LIPS}}$$

Special cases

For $2 \to 2$ scattering in the center-of-mass frame

$$p_1 + p_2 \to p_3 + p_4$$

Then

$$d\Pi_{\text{LIPS}} = \frac{d^3p_3}{(2\pi)^3}\frac{1}{2E_3}\frac{d^3p_4}{(2\pi)^3}\frac{1}{2E_4}(2\pi)^4\delta(p_3 + p_4 - p_1 - p_2)$$

We can now integrate over $\boldsymbol{p}_4$ to give

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2}d\Omega \int dp_f \frac{p_f^2}{E_3E_4}\delta(E_3 + E_4 - E_{CM})$$

where $p_f = |\boldsymbol{p}_3| = |\boldsymbol{p}_4|$, $E_3 = \sqrt{p_f^2 + m_3^2}$,$E_4 = \sqrt{p_f^2 + m_4^2}$ and $E_{CM} = E_1 + E_2$. Define $x(p_f) \equiv E_3 + E_4 - E_{CM}$, we can get

$$\frac{dx}{dp_f} = \frac{E_3 + E_4}{E_3E_4}p_f$$

So,

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2}d\Omega\frac{p_f}{E_{CM}}\theta(E_{CM} - m_3 - m_4)$$

Plugging this into the general equation for cross section, we can get

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2E_{CM}^2}\frac{|\boldsymbol{p}_f|}{|\boldsymbol{p}_i|}|\mathcal{M}|^2\theta(E_{CM} - m_3 - m_4)$$

If all the masses are equal then this formula simplifies further

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2E_{CM}^2}|\mathcal{M}|^2$$

23.8 LSZ reduction formula

23.8.1 Field strength renormalization

The completeness relation:

$$\mathbf{1} = |\Omega\rangle\langle\Omega| + \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\boldsymbol{p}}}|\lambda_{\boldsymbol{p}}\rangle\langle\lambda_{\boldsymbol{p}}|$$

Here, $E_{\boldsymbol{p}} = \sqrt{m_\lambda^2 + \boldsymbol{p}^2}$

Figure 23.4: Particle's energy-momentum relation

Assume for now $x^0 > y^0$ and define $\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \langle\Omega|\phi(x)\phi(y)|\Omega\rangle - \langle\Omega|\phi(x)|\Omega\rangle\langle\Omega|\phi(y)|\Omega\rangle$ as connected two point function. (The term $\langle\Omega|\phi(x)|\Omega\rangle\langle|\Omega\phi(y)|\Omega\rangle$ is usually zero by symmetry; for higher spin fields, it is zero by Lorentz invariance.) The connected two point function is

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\boldsymbol{p}}}\langle\Omega|\phi(x)|\lambda_{\boldsymbol{p}}\rangle\langle\lambda_{\boldsymbol{p}}|\phi(y)|\Omega\rangle$$

It can be verified that

$$\langle\Omega|\phi(x)|\lambda_{\boldsymbol{p}}\rangle = \langle\Omega|\phi(0)|\lambda_0\rangle e^{ipx}|_{p^0=E_{\boldsymbol{p}}}$$

So,

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \sum_\lambda \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+m_\lambda^2-i\epsilon}e^{ip(x-y)}|\langle\Omega|\phi(0)|\lambda_0\rangle|^2$$

Analogous expressions hold for the case $y^0 > x^0$, and both cases can be summarized as

$$\langle\Omega|T\phi(x)\phi(y)|\Omega\rangle_C = \int_0^\infty \frac{dM^2}{2\pi}\rho(M^2)D_F(x-y;M^2)$$

and

$$\rho(M^2) = \sum_\lambda (2\pi)\delta(M^2 - m^2)|\langle\Omega|\phi(0)|\lambda_0\rangle|^2$$

Figure 23.5: The structure of the spectral density function $\rho(M^2)$

The one-particle state contribute an isolated delta function to the spectral density function, so

$$\rho(M^2) = 2\pi\delta(M^2 - m^2)\cdot Z + \text{ (nothing else until } M^2 >\sim (2m)^2)$$

$Z = |\langle\Omega|\phi(0)|\lambda_0\rangle|^2$ is called field-strength renormalization. m is the physical mass of a single particle of the ϕ boson. The Fourier transformation of the two point function is

$$\begin{aligned}&\int d^4x e^{-ipx}\langle\Omega|T\phi(x)\phi(0)|\Omega\rangle_C \\ = &\int_0^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon} = \frac{-iZ}{p^2+m^2-i\epsilon} + \int_{\sim 4m^2}^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon}\end{aligned}$$

Figure 23.6: The structure of the two point function in Fourier space

23.8.2 LSZ reduction formula

Theorem 23.2 LSZ reduction formula

$$\prod_1^n \int d^4x_i e^{-ip_ix_i} \prod_1^m d^4y_j e^{ik_jy_j} \langle\Omega|T\{\phi(x_1)\cdots\phi(x_n)\phi(y_1)\cdots\phi(y_m)\}|\Omega\rangle$$
$$\underset{p_i^0\to E_{\boldsymbol{p}_i}\ k_i^0\to E_{\boldsymbol{k}_i}}{\sim} \left(\prod_1^n \frac{-\sqrt{Z}i}{p_i^2+m^2-i\epsilon}\right)\left(\prod_1^m \frac{-\sqrt{Z}i}{k_i^2+m^2-i\epsilon}\right)\langle\boldsymbol{p}_1\cdots\boldsymbol{p}_n|S|\boldsymbol{k}_1\cdots\boldsymbol{k}_m\rangle$$

The $\sim$ means the two sides of the expression share the same singular structure around $p_i^0 \to E_{\boldsymbol{p}_i}, k_i^0 \to E_{\boldsymbol{k}_i}$. The proof can be found in chapter 7.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*. To express the formula above in the language of Feynman diagrams, we consider the S-matrix element for 2-particle $\to$ 2-particle for example. Note the disconnected diagram should be disregarded because they do not have the singularity structure with a product of four poles indicated by on the right side of the LSZ reduction formula. So, the exact four point function

$$\prod_1^2 \int d^4x_i e^{-ip_ix_i} \prod_1^2 d^4y_i e^{ik_jy_j} \langle\Omega|T\{\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\}|\Omega\rangle$$

has the general form showed as below.

Figure 23.7: Amputated Feynman diagram

We can sum up the corrections to each external leg. Let $-iM^2(p^2)$ denote the sum of all 1PI insertions into the scalar propagator:

Figure 23.8: Diagram representation of 1PI propagator

Note: *One-particle-irreducible, or 1PI for short, refers to diagrams that is still connected after one line is cut*

Then the exact propagator can be written as a geometric series in Figure 21.9.

Figure 23.9: Diagram representation of exact propagator

The result is $\frac{-i}{p^2+m_0^2+M^2}$. If we expand each resummed propagator about the physical particle pole, we see that each external leg of the four-point amplitude contributes

$$\frac{-i}{p^2+m_0^2+M^2} \underset{p^0\to E_{\boldsymbol{p}}}{\sim} \frac{-iZ}{p^2+m^2} + \text{(regular)}$$

Thus, the sum of diagrams contains a product of four point poles:

$$\frac{-iZ}{p_1^2+m^2}\frac{-iZ}{p_2^2+m^2}\frac{-iZ}{k_1^2+m^2}\frac{-iZ}{k_2^2+m^2}$$

So, the S matrix element can be represented by

$$\langle \mathbf{p}_1\mathbf{p}_2|\, S\, |\mathbf{k}_1\mathbf{k}_2\rangle = \left(\sqrt{Z}\right)^4$$

,

Figure 23.10: Feynman diagram representation of LSZ reduction formula

It is easy to be generalized to the more complicated scattering cases. After Fourier transforming the n-point function to momentum space and cutting off the external legs, the Feynman rules for S-matrix element can be stated as follows:

1. For each propagator, $P = \frac{-i}{p^2+m_0^2-i\epsilon}$;
2. For each vertex, $V = -i\lambda_0$;
3. For each external point, $E = 1$;
4. Impose momentum conservation at each vertex;
5. Integrate over each undetermined loop momentum: $\int \frac{d^4p}{(2\pi)^4}$;
6. Divided by the symmetry factor;
7. Multiply the total momentum conservation factor $(2\pi)^4\delta(\sum p_f - \sum p_i)$

We can write $\langle f|S|i\rangle = (Z_1)^{\frac{n_s}{2}} i\mathcal{M}(2\pi)^4\delta(\sum p_f - \sum p_i)$ for convenience.

23.9 Renormalization

Renormalization, the procedure in quantum field theory by which divergent parts of a calculation, leading to nonsensical infinite results, are absorbed by redefinition into a few measurable quantities, so yielding finite answers.

23.9.1 Counting of ultraviolet divergence

Consider a pure scalar theory in d dimensions with a ϕ^n interaction term

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{n!}\phi^n$$

Let N be the number of external lines in the diagram, P the number of propagators, V the number of vertices. The number of the loops in the diagram is $L = P - V + 1$. There are n lines meeting at each vertex, so $nV = 2P + N$. Loosely speaking, each loop has an integral d^dp, each propagator has a factor p^{-2}, so the superficial degrees of divergence is

$$D = dL - 2P = d + [n(\frac{d-2}{2}) - d)]V - (\frac{d-2}{2})N$$

According the superficial degrees of divergence of the diagram. These three possible types of ultraviolet behaviour of quantum field theories. We will refer to them as follows

1. Super-renormalizable theory: Only a finite number of Feynman diagrams superficially diverge.
2. renormalizable theory: Only a finite number of amplitudes superficially diverge; however, divergences occur at all orders in perturbation theory.
3. Non-renormalizable theory: All amplitudes are divergent at a sufficiently high order in perturbation theory.

So, for ϕ^4 theory in four dimension, $D = 4 - N$. It is a renormalizable theory. For ϕ^3 theory in four dimension, $D = 4 - V - N$. It is a super-renormalizable theory. For ϕ^6 theory in four dimension, $D = 4 + 2V - N$. It is a Non-renormalizable theory.

The superficial degrees of freedom can also be derived from dimensional analysis. The dimension of λ is $d - \frac{n(d-2)}{2}$. Now consider an arbitrary diagram with N external lines. One way that such a diagram could arise is from an interaction term $\eta\phi^N$ in the Lagrangian. The dimension of η would then be $d - \frac{N(d-2)}{2}$, and therefore we conclude that any (amputated) diagram with N external lines has dimension $d - \frac{N(d-2)}{2}$. In our theory with only the $\lambda\phi^n$ vertex, if the diagram has V vertices, its divergent part is proportional to $\lambda^V\Lambda^D$, where Λ is a high momentum cut-off and D is the superficial degree of divergence. Applying dimensional analysis, we find

$$d - \frac{N(d-2)}{2} = V[d - \frac{n(d-2)}{2}] + D$$

Note that the quantity that multiplies V in this expression is just the dimension of the coupling constant λ. Thus we can characterize the three degrees of renormalizability in a second way:

1. Super-renormalizable: Coupling constant has positive mass dimension.
2. renormalizable: Coupling constant is dimensionless.
3. Non-renormalizable: Coupling constant has negative mass dimension.

23.9.2 Renormalized perturbation theory

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4$$

We write m_0 and λ_0, to emphasize that these are the bare values of the mass and coupling constant, not the values measured in experiments. Since the theory is invariant under $\phi \to -\phi$, all amplitudes with an odd number of external legs vanish. The only divergent amplitudes are therefore Ignoring the vacuum diagram, these amplitudes contain three infinite constants.

Figure 23.11: Divergence of ϕ^4 theory

Our goal is to absorb these constants into the three unobservable parameters of the theory: the bare mass, the bare coupling constant, and the field strength. To accomplish this goal, it is convenient to reformulate the perturbation expansion so that these unobservable quantities do not appear explicitly in the Feynman rules. Recall that the exact two-point function has the form

$$\int d^4x\langle\Omega|\phi(x)\phi(0)|\Omega\rangle e^{-ipx} = \frac{-iZ}{p^2+m^2} + \text{ terms regular at } p^2 = m^2$$

We can eliminate the Z from this equation by rescaling the field: $\phi = Z^{\frac{1}{2}}\phi_r$ We also define

$$\delta_Z = Z - 1 \quad \delta_m = Zm_0^2 - m^2 \quad \delta_\lambda = \lambda_0 Z^2 - \lambda$$

Then the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi_r\partial_\mu\phi_r - \frac{1}{2}m^2\phi_r^2 - \frac{\lambda}{4!}\phi_r^4 - \frac{1}{2}\delta_Z\partial^\mu\phi_r\partial_\mu\phi_r - \frac{1}{2}\delta_m\phi_r^2 - \frac{\delta\lambda}{4!}\phi_r^4$$

The last three terms, known as counter-terms, have absorbed the infinite but unobservable shifts between the bare parameters and the physical parameters.We give precise definitions of

Figure 23.12: Renormalization condition

the physical mass and coupling constant as Figure 21.12.

The renormalization scheme here is called on-shell (OS) scheme. Other renormalization scheme would be introduced later. These equations are called renormalization conditions. Our new Lagrangian gives a new set of Feynman rules as Figure 21.13.

Figure 23.13: Feynman rules for renormalized perturbation theory

We can use these new Feynman rules to compute any amplitude in ϕ^4 theory. The procedure is as follows. Compute the desired amplitude as the sum of all possible diagrams created from the propagator and vertices shown above. The loop integrals in the diagrams will often diverge, so one must introduce a regulator. The result of this computation will be a function of the three unknown parameters δ_Z, δ_m, and δ_λ. Adjust (or "renormalise") these three parameters as necessary to maintain the renormalization conditions. After this adjustment, the expression for the amplitude should be finite and independent of the regulator.
This procedure, using Feynman rules with counter-terms, is known as renormalized perturbation theory.

Mandelstam variable

In theoretical physics, the **Mandelstam variable** are numerical quantities that encode the energy, momentum, and angles of particles in a scattering process in a Lorentz-invariant fashion. They are used for scattering processes of two particles to two particles. The Mandelstam vari-

ables s, t, u are then defined by

$$\begin{aligned} s &= -(p_1+p_2)^2 = -(p_3+p_4)^2 \\ t &= -(p_1-p_3)^2 = -(p_2-p_4)^2 \\ u &= -(p_1-p_4)^2 = -(p_2-p_3)^2 \end{aligned}$$

Where p_1 and p_2 are the four-momenta of the incoming particles and p_3 and p_4 are the four-momenta of the outgoing particles. s is also known as the square of the center-of-mass energy (invariant mass) and t is also known as the square of the four-momentum transfer.
We can verify that

$$s+t+u = m_1^2+m_3^2+m_3^2+m_4^2$$

23.9.3 Techniques for renormalization

Feynman's formula

Theorem 23.3 Feynman's formula

$$\frac{1}{A_1\cdots A_n} = \int dF_n (x_1A_1+\cdots+x_nA_n)^{-n}$$

where the integration measure over the Feynman parameters x_i is

$$\int dF_n = (n-1)!\int_0^1 dx_1\cdots dx_n\delta(x_1+\cdots+x_n-1)$$

This measure is normalized so that

$$\int dF_n = 1$$

A generalization of Feynman's formula is

$$\frac{1}{A_1^{\alpha_1}\cdots A_n^{\alpha_n}} = \frac{\Gamma(\sum_i\alpha_i)}{\prod_i\Gamma(\alpha_i)}\frac{1}{(n-1)!}\int dF_n\frac{\prod_i x_i^{\alpha_i-1}}{(\sum_i x_iA_i)^{\sum_i\alpha_i}}$$

Wick rotation

For an integral $\int d^dq f(q^2-i\epsilon)$, if the integrand vanishes fast enough as $|q_0|\to\infty$, we can rotate this contour clockwise by $\frac{\pi}{2}$, so that it runs from $-i\infty$ to $i\infty$. In making this Wick rotation, the contour does not pass over any poles. (The $i\epsilon$ are needed to make this statement unambiguous.) Thus the value of the integral is unchanged. It is now convenient to define a Euclidean d-dimensional vector $\bar{q}$ via $q^0 = i\bar{q}_d$ and $q_j = \bar{q}_j$; then $q^2 = \bar{q}^2$, where

$$\bar{q}^2 = \bar{q}_1^2+\cdots+\bar{q}_d^2$$

Figure 23.14: Wick rotation

Also, $d^d q = i d^d \bar{q}$.Therefore, in general,

$$\int d^d q f(q^2 - i\epsilon) = i \int d^d \bar{q} f(\bar{q}^2)$$

Dimensional regularization

Dimensional regularization is a method for regularizing integrals in the evaluation of Feynman diagrams. For example, if one wishes to evaluate a loop integral which is logarithmically divergent in four dimensions, like

$$\int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2 + m^2)^2}$$

One first rewrites the integral in some way so that the number of variables integrated over does not depend on d, and then we formally vary the parameter d, to include non-integral values like $d = 4 - \epsilon$.

$$\int_0^\infty \frac{dp}{(2\pi)^{4-\varepsilon}} \frac{2\pi^{(4-\varepsilon)/2}}{\Gamma\left(\frac{4-\varepsilon}{2}\right)} \frac{p^{3-\varepsilon}}{(p^2+m^2)^2} = \frac{2^{\varepsilon-4}\pi^{\frac{\varepsilon}{2}-1}}{\sin(\frac{\pi\varepsilon}{2})\Gamma(1-\frac{\varepsilon}{2})} m^{-\varepsilon} = \frac{1}{8\pi^2\varepsilon} - \frac{1}{16\pi^2}\left(\ln\frac{m^2}{4\pi} + \gamma\right) + \mathcal{O}(\varepsilon)$$

There is a useful formula for calculating the integral

$$\int \frac{d^d\bar{q}}{(2\pi)^d} \frac{(\bar{q}^2)^a}{(\bar{q}^2 + D)^b} = \frac{\Gamma(b - a - \frac{1}{2}d)\Gamma(a + \frac{1}{2}d)}{(4\pi)^{d/2}\Gamma(b)\Gamma(\frac{1}{2}d)} D^{-(b-a-d/2)}$$

If $a = 0$, then the formula will be

$$\int \frac{d^d\bar{q}}{(2\pi)^d} \frac{1}{(\bar{q}^2 + D)^b} = \frac{\Gamma(b - \frac{1}{2}d)}{(4\pi)^{d/2}\Gamma(b)} D^{-(b-d/2)}$$

23.9.4 One loop structure of ϕ^4 theory

First consider the basic two-particle scattering amplitude, If we define $p = p_1 + p_2$, then the second diagram of Figure 21.15 is

$$\frac{(-i\lambda)^2}{2} \int \frac{d^4k}{(2\pi)^4} \frac{-i}{k^2 + m^2} \frac{-i}{(k+p)^2 + m^2} \equiv (-i\lambda)^2 i V(-p^2)$$

Figure 23.15: Feynman diagram representation of two-particle scattering to one loop

So the entire amplitude is therefore

$$i\mathcal{M} = -i\lambda + (-i\lambda)^2[iV(s) + iV(t) + iV(u)] - i\delta_\lambda + \mathcal{O}(\lambda^3)$$

To keep λ dimensionless in dimensional regularization, we can make the transformation $\lambda \to \lambda\tilde{\mu}^\epsilon$. Here, $\tilde{\mu}$ is an arbitrary number with mass dimension 1 and $\epsilon \equiv 4 - d$.
We can calculate that

$$V(-p^2) = -\frac{1}{32\pi^2}\int_0^1 (\frac{2}{\epsilon} + \ln(\frac{\mu^2}{D(-p^2)}))$$

where $\mu \equiv \sqrt{4\pi}e^{-\gamma/2}\tilde{\mu}$, $D(-p^2) = x(1-x)p^2 + m^2$
The renormalization condition implies that

$$\delta_\lambda = -\lambda^2[V(4m^2) + 2V(0)] + \mathcal{O}(\lambda^3)$$

So,

$$i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2}\int_0^1 dx\left[\ln(\frac{D(s)}{D(4m^2)}) + \ln(\frac{D(t)}{D(0)}) + \ln(\frac{D(u)}{D(0)})\right] + \mathcal{O}(\lambda^3)$$

To determine δ_Z and δ_m we must compute the two-point function. Define $-iM(p^2)$ as the sum of all one-particle-irreducible insertions into the propagator.The full two-point function is given by

$$\frac{-i}{p^2 + m^2 + M^2}$$

The renormalization conditions require that the pole in this full propagator occur at $p^2 = -m^2$ and have residue 1. These two conditions are equivalent, respectively, to

$$M^2(p^2)|_{p^2=-m^2} = 0 \qquad \frac{d}{dp^2}M^2(p^2)\bigg|_{p^2=-m^2} = 0$$

We can calculate that

$$-iM^2(p^2) = \frac{i\lambda}{32\pi^2}(\frac{2}{\epsilon} + \ln(\frac{\mu^2}{m^2}) + 1)m^2 - i(p^2\delta_Z + \delta_m)$$

So, to the order of λ,

$$\delta_Z = \mathcal{O}(\lambda^2) \quad \delta_m = \frac{\lambda}{32\pi^2}(\frac{2}{\epsilon} + \ln(\frac{\mu^2}{m^2}) + 1)m^2 + \mathcal{O}(\lambda^2) \quad M^2(p^2) = \mathcal{O}(\lambda^2)$$

The detailed calculation can be found in chapter 10.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)* and will be eliminated here.

Perturbation theory to all orders

We begin by summing all one-particle irreducible diagrams with two external lines; this gives us the self-energy M^2. We next sum all amputated diagrams with four external lines; this gives us the four-point vertex function $V_4(k_1, k_2, k_3, k_4)$. (We can prove that amputated diagrams with four external lines must be one-particle irreducible). Order by order in λ, we must adjust the value of the lagrangian coefficients δ_Z, δ_m, and δ_λ to maintain the conditions $M^2(-m^2) = 0$, $\frac{dM^2}{dp^2}(-m^2) = 0$, and $V_4(s = 4m^2) = 0$.
Next we will construct the n-point vertex functions V_n with $4 < n \leq E$, where E is the number of external lines in the process of interest. We compute these using a skeleton expansion. This means that we draw all the contributing 1PI diagrams, but omit diagrams that include either propagator or three-point vertex corrections. That is, we include only diagrams that are not only 1PI, but also 2PI and 4PI: they remain connected when any one, two, or four lines are cut. (Cutting four lines may isolate a single tree-level vertex, but nothing more complicated.) Then we take the propagators and vertices in these diagrams to be given by the exact propagator $\frac{-i}{p^2+m^2+M^2(p^2)}$ and vertex $V_4(k_1, k_2, k_3, k_4)$, rather than by the tree-level propagator and vertex. We then sum these skeleton diagrams to get V_n for $4 < n \leq E$. Order by order in λ, this procedure is equivalent to computing V_n by summing the usual set of contributing 1PI diagrams.
Next we draw all tree-level diagrams that contribute to the process of interest (which has E external lines), including not only four-point vertices, but also n-point vertices. Then we evaluate these diagrams using the exact propagator for internal lines, and the exact 1PI vertices V_n; external lines are assigned a factor of one. We sum these tree diagrams to get the scattering amplitude. Order by order in λ, this procedure is equivalent to computing the scattering amplitude by summing the usual set of contributing diagrams.
Thus we now know how to compute an arbitrary scattering amplitude to arbitrarily high order. The procedure is the same in any quantum field theory; only the form of the propagators and vertices change, depending on the spins of the fields.

23.9.5 General renormalization theory

Recall some of the major results and methods of renormalization theory:

1. In perturbation theory, bare and physical quantities are related by ultraviolet-divergent expressions
$$m_{phys} = m_0 + \Delta_m$$
where m_{phys} is finite, Δ_m is ultraviolet-divergent, and so m_0 is necessarily ultraviolet-divergent.

2. We express the Lagrangian in terms of physical quantites, and separate it into
$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_{CT}$$
where $\mathcal{L}_0$ is the canonically normalized free Lagrangian for physical fields and masses, $\mathcal{L}_I$ contains the interaction, again in terms of physical parameters, and $\mathcal{L}_{CT}$ contains the counterterms with ultraviolet divergent coefficients. From $\mathcal{L}_0$, we obtain the propagators of the physical fields. $\mathcal{L}_I$ and $\mathcal{L}_{CT}$ give interaction vertices.

3. At the one-loop level, the self-energy is given by the effective two-point vertices: the 1PI two-point vertex of the interaction and the counter-term two-point vertex. The counterterms absorb ultraviolet divergences, and the finite parts of the counterterms are determined by renormalization conditions, which ensure the quantities in $\mathcal{L}_0 + \mathcal{L}_I$ are physical. The conditions constrain the self-energy and the effective vertices, and give a finite, uniquely-determined value for the counterterms.

Degrees of divergences

Now, for a general theory in d spacetime dimensions, the field content is given by ϕ_f, $f = 1, 2, \cdots$, where f labels the field type. $[\phi_f] = \Delta f$ and Δf in all physical theories. We have interaction vertices of type i, $i = 1, 2, \cdots$ contributing a term of the form

$$\lambda_i \partial^{n_i} \prod_f \phi_f^{n_{if}}$$

Here, λ_i is the coupling constant, with dimension

$$[\lambda_i] = \kappa_i = d - n_i - \sum_f n_{if} \Delta f$$

Now consider a 1PI diagram in such a theory:

$$E_f \equiv \text{ number of external lines of } \phi_f$$

$$V_i \equiv \text{ number of vertices of type i .}$$

Then $M \sim \Lambda^D \prod_i \lambda_i^{V_i}$ and so $D = [M] - \sum_i V_i \kappa_i$. Again

$$[M] - d = -\sum E_f \Delta f$$

and so the general expression for the superficial degree of divergence is given by

$$D = d - \sum_f E_f \Delta f - \sum_i V_i \kappa_i$$

Diagrams which are ultraviolet divergent satisfy

$$\sum_f E_f \Delta f + \sum_i V_i \kappa_i < d$$

We can now divide all theories into

1. All $\kappa_i > 0$: superrenormalizable theories.
2. All $\kappa_i \geq 0$: renormalizable theories.
3. There exists at least one $\delta_i < 0$: non-renormalizable theories.

These terms also apply to individual interactions for a vertex of type i:

1. $\kappa_i > 0$: super-renormalizable, relevant interaction.
2. $\kappa_i \geq 0$: renormalizable, marginal interaction.
3. $\kappa_i < 0$: non-renormalizable, irrelevant interaction.

Cancellation of divergences

Consider a generic divergent diagram M of degree D; that is,

$$M = \int^{\Lambda} ds s^{D-1}$$

if all loop momenta are taken proportional to s. Generically, internal propagators have the form

$$\frac{1}{(as+p)^{\alpha}\cdots} \sim \frac{1}{s^{\alpha}}$$

for large s, where a is a numerical constant and p is a combination of the external momenta. Differentiating M, with respect to p, n times gives a term proportional to

$$\frac{1}{(as+p)^{\alpha+n}} \sim \frac{1}{s^{\alpha+n}}$$

and so $D+1$ derivatives with respect to the external momenta will make M finite. This means that

$$M(p) = M_0 + M_1 p + \cdots + M_D p^D + \text{ finite terms}$$

where the argument p of the function represents the collection of external momenta, we have suppressed the index structure, and $M_0, M_1, \cdots, M_D$ are potentially divergent constants. Suppose that M has E_f external lines of the field ϕ_f. Then divergences of $M(p)$ can be cancelled by counterterms of the form

$$\sum_{j=0}^{D} A_j (\partial)^j \prod_f \phi_f^{E_f}$$

where the A_j are divergent coefficients in order to cancel divergences in M_j. The index structure in $A_j \partial^j$ should match the suppressed index structure of $M(p)$.

23.10 Renormalization group

23.10.1 Modified minimal-subtraction scheme

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 - \frac{1}{2}\delta_Z\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}\delta_m\phi_r - \frac{\delta\lambda}{4!}\phi^4$$

For minimal-subtraction scheme, we do not demand that m be the physics mass of the field and ϕ create a normalized one-particle state. The physical meaning of λ is not expressed directly as well. Instead we choose δ_Z, δ_m and δ_λ to cancel the infinities, and nothing more; we say that δ_Z, δ_m and δ_λ have no finite parts. It is called the modified minimal-subtraction or $\overline{\text{MS}}$ scheme. ("modified" because we introduced μ via $\lambda \to \lambda\tilde{\mu}^\epsilon$, with $\mu \equiv \sqrt{4\pi}e^{-\gamma/2}\tilde{\mu}$; had we set $\mu = \tilde{\mu}$ instead, the scheme would be just plain minimal subtraction or MS.)
For loop corrections to propagator,

$$\delta_Z = \mathcal{O}(\lambda^2) \quad \delta_m = \left[\frac{\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)\right]\frac{1}{\epsilon}m^2 \quad M^2(p^2) = \frac{\lambda}{32\pi^2}(\ln(\frac{m^2}{\mu^2}) - 1)m^2 + \mathcal{O}(\lambda^2)$$

Firstly, in the $\overline{\text{MS}}$ scheme, the propagator will no longer have a pole at $k^2 = -m^2$. The pole will be somewhere else. However, by definition, the actual physical mass m_{ph} of the particle is determined by the location of this pole: $k^2 = -m_{ph}^2$. Thus, the Lagrangian parameter m is no longer the same as m_{ph}. The relation of m and m_{ph} is

$$m_{ph}^2 = M^2(-m_{ph}^2) + m^2$$

To the lowest order,

$$m_{ph}^2 = \left[1 + \frac{\lambda}{32\pi^2}(\ln(\frac{m^2}{\mu^2}) - 1)\right] m^2$$

Because m_{ph} is a independent of μ, according to $\frac{d}{d\mu} m_{ph} = 0$, it can be derived that

$$\frac{dm}{d\ln\mu} = \left[\frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)\right] m$$

Furthermore, the residue of this pole is no longer one. Let us call the residue R. So, in the LSZ formula, we get a net factor of $\sqrt{R}$ for each external line when using the $\overline{\text{MS}}$ scheme. And in ϕ^4 theory,

$$R = 1 + \mathcal{O}(\lambda^2)$$

For loop corrections to vertex,

$$\delta_\lambda = \left[\frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3)\right] \frac{1}{\epsilon}$$

$$i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2} \int_0^1 dx \left[\ln(\frac{D(s)}{\mu^2}) + \ln(\frac{D(t)}{\mu^2}) + \ln(\frac{D(u)}{\mu^2})\right] + \mathcal{O}(\lambda^3)$$

For a process with $p^2 \gg m^2$, we have

$$D \approx x(1-x)p^2$$

In OS renormalization scheme, the one-loop correction to propagator or vertex generally includes a factor

$$\ln\left(\frac{D}{D_0}\right) \sim \ln\frac{p^2}{m^2}$$

so perturbation theory is no longer a good approximation when $p^2 \gg m^2$. In $\overline{MS}$ renormalization scheme, introducing μ allows us to address this problem: if we choose $\mu \sim p$, no such logarithm arises. If we choose μ appropriately, that is, to be comparable to the momentum scale of the physical process, we can improve our perturbation expansion. So $\lambda(\mu)$ and $m(\mu)$ can be considered as the scale-dependent coupling constants. And the reason we get large logarithmic terms in the on-shell scheme is that we are trying to use coupling defined at one scale to describe physics at very different scales.

23.10.2 Beta and gamma function

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi_0\partial_\mu\phi_0 - \frac{1}{2}m_0^2\phi_0^2 - \frac{\lambda_0}{4!}\phi_0^4$$

It can be written as

$$\mathcal{L} = -\frac{1}{2}Z_\phi\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}Z_m m^2\phi^2 - Z_\lambda\tilde{\mu}^\epsilon\frac{\lambda}{4!}\phi^4$$

So,

$$\phi_0 = Z_\phi^{1/2}\phi \quad m_0 = Z_\phi^{-1/2}Z_m^{1/2}m \quad \lambda_0 = Z_\phi^{-2}Z_\lambda\lambda\tilde{\mu}^\epsilon$$

After using dimensional regularization, the infinities coming from loop integrals take the form of inverse powers of ϵ. In the $\overline{\text{MS}}$ renormalization scheme, we choose the Zs to cancel off these powers of $1/\epsilon$, and nothing more. Therefore the Zs can be written as

$$Z_\phi = 1 + \sum_{n=1}^{\infty}\frac{a_n(\lambda)}{\epsilon^n}$$
$$Z_m = 1 + \sum_{n=1}^{\infty}\frac{b_n(\lambda)}{\epsilon^n}$$
$$Z_\lambda = 1 + \sum_{n=1}^{\infty}\frac{c_n(\lambda)}{\epsilon^n}$$

In ϕ^4 theory, $a_1 = \mathcal{O}(\lambda^2)$, $b_1 = \frac{\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$,$c_1 = \frac{3\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$
Remember that bare fields and parameters must be independent of μ.Define

$$G(\lambda,\epsilon) \equiv \ln(Z_\phi^{-2}Z_\lambda) = \sum_{n=1}^{\infty}\frac{G_n(\lambda)}{\epsilon^n}$$

We can calculate $G_1 = c_1 - 2a_1 = \frac{3\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$. As $\ln\lambda_0 = G + \ln\lambda + \epsilon\ln\tilde{\mu}$. From the independence of λ_0, we can derive

$$\left(1 + \frac{\lambda G_1'}{\epsilon} + \cdots\right)\frac{d\lambda}{d\ln\mu} + \epsilon\lambda = 0$$

$\frac{d\lambda}{d\ln\mu}$ is the rate at which λ must change to compensate for a small change in $\ln\mu$. If compensation is possible at all, this rate should be finite in the $\epsilon \to 0$ limit. In a renormalizable theory, we should have

$$\frac{d\lambda}{d\ln\mu} = -\epsilon\lambda + \beta(\lambda)$$

So

$$\beta(\lambda) = \lambda^2 G_1'(\lambda)$$

In ϕ^4 theory, we have

$$\beta(\lambda) = \frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3)$$

Define

$$M(\lambda,\epsilon) \equiv \ln(Z_m^{1/2} Z_\phi^{-1/2}) = \sum_{n=1}^{\infty} \frac{M_n(\lambda)}{\epsilon^n}$$

We can calculate $M_1 = \frac{1}{2}b_1 - \frac{1}{2}a_1 = \frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)$. As $\ln m_0 = M + \ln m$, define the anomalous dimension of the mass

$$\gamma_m(\lambda) \equiv \frac{1}{m}\frac{dm}{d\ln\mu}$$

From the independence of m_0, we can derive

$$\gamma_m(\lambda) = \lambda M_1'$$

In ϕ^4 theory, we have

$$\gamma_m(\lambda) = \frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)$$

We can expand $\ln Z_\phi$ as

$$\ln Z_\phi = \frac{a_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the field

$$\gamma_\phi(\lambda) \equiv \frac{1}{2}\frac{d\ln Z_\phi}{d\ln\mu}$$

We can derive

$$\gamma_\phi(\lambda) = -\frac{1}{2}\lambda a_1'$$

In ϕ^4 theory, we have

$$\gamma_\phi(\lambda) = \mathcal{O}(\lambda^2)$$

23.10.3 Callen-Symanzik equation

Consider the n point green function,

$$G^{(n)}(x_1,\cdots,x_n) \equiv \langle\Omega|T\phi(x_1)\cdots\phi(x_n)|\Omega\rangle_C$$

As $G_0^{(n)} = Z_\phi^{n/2} G^{(n)}$, from the independence of bare Green's function, we have

$$\left(\frac{\partial}{\partial\ln\mu} + \beta(\lambda)\frac{\partial}{\partial\lambda} + \gamma_m(\lambda)m\frac{\partial}{\partial m} + n\gamma_\phi(\lambda)\right) G^n(x_1,\cdots,x_n;\lambda,m,\mu) = 0$$

From now on, we will focus on the ϕ^4 theory in massless limit. Firstly, consider the two point Green's function in momentum space $G^{(2)}(p)$, we can express its dependence on p and μ as

$$G^{(2)} = \frac{-i}{p^2} g(-p^2/\mu^2, \lambda(\mu))$$

Then the C-S equation can be written as

$$\left[p\frac{\partial}{\partial p} - \beta(\lambda)\frac{\partial}{\partial\lambda} + 2 - 2\gamma_\phi\right] G^{(2)}(p,\lambda(\mu),\mu) = 0$$

Here, $p \equiv (-p^2)^{1/2}$ In the free theory, β and γ vanish and we recover the trivial result

$$G^{(2)}(p) = \frac{-i}{p^2}$$

In an interacting theory, the solution to the C-S equation can be expressed as

$$G^{(2)}(p, \lambda_0, \mu_0) = G^{(2)}(\mu_0, \lambda_p, \mu_0) \exp\left(-\int_{p'=\mu_0}^{p'=p} d\ln(p'/\mu_0) \cdot 2[1 - \gamma_\phi(\lambda_{p'})]\right)$$

Here, $\lambda_0 = \lambda(\mu_0)$, λ_p satisfy the following equation

$$\frac{\partial}{\partial \ln p}\lambda_p(p, \lambda_0) = \beta(\lambda_p) \quad \lambda_p(\mu_0, \lambda_0) = \lambda_0$$

The solution can be checked directly by noticing that

$$\left(p\frac{\partial}{\partial p} - \beta(\lambda_0)\frac{\partial}{\partial \lambda_0}\right)\lambda_p(p, \lambda_0) = 0$$

A convenient way of writing the solution is

$$G^{(2)}(p, \lambda_0, \mu_0) = \frac{-i}{p^2} g^{(2)}(\mu_0, \lambda_p, \mu_0) \exp\left(2\int_{\mu_0}^{p} d\ln(p'/\mu_0)\gamma_\phi(\lambda_{p'})\right)$$

And $g(\lambda_p) = 1 + O(\lambda_p^2)$

Now consider the connected four-point function of ϕ^4 theory evaluated at spacelike momenta p_i such that $p_i^2 = P^2$, $p_i \cdot p_j = 0$, so that s, t, and u are of order $-P^2$. To leading order in perturbation theory, this function is given by

$$G^{(4)}(P) = \left(\frac{i}{P^2}\right)^4 (-i\lambda)$$

The solution of C-S equation is

$$G^{(4)}(p, \lambda_0, \mu_0) = \frac{1}{P^8} g^{(4)}(\mu_0, \lambda_p, \mu_0) \exp\left(4\int_{\mu_0}^{p} d\ln(p'/\mu_0)\gamma_\phi(\lambda_{p'})\right)$$

And $g(\lambda_p) = -i\lambda_p + O(\lambda_p^2)$.

The ordinary Feynman perturbation series for a Green's function depends both on the coupling constant λ and on the dimensionless parameter $\ln(p^2/\mu_0^2)$. The perturbation theory can be badly behaved even when λ is small if the ratio p^2/μ_0^2 is large. The solutions of C-S equation reorganize this dependence into a function of the running coupling constant and an exponential scale factor.

The first factor of the solution is a function of the running coupling constant, evaluated at the momentum scale p. If p were of order μ_0, this function would essentially be the ordinary perturbation evaluation of the Green's function. We can make use of this same expression at the scale p, but to replace λ_0 with a new coupling constant λ_p appropriate to that scale. The exponential factor is the accumulated field strength rescaling of the correlation function from the reference point μ_0 to the actual momentum p at which the Green's function is evaluated.

This factor receives a multiplicative contribution from each intermediate scale between μ and p.
In ϕ^4 theory, we have

$$\frac{d}{d\ln p}\lambda_p = \frac{3\lambda_p^2}{16\pi^2}$$

The solution is

$$\lambda_p = \frac{\lambda_0}{1 - \frac{3\lambda_0}{16\pi^2}\ln(p/\mu_0)}$$

This expression for the running coupling constant goes to zero at a logarithmic rate as $p \to 0$. If we expand the running coupling constant λ_p in powers of λ_0, we find that the successive powers of the coupling constant are multiplied by powers of logarithms,

$$\lambda_0^{n+1}(\ln p/\mu_0)^n$$

which become large and invalidate a simple perturbation expansion for p much greater or much less than μ_0. If the running coupling constant becomes large, as happens in ϕ^4 theory for $p \to \infty$, the perturbation expansion will break down anyway, and we will need more advanced methods. However, if the running coupling constant becomes small, as for ϕ^4 theory as $p \to 0$, we will have successfully organized the powers of logarithms into a meaningful and controlled expression.

23.10.4 Running of coupling constants

In the limit of $\epsilon \to 0$, the coupling constant satisfies the differential equation

$$\frac{\partial\lambda}{\partial\ln\mu} = \beta(\lambda)$$

Three behaviours are possible in the region of small λ:

1. $\beta(\lambda) > 0$
2. $\beta(\lambda) = 0$
3. $\beta(\lambda) < 0$

In theories of the first class, the running coupling constant goes to zero in the infra-red, leading to definite predictions about the small-momentum behaviour of the theory. However, the running coupling constant becomes large in the region of high momenta. Thus the short-distance behaviour of the theory cannot be computed using Feynman diagram perturbation theory.A Feynman diagram analysis is useful in such theories if one is mainly interested in large-distance or macroscopic behaviour.
In theories of the second class, the coupling constant does not flow. In these theories, the running coupling constant is independent of the momentum scale, and thus equal to the bare coupling. This means that there can be no ultraviolet divergences in the relation of coupling constants. The only possible ultraviolet divergences in such theories are those associated with field rescaling, which automatically cancel in the computation of S-matrix elements. Such

theories are called finite quantum field theories. Before the emergence of our modern understanding of renormalization, these theories would have been embraced as the solution to the problem of ultraviolet infinities. But in fact the known finite field theories in four dimensions are very special constructions the so-called gauge theories with extended supersymmetry with no known physical application.
In theories of the third class, the running coupling constant becomes large in the large-distance regime and becomes small at large momenta or short distances. Such theories are called asymptotically free. In theories of this class, the short-distance behaviour is completely solvable by Feynman diagram methods. Though ultraviolet divergences appear in every order of perturbation theory, the renormalization group tells us that the sum of these divergences is completely harmless.
In the region of strong coupling, the approximation we have made, ignoring the higher-order terms in the β function is no longer valid. It is a logical possibility that the leading order term is positive while the higher terms of the β function are negative, so that the β function has a zero at a non-zero value λ_*. When λ_* approaches this value, the renormalization group flow slows to a halt; thus $\lambda = \lambda_*$ would be a non-trivial fixed point of the renormalization group.
If the β function behaves in the vicinity of the fixed point as

$$\beta \approx -B(\lambda - \lambda_*)$$

where B is a positive constant. For λ near λ_*

$$\frac{d}{d\ln\mu}\lambda \approx -B(\lambda - \lambda_*)$$

The solution of this equation is

$$\lambda(\mu) = \lambda_* + C(\frac{\mu_0}{\mu})^B$$

Thus, λ indeed tends to λ_* as $\mu \to \infty$, and the rate of approach is governed by the slope of the β function at the fixed point. The fixed point here is called ultraviolet-stable fixed point.
If p is sufficiently large, $\lambda(p)$ is close to λ_*.In the massless limit, the solution of C-S equation for two point Green function in momentum space becomes

$$G^{(2)}(p, \lambda_0, \mu_0) \approx G^{(2)}(\mu_0, \lambda_*, \mu_0) \exp\left(-\ln(p/\mu_0) \cdot 2[1 - \gamma_\phi(\lambda_*)]\right)$$

Thus the naive scaling law $G(p) \sim p^{-2}$ is changed to $G(p) \sim p^{-2[1-\gamma_\phi(\lambda_*)]}$. This has applications in the theory of critical phenomena.
A similar behaviour is possible in an asymptotically free theory. If the β function behaves in the vicinity of the fixed point as

$$\beta \approx B(\lambda - \lambda_*)$$

where B is a positive constant. the running coupling constant will tend to a fixed point as $\mu \to 0$. The fixed point is called infrared-stable fixed points.

23.11 Spontaneous symmetry breaking

23.11.1 Effective action

$$Z[J] = e^{-iE[J]} = \int \mathcal{D}\phi \exp\left[i\int d^4x(\mathcal{L}[\phi] + J\phi)\right]$$

Define

$$\phi_{\text{cl}}(x) \equiv \langle\Omega|\phi(x)|\Omega\rangle_J$$

So, we can derive

$$\frac{\delta}{\delta J(x)}E[J] = -\phi_{\text{cl}}(x)$$

Define effective action as

$$\Gamma[\phi_{\text{cl}}] \equiv -E[J] - \int d^4y J(y)\phi_{\text{cl}}(y)$$

Suppose $\mathcal{L}$ is invariant under transformation U, i.e. $\mathcal{L}(U\phi) = \mathcal{L}(\phi)$. Then we can prove than effective action Γ is also invariant under transformation U, i.e. $\Gamma(U\phi_{\text{cl}}) = \Gamma(\phi_{\text{cl}})$.

☞ **Proof:**

$$U\phi_{\text{cl}}(x) = \langle\Omega|U\phi(x)|\Omega\rangle_J = \frac{\int \mathcal{D}\phi e^{i\int \mathcal{L}(\phi)+J\phi}U\phi(x)}{\int \mathcal{D}\phi e^{i\int \mathcal{L}(\phi)+J\phi}}$$

Define $J' = \frac{J\phi}{U\phi}$, and we suppose the measure of path integral is invariant under transformation U, then we have

$$U\phi_{\text{cl}}(x) = \frac{\int \mathcal{D}U\phi e^{i\int \mathcal{L}(U\phi)+J'U\phi}U\phi(x)}{\int \mathcal{D}U\phi e^{i\int \mathcal{L}(U\phi)+J'U\phi}} = \frac{\int \mathcal{D}\phi e^{i\int \mathcal{L}(\phi)+J'\phi}\phi(x)}{\int \mathcal{D}\phi e^{i\int \mathcal{L}(\phi)+J'\phi}} = \langle\Omega|\phi(x)|\Omega\rangle_{J'}$$

On the one hand, we have

$$\Gamma[U\phi_{\text{cl}}] = E[J'] - \int d^4y J'(y)U\phi_{\text{cl}}(y) = E[J'] - \int d^4y J(y)\phi_{\text{cl}}(y)$$

On the other hand, we have

$$\begin{aligned} Z[J'] &= \int \mathcal{D}\phi \exp\left[i\int d^4x\mathcal{L}(\phi) + J'\phi\right] = \int \mathcal{D}U\phi \exp\left[i\int d^4x\mathcal{L}(U\phi) + J'U\phi\right] \\ &= \int \mathcal{D}\phi \exp\left[i\int d^4x\mathcal{L}(\phi) + J\phi\right] = Z[J] \end{aligned}$$

So, $\Gamma(U\phi_{\text{cl}}) = \Gamma(\phi_{\text{cl}})$. □

We can further verify that

$$\frac{\delta}{\delta\phi_{\text{cl}}(x)}\Gamma[\phi_{\text{cl}}] = -J(x)$$

If the external source is set to zero, the effective action satisfy the equation

$$\frac{\delta}{\delta\phi_{\text{cl}}(x)}\Gamma[\phi_{\text{cl}}] = 0$$

The solution to this equation are the values of $\langle\phi(x)\rangle$ in the stable quantum states of the theory. For a translational-invariant vacuum state, we will find a solution in which ϕ_{cl} is independent of x. For simplicity, will assume the vacuum state in our following discussion are all

translational-invariant. So, if T is the time extent of the region and V is its three dimensional volume, we can define the effective potential of the field by

$$\Gamma[\phi_{\rm cl}] = -(VT)\cdot V_{\rm eff}(\phi_{\rm cl})$$

The condition that $\Gamma[\phi_{\rm cl}]$ has an extreme then reduces to the simple equation

$$\frac{\partial}{\partial\phi_{\rm cl}}V_{\rm eff}(\phi_{\rm cl}) = 0$$

A system with spontaneously broken symmetry will have several minimum of $V_{\rm eff}$, all with the same energy by virtue of the symmetry. The choice of one among these vacuum is the spontaneous symmetry breaking.

23.11.2 Computation of the effective action

Decompose the Lagrangian into a piece depending on renormalized parameters and one containing the counter-terms

$$\mathcal{L} = \mathcal{L}_1 + \delta\mathcal{L}$$

Define J_1 by

$$\left.\frac{\delta\mathcal{L}_1}{\delta\phi}\right|_{\phi=\phi_{\rm cl}} + J_1(x) = 0$$

Define δJ by

$$J(x) = J_1(x) + \delta J(x)$$

So, we have

$$e^{-iE[J]} = \int \mathcal{D}\phi e^{i\int d^4x(\mathcal{L}_1+J_1\phi)}e^{i\int d^4x(\delta\mathcal{L}+\delta J\phi)}$$

Replace ϕ by $\phi_{\rm cl}+\eta$,

$$\begin{aligned}\int d^4x\,(\mathcal{L}_1+J_1\phi) &= \int d^4x\,(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl}) + \int d^4x\,\eta(x)\left(\frac{\delta\mathcal{L}_1}{\delta\phi}+J_1\right)\\ &+ \frac{1}{2}\int d^4x\,d^4y\,\eta(x)\eta(y)\frac{\delta^2\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)}\\ &+ \frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)}+\cdots\end{aligned}$$

The term linear in η vanishes by definition of J_1. Then, put back the effects of the counter-term Lagrangian,writing it as

$$(\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\phi_{\rm cl}) + (\delta\mathcal{L}[\phi_{\rm cl}+\eta] - \delta\mathcal{L}[\phi_{\rm cl}] + \delta J\eta)$$

Define

$$\mathcal{L}_2 = \left(\frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)}+\cdots\right)+(\delta\mathcal{L}[\phi_{\rm cl}+\eta]-\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\eta)$$

So

$$e^{-iE[J]} = C_1 e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int\left(\frac{1}{2}\eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta+I\eta\right)}\Bigg|_{I=0}$$

where

$$C_1 \equiv \exp\left[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}} + \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}})\right]$$

If we define propagator D_F as

$$D_F \equiv i\left(\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right)^{-1}$$

We have

$$Z[J] = e^{-iE[J]} = C_1 Z_0[0] e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int\left(-\frac{1}{2}ID_FI\right)}\Big|_{I=0}$$

where

$$Z_0[0] \equiv \int\mathcal{D}\eta e^{\frac{i}{2}\int\eta\left(\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right)\eta}$$

Similar to the procedure in the perturbation theory for path integral quantization, we can get a perturbation expansion for $iE[J]$ using connected Feynman diagram,

$$-iE[J] = i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}} + \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}) + \log(Z_0[0]) + \text{ connected diagrams}$$

From this equation, Γ follows directly:

$$\Gamma[\phi_{\text{cl}}] = \int d^4x\mathcal{L}_1[\phi_{\text{cl}}] - i\log(Z_0[0]) - i\text{ connected diagrams} + \int d^4x\delta\mathcal{L}[\phi_{\text{cl}}]$$

Notice that there are no terms remaining that depend explicitly on J; thus, Γ is expressed as a function of ϕ_{cl} , as it should be. The Feynman diagrams contributing to $\Gamma[\phi_{\text{cl}}]$ have no external lines, and the simplest ones turn out to have two loops. The lowest-order quantum correction to Γ is given by the functional determinant.
The last term provides a set of counter-terms that can be used to satisfy the renormalization conditions on Γ and, in the process, to cancel divergences that appear in the evaluation of the functional determinant and the diagrams. The renormalization conditions will determine all of the counter-terms in $\delta\mathcal{L}$. However, the formalism we have constructed contains a new counter-term δJ. That coefficient is determined by $\langle\eta\rangle = 0$. In practice, we will satisfy this condition by simply ignoring any one-particle-irreducible one-point diagram, since any such diagram will be cancelled by adjustment of δJ.

23.11.3 The effective action as a generating functional

$E[J]$ is called the generating of connected correlation functions,

$$\frac{\delta^n E[J]}{\delta J(x_1)\cdots\delta J(x_n)} = i^{n+1}\langle\phi(x_1)\cdots\phi(x_n)\rangle\text{conn}$$

The effective action $\Gamma[\phi_{\text{cl}}]$ is the generating functional of one-particle-irreducible correlation functional,

$$\frac{\delta\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x)} = 0$$

$$\frac{\delta^2\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x)\delta\phi_{\text{cl}}(y)} = iD^{-1}(x,y)$$

Here, $D(x,y) = \langle\phi(x)\phi(y)\rangle_{\text{conn}}$. When $n \geq 3$,

$$\frac{\delta^n\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x_1)\cdots\delta\phi_{\text{cl}}(x_n)} = -i\langle\phi(x_1)\cdots\phi(x_n)\rangle_{\text{1PI}}$$

The proof of statements above can be found in chapter 10.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

The chapter 21 of *Quantum field theory (M. Srednicki)* gives an constructive way to define the effective action.

$$\begin{aligned}\Gamma[\phi] &\equiv \frac{1}{2}\int\frac{d^dk}{(2\pi)^d}\tilde{\phi}(-k)(-k^2-m^2-M^2(k^2))\tilde{\phi}(k)\\ &+ \frac{1}{n!}\int\frac{d^dk_1}{(2\pi)^d}\cdots\frac{d^dk_n}{(2\pi)^d}(2\pi)^d\delta(k_1+\cdots+k_n)V_n(k_1,\cdots,k_n)\tilde{\phi}(k_1)\cdots\tilde{\phi}(k_n)\end{aligned}$$

Here $\tilde{\phi}(k) = \int d^dxe^{-ikx}\phi(x)$, and $iV_n(k_1,\cdots,k_n)$ equals the value of 1PI Feynman diagram in momentum space. The effective action has the property that the tree-level Feynman diagrams it generates give the complete scattering amplitude of the original theory. The author also proved that this definition is equivalent to the definition from *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*.

23.11.4 Renormalization and symmetry

Consider first the computation of the effective potential for constant classical fields, in a field theory with an arbitrary number of fields ϕ^i. The effective potential has mass dimension 4, so we expect that $V_{\text{eff}}(\phi_{\text{cl}})$ will have divergent terms up to Λ^4. To understand these divergences, expand $V_{\text{eff}}(\phi_{\text{cl}})$ in a Taylor series:

$$V_{\text{eff}}(\phi_{\text{cl}}) = A_0 + A_2^{ij}\phi^i_{\text{cl}}\phi^j_{\text{cl}} + A_4^{ijkl}\phi^i_{\text{cl}}\phi^j_{\text{cl}}\phi^k_{\text{cl}}\phi^l_{\text{cl}} + \cdots$$

In theories without a symmetry of $\phi \to -\phi$, there might also be terms linear and cubic in ϕ^i; we omit these for simplicity. The coefficients A_0, A_2, A_4 have mass dimension, respectively, 4, 2, and 0; thus we expect them to contain Λ^4, Λ^2, and $\log\Lambda$ divergences, respectively. The power-counting analysis predicts that all higher terms in the Taylor series expansion should be finite.

The constant term A_0 is independent of ϕ_{cl}; it has no physical significance. However, the divergences in A_2 and A_4 appear in physical quantities, since these coefficients enter the inverse propagator and the irreducible four-point function and therefore appear in the computation of S-matrix elements. There is one further coefficient in the effective action that has non-negative mass dimension by power counting; this is the coefficient of the term quadratic in $\partial_\mu\phi_{\text{cl}}$, which appears when the effective action is evaluated for a non-constant background field:

$$\Delta\Gamma[\phi_{\text{cl}}] = \int d^4xB_2^{ij}\partial_\mu\phi^i_{\text{cl}}\partial^\mu\phi^j_{\text{cl}}$$

All other coefficients in the Taylor expansion of the effective action in powers of $\phi_{\rm cl}$ are finite by power counting.

We can now argue that the counter-terms of the original Lagrangian suffice to remove the divergences that might appear in the computation of $\Gamma[\phi_{\rm cl}]$. The argument proceeds in two steps. We first use the BPHZ theorem to argue that the divergences of Green's functions can be removed by adjusting a set of counter-terms corresponding to the possible operators that can be added to the Lagrangian with coefficients of mass dimension greater than or equal to zero. The coefficients of these counter-terms are in 1-to-1 correspondence with the coefficients A_2, A_4, and B_2 of the effective action. Next, we use the fact that the effective action is manifestly invariant to the original Symmetry group of the model. This is true even if the vacuum state of the model has spontaneous symmetry breaking, since the method we presented for computing the effective action is manifestly invariant to the original symmetry of the Lagrangian. Combining these two results, we conclude that the effective action can always be made finite by adjusting the set of counter-terms that are invariant to the original symmetry of the theory, even if this symmetry is spontaneously broken.

23.11.5 Goldstone's theorem

Theorem 23.4 Goldstone's theorem

Goldstone's theorem examines a generic continuous symmetry which is spontaneously broken; i.e., its currents are conserved, but the ground state is not invariant under the action of the corresponding charges. Then, necessarily, new massless (or light, if the symmetry is not exact) scalar particles appear in the spectrum of possible excitations. There is one scalar particle - called a Nambu-Goldstone boson — for each generator of the symmetry that is broken, i.e., that does not preserve the ground state. ♣

☞ **Proof:** A general continuous symmetry transformation has the form

$$\phi^a \to \phi^a + \alpha\Delta^a(\phi)$$

where α is an infinitesimal parameter and Δ^a is some function of all the ϕ's. Specialize to constant fields; then the derivative terms in $\mathcal{L}$ vanish and the potential alone must be invariant. This condition can be written

$$V(\phi^a) = V(\phi^a + \alpha\Delta^a(\phi)) \quad \text{or} \quad \Delta^a(\phi)\frac{\partial}{\partial\phi^a}V(\phi) = 0$$

The effective potential $V_{\rm eff}$ encapsulates the full solution to the theory, including all orders of quantum corrections. At the same time, it satisfies the general properties of the classical potential: It is invariant to the symmetries of the theory, and its minimum gives the vacuum expectation value of $\phi_{\rm cl}$. So

$$\Delta^a(\phi)\frac{\partial}{\partial\phi^a}V_{\rm eff}(\phi) = 0$$

Now differentiate with respect to ϕ^b, and set $\phi = \phi_{\rm cl}$

$$0 = \left(\frac{\partial\Delta^a}{\partial\phi^b}\right)_{\phi_{\rm cl}}\left(\frac{\partial V_{\rm eff}}{\partial\phi^a}\right)_{\phi_{\rm cl}} + \Delta^a(\phi_{\rm cl})\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V_{\rm eff}\right)_{\phi_{\rm cl}}$$

The first term vanishes since $\phi_{\rm cl}$ is a minimum of $V_{\rm eff}$, so the second term must also vanish. If the transformation leaves $\phi_{\rm cl}$ unchanged (i.e., if the symmetry is respected by the ground state), then $\Delta^a(\phi_{\rm cl}) = 0$ and this relation is trivial. A spontaneously broken symmetry is precisely one for which $\Delta^a(\phi_{\rm cl}) \neq 0$; in this case $\Delta^a(\phi_{\rm cl})$ is the vector with eigenvalue zero. We now argue that the presence of such a zero eigenvalue implies the existence of a massless scalar particle. Effective action's second functional derivative is the inverse propagator

$$i\tilde{D}^{-1}_{ij}(p^2) = \int d^4x e^{-ip(x-y)} \frac{\delta\Gamma}{\delta\phi^i\delta\phi^j}(x,y)\bigg|_{\phi=\phi_{\rfloor\ddagger}}$$

A particle of mass 0 corresponds to a zero eigenvalue of this matrix equation at $p^2 = 0$. Now set $p = 0$. This implies $\frac{\delta\Gamma}{\delta\phi^i\delta\phi^j}(x,y)$ has a zero eigenvalue. This is equivalent to

$$\frac{\partial^2}{\partial\phi^i_{\rm cl}\partial\phi^j_{\rm cl}} V_{\rm eff}$$

has a zero eigenvalue. This completes the proof of Goldstone's theorem. □

23.12 Linear sigma model

23.12.1 Introduction

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi^i\partial^\mu\phi^i + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}[(\phi^i)^2]^2$$

The Lagrangian is invariant under the symmetry

$$\phi^i \to R^{ij}\phi^j$$

for any $N \times N$ orthogonal group in N dimensions, also called the N-dimensional orthogonal group or simply $O(N)$. In classical theory, the lowest-energy classical configuration is a constant field ϕ^i_0, whose value is chosen to minimize the potential

$$V = -\frac{1}{2}\mu^2(\phi^i)^2 + \frac{\lambda}{4}[(\phi^i)^2]^2$$

This potential is minimized for any ϕ^i_0 that satisfies

$$(\phi^i)^2 = \frac{\mu^2}{\lambda}$$

This condition determines only the length of the vector ϕ^i_0, its direction is arbitrary. It is conventional to choose coordinates so that ϕ^i_0 points in the Nth direction

$$\phi^i_0 = (0, 0, \cdots, 0, v) \quad v \equiv \frac{\mu}{\sqrt{\lambda}}$$

We can now define a set of shifted fields by writing

$$\phi^i = (\pi^k, v + \sigma) \quad k = 1, \cdots, N-1$$

It is now straightforward to rewrite the Lagrangian in terms of the π and σ fields. The result is

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2}(\partial_\mu \pi^k)^2 - \frac{1}{2}(\partial_\mu \sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 \\ &- \sqrt{\lambda}\mu\sigma^3 - -\sqrt{\lambda}\mu(\pi^k)^2\sigma - \frac{\lambda}{4}\sigma^4 - \frac{\lambda}{2}(\pi^k)^2\sigma^2 - \frac{\lambda}{4}[(\pi^k)^2]^2\end{aligned}$$

We obtain a massive σ field and also a set of $N-1$ massless π fields. The original $O(N)$ symmetry is hidden when we choose a specific ϕ_0^i for vacuum state, leaving only the subgroup $O(N-1)$, which rotates the π fields among themselves. Given the theory in the previous section, our discussion here will remain valid when quantum correction is considered.

23.12.2 Renormalization

From this expression of the Lagrangian written in terms of shifted fields, we can read off the Feynman rules:

Figure 23.16: Feynman rules for the linear sigma model

Using these Feynman rules, we can compute tree-level amplitudes without difficulty. Diagrams with loops, however, will often diverge. For the amplitude with N_e external legs, the superficial degree of divergence is

$$D = 4 - N_e$$

The linear sigma model has eight different superficially divergent amplitudes and several of these have $D > 0$ and therefore can contain more than one in nite constant.
Yet we have only three counterterms:

$$\mathcal{L}_{ct} = -\frac{1}{2}\delta_Z \partial_\mu \phi^i \partial^\mu \phi^i - \frac{1}{2}\delta_\mu (\phi^i)^2 - \frac{\delta_\lambda}{4}[(\phi^i)^2]^2$$

Written in terms of σ and π fields, we have

$$\begin{aligned}\mathcal{L}_{ct} &= -\frac{\delta_Z}{2}(\partial_\mu \pi^k)^2 - \frac{1}{2}(\delta_\mu + \delta_\lambda v^2)(\pi^k)^2 - \frac{\delta_Z}{2}(\partial_\mu \sigma)^2 - \frac{1}{2}(\delta_\mu + 3\delta_\lambda v^2)\sigma^2 \\ &- (\delta_\mu v + \delta_\lambda v^3)\sigma - \delta_\lambda v\sigma(\pi^k)^2 - \delta_\lambda v\sigma^3 - \frac{\delta_\lambda}{4}[(\pi^k)^2]^2 - \frac{\delta_\lambda}{2}\sigma^2(\pi^k)^2 - \frac{\delta_\lambda}{4}\sigma^4\end{aligned}$$

Figure 23.17: Divergent amplitudes in the linear sigma model

Figure 23.18: Feynman rules for counterterm vertices in the linear sigma model

Figure 23.19: Renormalization conditions ($m^2 = 2\mu^2$)

We can get the Feynman rules associated with these counterterms.
Following the renormalization conditions above, the one-loop corrections for linear sigma model has been calculated in section 11.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*. There are two important results.

- All the divergence up to one loop will be cancelled by adjusting three counterterms. Apparently, the divergent part of the diagram is unaffected by the symmetry breaking.
- The propagator of $\pi\pi$ has a pole at $p^2 = 0$ after one loop correction, i.e. π particles remain massless after one loop correction.

23.12.3 Effective action

We begin again with the Lagrangian

$$\mathcal{L}_1 = -\frac{1}{2}\partial_\mu\phi^i\partial^\mu\phi^i + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}[(\phi^i)^2]^2$$

Expand about the classical field $\phi^i = \phi^i_{\text{cl}} + \eta^i$, and we assume the vacuum is translational invariant. Then we have

$$\mathcal{L}_1 = -\frac{1}{2}(\partial_\mu \eta)^2 + \frac{1}{2}\mu^2(\eta^i)^2 - \frac{\lambda}{2}[(\phi^2_{\text{cl}})(\eta^i)^2 + 2(\phi^i_{\text{cl}}\eta^i)^2] + \cdots$$

From the terms quadratic in η, we can read off

$$\frac{\delta^2 \mathcal{L}_1}{\delta\phi^i \delta\phi^j} = \partial^2\delta_{ij} + \mu^2\delta_{ij} - \lambda[(\phi^k_{\text{cl}})^2\delta_{ij} + 2\phi^i_{\text{cl}}\phi^j_{\text{cl}}]$$

We choose the vacuum state by demanding ϕ^i_{cl} points in the Nth direction

$$\phi^i_{\text{cl}} = (0, \cdots, \phi_{\text{cl}})$$

Then the operator is just equal to the Klein-Gordon operator $(\partial^2 - m_i^2)$, where

$$m_i^2 = \begin{cases} \lambda\phi^2_{\text{cl}} - \mu^2 & i = 1, \cdots, N-1 \\ 3\lambda\phi^2_{\text{cl}} - \mu^2 & i = N \end{cases}$$

We would perform the calculation of $\log Z_0[0]$ in the next subsection. Here, we just list the result:

$$\log Z_0[0] = \frac{i}{2}\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{\frac{d}{2}}VT$$

So, up to one loop corrections, we can get

$$V_{\text{eff}} = -\frac{1}{2}\mu^2\phi^2_{\text{cl}} + \frac{\lambda}{4}\phi^4_{\text{cl}} - \frac{1}{2}\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}[(N-1)(\lambda\phi^2_{\text{cl}} - \mu^2)^{\frac{d}{2}} + (3\lambda\phi^2_{\text{cl}} - \mu^2)^{\frac{d}{2}}] + \frac{1}{2}\delta_\mu\phi^2_{\text{cl}} + \frac{1}{4}\delta_\lambda\phi^4_{\text{cl}}$$

And if we want V_{eff} is finite for terms involving ϕ_{cl}, we can get

$$\delta_\lambda = \frac{2\lambda^2(N+8)}{(4\pi)^2} \times \frac{1}{4-d} + \text{ finite terms}$$

$$\delta_\mu = -\frac{2\lambda\mu^2(N+2)}{(4\pi)^2} \times \frac{1}{4-d} + \text{ finite terms}$$

It is the same result as that in previous section.

23.12.4 Functional determinants

$$Z_0[0] \equiv \prod_{i=1}^N Z_i = \prod_{i=1}^N \int \mathcal{D}\eta e^{\frac{i}{2}\int \eta(\partial^2 - m_i^2)\eta}$$

Here, m_i is a function of ϕ_{cl}. We want to get $\log Z_0[0]$ as a function of ϕ_{cl} and the constant infinite shift of $\log Z_0[0]$ will be dropped in our calculation. We treat $-\frac{1}{2}m_i^2\eta^2$ as a perturbation, so we have

$$Z_i \propto e^{-\frac{im_i^2}{2}(\frac{1}{i}\frac{\delta}{\delta I})^2} \int \mathcal{D}\eta e^{i\int(-\frac{1}{2}ID_F I)}\Big|_{I=0}$$

where

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2}e^{ip(x-y)}$$

Now we can have the following Feynamn rules.

- A line from x to y is associated with $D_F(x-y)$
- A vertex joining two lines at x is associated with $-im_i^2 \int d^4x$

So we have

$$\log Z_i = \sum_I C_I$$

where C_I represents connected diagram without external source. A connected diagram without external source must have the following form

Figure 23.20: Connected Feymann diagram without external source

So, we have

$$C_n = \frac{1}{2n}\int \prod_{k=1}^{n} \frac{d^4p_k d^4x_k}{(2\pi)^4} \frac{-m_i^2}{p_k^2} \exp(ip_k(x_k - x_{k+1})) = \int d^4p\delta(0)\left(-\frac{m_i^2}{p^2}\right)^n$$

$$\log Z_i = -\frac{1}{2}VT\int \frac{d^4p}{(2\pi)^4}\sum -\frac{1}{n}\left(-\frac{m_i^2}{p^2}\right)^n = -\frac{1}{2}VT\int \frac{d^4p}{(2\pi)^4}\log(1+\frac{m_i^2}{p^2})$$

The following calculation needs tricks of wick rotation and dimension regularization and you can refer to the equation 11.72 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*.

However, recall the Gaussian integral

$$\int_{-\infty}^{\infty} e^{\left(-\frac{i}{2}\sum\limits_{i,j=1}^{n} A_{ij}x_ix_j\right)} d^nx = \sqrt{\frac{(-2\pi i)^n}{\det A}}$$

So formally, we have

$$\log Z_i = -\frac{1}{2}\log\det(-\partial_x^2 + m_i^2)\delta(x-y)$$

Define

$$M(x-y) \equiv (-\partial_x^2+m_i^2)\delta(x-y) \quad M_0(x,y) \equiv -\partial_x^2\delta(x-y) \quad M_1(y,z) \equiv \delta(y-z)+im_i^2D_F(y-z)$$

We can verify that

$$M(x,z) = \int d^4yM_0(x-y)M_1(y-z)$$

So,

$$\log\det M = \log\det M_0 + \log\det M_1 \approx \log\det M_1$$

We drop out $\log\det M_0$ because it do not contain $m(\phi_{\rm cl})$. Furthermore, we have $M_1 = I - G$, where $I = \delta(x-y)$ is the identity matrix and $G = -im_i^2 D_F$. So

$$\log\det M_1 = \mathrm{Tr}\log M_1 = \mathrm{Tr}\log(I-G) = -\frac{1}{n}\sum_{n=1}^{\infty}\mathrm{Tr}G_n$$

We can verify that

$$C_n = \frac{1}{2n}\mathrm{Tr}G_n$$

So, the function determinant method is valid.

23.13 Non-relativistic limit

23.13.1 Complex Klein-Gordon field

The Lagrangian of complex Klein-Gordon field is

$$\mathcal{L} = -\partial^\mu\Phi^\dagger\partial_\mu\Phi - m^2\Phi^\dagger\Phi$$

The canonical momentum is

$$\pi = \dot{\Phi}^\dagger \quad \pi^\dagger = \dot{\Phi}$$

The commutation relations are

$$[\phi(\boldsymbol{x},t), \pi(\boldsymbol{y},t)] = i\delta(\boldsymbol{x}-\boldsymbol{y})$$

Since the equation of motion is

$$(\partial^2 - m^2)\Phi = 0$$

we have the Fourier expansion

$$\Phi = \int \widetilde{dp}[b(\boldsymbol{p})e^{ipx} + c^\dagger(\boldsymbol{p})e^{-ipx}] \quad \Phi^\dagger = \int \widetilde{dp}[b^\dagger(\boldsymbol{p})e^{-ipx} + c(\boldsymbol{p})e^{ipx}]$$

We can get the commutation relations for creation and annihilation operators

$$[b(\boldsymbol{p}), b^\dagger(\boldsymbol{q})] = (2\pi)^3 2\omega\delta(\boldsymbol{p}-\boldsymbol{q}) \quad [c(\boldsymbol{p}), c^\dagger(\boldsymbol{q})] = (2\pi)^3 2\omega\delta(\boldsymbol{p}-\boldsymbol{q})$$

After work out the commutation relations between H, P and $b, b^\dagger, c, c^\dagger$, we can conclude that $b^\dagger(\boldsymbol{p})(c^\dagger(\boldsymbol{p}))$ creates a b(c) particle with momentum $\boldsymbol{p}$, and $b(\boldsymbol{p})(c(\boldsymbol{p}))$ annihilates a b(c) particle with momentum $\boldsymbol{p}$. They share the same mass m.
We notice that $\mathcal{L}$ is invariant under change $\Phi \to \Phi e^{i\alpha}$. Noether theorem implies that complex Klein-Gordon field has a conserve charge

$$Q = i\int d^3x(\dot{\Phi}^\dagger\Phi - \Phi^\dagger\dot{\Phi}) = \int \widetilde{dp}(c^\dagger(\boldsymbol{p})c(\boldsymbol{p}) - b^\dagger(\boldsymbol{p})b(\boldsymbol{p})) = N_c - N_b$$

Now we can interpret c-particle as the antiparticle of b-particle. And The number of anti-particles minus the number of particles is a conserved quantity, i.e. particles and anti-particles must be created in pair.

23.13.2 Non-relativistic limit

We decompose the complex Klein-Gordon equation as

$$\Phi(x) = \frac{1}{\sqrt{2m}} e^{-imt} \psi(x)$$

The The non-relativistic limit of a particle is $|\boldsymbol{p}| \ll m$. After a Fourier transform, this is equivalent to saying that $\dot{\psi} \ll m\psi$. In this limit, we have

$$\frac{\partial \phi^\dagger}{\partial t} \frac{\partial \phi}{\partial t} - m^2 \phi^\dagger \phi \approx \frac{i}{2} \left(\phi^\dagger \frac{\partial \phi}{\partial t} - \frac{\partial \phi^\dagger}{\partial t} \phi \right)$$

After integration by parts, the Lagrangian of the complex Klein-Gordon can be written as

$$\mathcal{L} = i\phi^\dagger \left(\partial_t + \frac{\nabla^2}{2m} \right) \phi$$

It is exactly the Schrödinger field in non-relativistic quantum field theory.

Chapter 24
Spin 1/2 Field

24.1 Representation of the Lorentz group

Recall the rotation of the field

$$U^{-1}(\Lambda)\phi_a(x)U(\Lambda) = \mathcal{S}_a{}^b\phi_b(\Lambda^{-1}x)$$

For infinitesimal rotation,

$$\mathcal{S}_a{}^b = \delta_a{}^b + \frac{i}{2}\delta\omega_{\alpha\beta}(S^{\alpha\beta})_a{}^b$$

The matrix $S^{\alpha\beta}$ satisfy the following commutation relation

$$[S^{\mu\nu}, S^{\rho\sigma}] = -\eta^{\nu\rho}S^{\mu\sigma} + \eta^{\sigma\mu}S^{\rho\nu} + \eta^{\mu\rho}S^{\nu\sigma} - \eta^{\sigma\nu}S^{\rho\mu}$$

Define $C_i \equiv \frac{1}{2}\epsilon_{ijk}S^{jk}$,$D_i \equiv S^{i0}$, so we can get

$$[C_i, C_j] = i\epsilon_{ijk}C_k \quad [C_i, D_j] = i\epsilon_{ijk}D_k \quad [D_i, D_j] = -i\epsilon_{ijk}C_k$$

We further define $N_i \equiv \frac{1}{2}(C_i - iD_i)$ and $N_i^\dagger \equiv \frac{1}{2}(C_i + iD_i)$, then the commutation relation will be

$$[N_i, N_j] = i\epsilon_{ijk}N_k \quad [N_i^\dagger, N_j^\dagger] = i\epsilon_{ijk}N_k^\dagger \quad [N_i, N_j^\dagger] = 0$$

We see that we have two different $SU(2)$ Lie algebras that are exchanged by hermitian conjugation. As we just discussed, a representation of the $SU(2)$ Lie algebra is specified by an integer or half integer; we therefore conclude that a representation of the Lie algebra of the Lorentz group in four spacetime dimensions is specified by two integers or half-integers n and n'.

We will label these representations as $(2n+1, 2n'+1)$; the number of components of a representation is then $(2n+1)(2n'+1)$. Different components within a representation can also be labeled by their angular momentum representations. We have $C_i = N_i + N_i^\dagger$. Thus, deducing the allowed values of j given n and n' becomes a standard problem in the addition of angular momenta. The general result is that the allowed values of j are $|n-n'|, |n-n'|+1, \cdots, n+n'$, and each of these values appears exactly once.

24.2 Spin-statistics theorem

Theorem 24.1 Spin-statistics theorem

States with identical particles of integer spin are symmetric under the interchange of the particles, while states with identical particles of half-integer spin are antisymmetric under the interchange of the particles.
This is equivalent to the statement that the creation and annihilation operators for integer spin particles satisfy canonical commutation relations, while creation and annihilation operators for halfinteger spin particles satisfy canonical anticommutation relations.
Particles quantized with canonical commutation relations are called bosons, and satisfy Bose Einstein statistics, and particles quantized with canonical anticommutation relations are called fermions, and satisfy Fermi Dirac statistics.

Roughly speaking, one way to interchange two particles is to rotate them around their midpoint by π. For a particle of spin s, this rotation will introduce a phase factor of $e^{i\pi s}$. Thus, a two-particle state with identical particles both of spin s will pick up a factor of $e^{i2\pi s}$. For s a half-integer, this will give a factor of -1; for s an integer, it will give a factor of $+1$. So, the creation and annihilation operators for integer spin particles satisfy canonical commutation relations, while creation and annihilation operators for halfinteger spin particles satisfy canonical anticommutation relations. The detailed proof can be found in chapter 12.1 and 12.2 from *Quantum Field Theory and the Standard Model (Matthew D. Schwartz)*.

24.3 Spinor field

Consider a left-handed spinor field $\psi_a(x)$, also known as a left-handed Weyl field, which is in the $(2,1)$ representation of the Lie algebra of the Lorentz group. Here the index a is a left-handed spinor index that takes on two possible values. Under a Lorentz transformation, we have

$$U(\Lambda)^{-1}\psi_a(x)U(\Lambda) = L_a{}^b(\Lambda)\psi_b(\Lambda^{-1}x)$$

For an infinitesimal transformation, we can write

$$L_a{}^b(1+\delta\omega) = \delta_a{}^b + \frac{i}{2}\delta\omega_{\mu\nu}(S_L^{\mu\nu})_a{}^b$$

$n=1, \quad n'=0$ implies that

$$(S_L^{ij}) = \frac{1}{2}\epsilon^{ijk}\sigma_k \quad (S_L^{k0}) = \frac{1}{2}i\sigma_k$$

where σ_k is Pauli matrix. Then the infinitesimal transformation can be written as

$$L(1+\delta\omega) = I - \frac{i}{2}\theta_i\sigma_i - \frac{1}{2}\beta_i\sigma_i$$

Recall that hermitian conjugation swaps the two $SU(2)$ Lie algebras that comprise the Lie algebra of the Lorentz group. Therefore, the hermitian conjugate of a field in the $(2,1)$ representation should be a field in the $(1,2)$ representation; such a field is called a right-handed spinor field or a right-handed Weyl field. We will distinguish the indices of the $(1,2)$ representation from those of the $(2,1)$ representation by putting dots over them. Thus, we write

$$[\psi_a(x)]^\dagger = \psi^\dagger_{\dot{a}}(x)$$

Under a Lorentz transformation, we have

$$U(\Lambda)^{-1}\psi^\dagger_{\dot{a}}(x)U(\Lambda) = R_{\dot{a}}{}^{\dot{b}}(\Lambda)\psi^\dagger_{\dot{b}}(x)(\Lambda^{-1}x)$$

For an infinitesimal transformation, we can write

$$R_{\dot{a}}{}^{\dot{b}}(1+\delta\omega) = \delta_{\dot{a}}{}^{\dot{b}} + \frac{i}{2}\delta\omega_{\mu\nu}(S_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}}$$

We can prove that

$$(S_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}} = -[(S_L^{\mu\nu})_a{}^b]^*$$

Define

$$\epsilon_{ab} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Group theory can show that

$$L_a{}^c(\Lambda)L_b{}^d(\Lambda)\epsilon_{cd} = \epsilon_{ab}$$

which means that ϵ_{ab} is an invariant symbol of the Lorentz group: it does not change under a Lorentz transformation that acts on all of its indices. The inverse matrix of ϵ_{ab} is

$$\epsilon^{ab} \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

So,

$$\epsilon_{ab}\epsilon^{bc} = \delta_a{}^c \quad \epsilon^{ab}\epsilon_{bc} = \delta^a{}_c$$

We can use ϵ_{ab} and and its inverse ϵ^{ab} to raise and lower left-handed spinor indices,

$$\psi^a(x) \equiv \epsilon^{ab}\psi_b(x)$$

We also notice the minus sign when we contract indices,

$$\psi^a\chi_a = -\psi_a\chi^a$$

And we can verify the following equations,

$$-L^a{}_bL_a{}^c = \delta^c_b$$

$$\psi_a(x) = \epsilon_{ab}\psi^b(x)$$

$$L^a{}_c(\Lambda)L^b{}_d(\Lambda)\epsilon^{cd} = \epsilon^{ab}$$

$$U(\Lambda)^{-1}\psi^a(x)U(\Lambda) = -L^a{}_b(\Lambda)\psi^b(\Lambda^{-1}x)$$

In right handed representation; we can also deduce the existence of an invariant symbol $\epsilon_{\dot{a}\dot{b}}$. The value of $\epsilon_{\dot{a}\dot{b}}$ is the same as that of ϵ_{ab}.
Define

$$\sigma^\mu_{a\dot{a}} \equiv (I, \vec{\sigma})$$

It is an invariant symbol of the group $(2,1)\otimes(1,2)\otimes(2,2)$
The properties of invariance symbol can be used to derive the following equations. The detailed derivation can be found in chapter 35 from *Quantum Field Theory (Mark Srednicki)*.

Proposition 24.1

1.
$$\sigma^\mu_{a\dot{a}}\sigma_{\mu b\dot{b}} = -2\epsilon_{ab}\epsilon_{\dot{a}\dot{b}}$$
$$\epsilon^{ab}\epsilon^{\dot{a}\dot{b}}\sigma^\mu_{a\dot{a}}\sigma^\nu_{b\dot{b}} = -2\eta^{\mu\nu}$$

2.
$$(S^{\mu\nu}_L)_{ab} = (S^{\mu\nu}_L)_{ba}$$
$$(S^{\mu\nu}_R)_{\dot{a}\dot{b}} = (S^{\mu\nu}_R)_{\dot{b}\dot{a}}$$

3. Define
$$\overline{\sigma}^{\mu\dot{a}a} \equiv \epsilon^{ab}\epsilon^{\dot{a}\dot{b}}\sigma^\mu_{b\dot{b}}$$
Numerically,
$$\overline{\sigma}^{\mu\dot{a}a} = (I, -\vec{\sigma})$$
So,
$$\begin{aligned}(S^{\mu\nu}_L)_a{}^b &= +\frac{i}{4}(\sigma^\mu\overline{\sigma}^\nu - \sigma^\nu\overline{\sigma}^\mu)_a{}^b \\ (S^{\mu\nu}_R)^{\dot{a}}{}_{\dot{b}} &= -\frac{i}{4}(\overline{\sigma}^\mu\sigma^\nu - \overline{\sigma}^\nu\sigma^\mu)^{\dot{a}}{}_{\dot{b}}\end{aligned}$$

We adopt the following convention: a missing pair of contracted, undotted indices is understood to be written as ${}^c{}_c$, and a missing pair of contracted, dotted indices is understood to be written as ${}_{\dot{c}}{}^{\dot{c}}$. Thus, if χ and ψ are two left-handed Weyl fields, we have

$$\chi\psi = \chi^a\psi_a \quad \chi^\dagger\psi_\dagger = \chi^\dagger_{\dot{a}}\psi^{\dagger\dot{a}}$$

We expect Weyl fields to describe spin-one-half particles, and (by the spinstatistics theorem) these particles must be fermions. Therefore the corresponding fields must anticommute, rather than commute. That is, we should have

$$\chi_a(x)\psi_b(y) = -\psi_b(x)\chi_a(x)$$

Thus, we can get

$$\chi\psi = \psi\chi$$

Using the above convention, we can derive

Proposition 24.2

1. $$(\chi\psi)^\dagger = \psi^\dagger\chi^\dagger$$

2. $$[\psi^\dagger\overline{\sigma}^\mu\chi]^\dagger = \chi^\dagger\overline{\sigma}^\mu\psi$$

24.4 Lagrangians for spinor fields

Weyl field

$$\mathcal{L} = i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi$$

Note:

$$(i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi)^\dagger = i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi - \partial_\mu(i\psi^\dagger\overline{\sigma}^\mu\psi)$$

The second term is a total divergence, and vanishes (with suitable boundary conditions on the fields at infinity) when we integrate it over d^4x to get the action S. Thus $i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi$ has the hermiticity properties necessary for a term in $\mathcal{L}$.

We can derive the equation of motion from Hamilton principle,

$$\overline{\sigma}^\mu\partial_\mu\psi = 0$$

Majorana field

If we add mass terms to the Lagrangian, we have

$$\mathcal{L} = i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi - \frac{1}{2}m\psi\psi - \frac{1}{2}m\psi^\dagger\psi^\dagger$$

The equation of motion is

$$-i\overline{\sigma}^\mu\partial_\mu\psi + m\psi^\dagger = 0 \quad -i\sigma^\mu\partial_\mu\psi^\dagger + m\psi = 0$$

Define gamma matrix as

$$\gamma^\mu \equiv \begin{bmatrix} 0 & \sigma^\mu_{a\dot{c}} \\ \overline{\sigma}^{\mu\dot{a}c} & 0 \end{bmatrix}$$

We can prove that

$$\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu}$$

Define a four-component Majorana field as

$$\Psi \equiv \begin{bmatrix} \psi_c \\ \psi^{\dagger\dot{c}} \end{bmatrix}$$

The equation of motion can be written as

$$(-i\gamma^\mu\partial_\mu + m)\Psi = 0$$

Dirac field

$$\mathcal{L} = i\chi^\dagger \overline{\sigma}^\mu \partial_\mu \chi + i\xi^\dagger \overline{\sigma}^\mu \partial_\mu \xi - \frac{1}{2} m\chi\xi - \frac{1}{2} m\xi^\dagger \chi^\dagger$$

It is invariant under

$$\chi \to e^{-i\alpha}\chi \quad \xi \to e^{i\alpha}\xi$$

Define a four-component Dirac field

$$\Psi \equiv \begin{bmatrix} \chi_a \\ \xi^{\dagger \dot{a}} \end{bmatrix}$$

So we take the hermitian conjugate of Ψ to get

$$\Psi^\dagger = (\chi^\dagger_{\dot{a}}, \xi^a)$$

Introduce the matrix

$$\beta \equiv \begin{bmatrix} 0 & \delta^{\dot{a}}{}_{\dot{c}} \\ \delta_a{}^c & 0 \end{bmatrix}$$

Given β, we define

$$\overline{\Psi} \equiv (\xi^a, \chi^\dagger_{\dot{a}})$$

Detailed calculation shows that

$$\overline{\Psi}\gamma^\mu \partial_\mu \Psi = \chi^\dagger \overline{\sigma}^\mu \partial_\mu \chi + \xi^\dagger \overline{\sigma}^\mu \partial_\mu \xi + \partial_\mu (\xi^\dagger \sigma^\mu \xi)$$

So, the Lagrangian can be written as

$$\mathcal{L} = i\overline{\Psi}\gamma^\mu \partial_\mu \Psi - m\overline{\Psi}\Psi$$

It is invariant under the $U(1)$ transformation

$$\Psi \to e^{-i\alpha}\Psi \quad \overline{\Psi} \to e^{i\alpha}\overline{\Psi}$$

The corresponding Noether current is

$$j^\mu = \overline{\Psi}\gamma^\mu \Psi = \chi^\dagger \overline{\sigma}^\mu \chi - \xi^\dagger \overline{\sigma}^\mu \xi$$

The equation of motion is

$$(-i\gamma^\mu \partial_\mu + m)\Psi = 0$$

Charge conjugation

Charge conjugation simply exchanges ξ and χ. We can define a unitary charge conjugation operator C that implements this,

$$C^{-1}\xi_a(x)C = \chi_a(x)$$

$$C^{-1}\chi_a(x)C = \xi_a(x)$$

We then have $C^{-1}\mathcal{L}(x)C = \mathcal{L}(x)$.
Introduce the charge conjugation matrix

$$\mathcal{C} \equiv \begin{bmatrix} \varepsilon_{ab} & \\ & \varepsilon^{\dot{a}\dot{b}} \end{bmatrix}$$

Take the transpose of Ψ

$$\overline{\Psi}^T = \begin{bmatrix} \xi^a \\ \chi^\dagger_{\dot{a}} \end{bmatrix}$$

Define the charge conjugate of Ψ,

$$\Psi^C \equiv \mathcal{C}\overline{\Psi}^T = \begin{bmatrix} \xi_a \\ \chi^{\dagger\dot{a}} \end{bmatrix}$$

We therefore have

$$C^{-1}\Psi(x)C = \Psi^C(x)$$

for a Dirac field.
The charge conjugation matrix has a number of useful properties.

$$\mathcal{C}^T = \mathcal{C}^\dagger = \mathcal{C}^{-1} = -\mathcal{C}$$

$$\mathcal{C}^{-1}\gamma^\mu\mathcal{C} = -(\gamma^\mu)^T$$

Now let us return to the Majorana field. It is obvious that a Majorana field is its own charge conjugate, that is, $\Psi^C = \Psi$. This condition is analogous to the condition $\phi^\dagger = \phi$ that is satisfied by a real scalar field. A Dirac field, with its $U(1)$ symmetry, is analogous to a complex scalar field, while a Majorana field, which has no $U(1)$ symmetry, is analogous to a real scalar field. For a Majorana field, we have $\overline{\Psi} = \Psi^T\mathcal{C}$. Then the Lagrangian can be written as

$$\mathcal{L} = \frac{i}{2}\Psi^T\mathcal{C}\gamma^\mu\partial_\mu\Psi - \frac{1}{2}m\Psi^T\mathcal{C}\Psi$$

Projection matrix

We can also recover the Weyl components of a Dirac or Majorana field by means of a suitable projection matrix. Define

$$\gamma^5 \equiv \begin{bmatrix} -\delta_a{}^c & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{bmatrix}$$

Then we can define left and right projection matrices

$$P_L \equiv \frac{1}{2}(1-\gamma^5) = \begin{bmatrix} \delta_a{}^c & 0 \\ 0 & 0 \end{bmatrix}$$

$$P_R \equiv \frac{1}{2}(1+\gamma^5) = \begin{bmatrix} 0 & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{bmatrix}$$

The matrix γ^5 can also be expressed as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$$

The γ^5 has the following properties,

$$(\gamma^5)^\dagger = \gamma^5$$
$$(\gamma^5)^2 = 1$$
$$\{\gamma^5, \gamma^\mu\} = 0$$

The behavior of a Dirac field under a Lorentz transformation

Define

$$S^{\mu\nu} \equiv \begin{bmatrix} (S_L^{\mu\nu})_a{}^b & 0 \\ 0 & -(S_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}} \end{bmatrix} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$$

Numerically, we have

$$S^{i0} = \frac{i}{2}\begin{bmatrix} \sigma^i & \\ & -\sigma^i \end{bmatrix} \quad S^{ij} = \frac{1}{2}\epsilon_{ijk}\begin{bmatrix} \sigma^k & \\ & \sigma^k \end{bmatrix}$$

Then, for either a Dirac or Majorana field Ψ, we can write

$$U(\Lambda)^{-1}\Psi(x)U(\Lambda) = D(\Lambda)\Psi(\Lambda^{-1}x)$$

where, for an infinitesimal transformation,

$$D(1+\delta\omega) = I + \frac{i}{2}\delta\omega_{\mu\nu}S^{\mu\nu}$$

So the $D(\Lambda)$ can be written as

$$D = \begin{bmatrix} L_a{}^b & 0 \\ 0 & -R^{\dot{a}}{}_{\dot{b}} \end{bmatrix}$$

We can verify that

$$U(\Lambda)^{-1}\overline{\Psi}(x)U(\Lambda) = \overline{\Psi}(\Lambda^{-1}x)[D(\Lambda)]^{-1}$$

From the identity

$$\sigma^\mu_{a\dot{a}} = L_a{}^b R_{\dot{a}}{}^{\dot{b}} \Lambda^\mu{}_\nu \sigma^\nu_{b\dot{b}} \quad \overline{\sigma}^{\mu\dot{a}a} = L^a{}_b R^{\dot{a}}{}_{\dot{b}} \Lambda^\mu{}_\nu \overline{\sigma}^{\nu\dot{b}b}$$

We have

$$L^a{}_b R^{\dot{a}}{}_{\dot{b}} \sigma^\mu_{a\dot{a}} = \Lambda^\mu{}_\nu \sigma^\nu_{b\dot{b}} \quad L_a{}^b R_{\dot{a}}{}^{\dot{b}} \overline{\sigma}^{\mu\dot{a}a} = \Lambda^\mu{}_\nu \overline{\sigma}^{\nu\dot{b}b}$$

So we can verify that

$$D^{-1}\gamma^\mu D = \Lambda^\mu{}_\nu \gamma^\nu$$

Recall the commutation relation of $M^{\mu\nu}$ and P^μ, we have

$$[\gamma^\mu, S^{\rho\sigma}] = i(\eta^{\mu\sigma}\gamma^\mu - \eta^{\mu\rho}\gamma^\sigma)$$

We also have

$$[\gamma^5, S^{\mu\nu}] = 0$$

which means that

$$D^{-1}\gamma^5 D = \gamma^5$$

24.5 Canonical quantization of Dirac field

Canonical momentum and Hamiltonian

$$\Pi \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi)} = i\overline{\Psi}\gamma^0 = i\Psi^\dagger \quad (\overline{\Psi} = -i\Pi\gamma^0 \quad \Psi^\dagger = -i\Pi)$$

$$\mathcal{H} = -\Pi(\vec{\alpha}\cdot\vec{\nabla} + i\beta m)\Psi \quad (\alpha_i = \gamma^0\gamma^i \quad \beta = \gamma_0)$$

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

Momentum and angular momentum

$$T^{\mu\nu} = i\overline{\Psi}\left[\eta^{\mu\nu}\gamma^\rho\partial_\rho - \gamma^\mu\partial^\nu\right]\Psi - m\eta^{\mu\nu}\overline{\Psi}\Psi$$

$$P^0 = H \quad P^i = -\int \Pi\nabla^i\Phi\, d^3x$$

$$J_i = -\epsilon_{ijk}\int \Pi(x^j\nabla^k + \frac{i}{2}S^{jk})\Psi\, d^3x$$

Define $\Sigma_i \equiv \frac{1}{2}\epsilon_{ijk}S^{jk}$, so

$$\Sigma_i = \frac{1}{2}\begin{bmatrix}\sigma^i & \\ & \sigma^i\end{bmatrix}$$

The above equation can be written as

$$\vec{J} = -\int \Pi(\vec{x}\times\vec{\nabla} + i\vec{\Sigma})\Psi\, d^3x$$

Canonical quantization

$$\{\Psi_a(\boldsymbol{x},t), \Psi_b(\boldsymbol{x},t)\} = 0$$

$$\{\Psi_a(\boldsymbol{x},t), \Pi^b(\boldsymbol{y},t)\} = i\delta_a^b\delta(\boldsymbol{x}-\boldsymbol{y})$$

$$\{\Psi_a(\boldsymbol{x},t), \Psi^{\dagger b}(\boldsymbol{y},t)\} = \delta_a^b\delta(\boldsymbol{x}-\boldsymbol{y})$$

Solution of Dirac equation

$$\Psi(x) = \sum_{s=\pm}\int \widetilde{dp}\left[b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ipx} + d_s^\dagger(\boldsymbol{p})v_s(\boldsymbol{p})e^{-ipx}\right]$$

Here, we introduce the Feynman slash: given any four-vector a^μ, we define

$$\not{a} \equiv a_\mu\gamma^\mu$$

The Dirac equation implies that

$$(\not{p} + m)u(\boldsymbol{p}) = 0$$

$$(-\not{p} + m)v(\boldsymbol{p}) = 0$$

Each of these equations has two solutions, which we label via $s = +$ and $s = -$. For $m \neq 0$, we can go to the rest frame, $\boldsymbol{p} = 0$. We will then distinguish the two solutions by the eigenvalue of the spin matrix Σ_3. Specifically, we will require

$$\Sigma_3 u_\pm(\mathbf{0}) = \pm\frac{1}{2}u_\pm(\mathbf{0})$$

$$\Sigma_3 v_\pm(\mathbf{0}) = \mp\frac{1}{2}v_\pm(\mathbf{0})$$

The solutions are

$$u_+(\mathbf{0}) = \sqrt{m}\begin{pmatrix}1\\0\\1\\0\end{pmatrix} \qquad u_-(\mathbf{0}) = \sqrt{m}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$$

$$v_+(\mathbf{0}) = \sqrt{m}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix} \qquad v_-(\mathbf{0}) = \sqrt{m}\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}$$

For later use we also compute the barred spinors

$$\overline{u}_s(\boldsymbol{p}) \equiv u_s^\dagger(\boldsymbol{p})\beta \quad \overline{v}_s(\boldsymbol{p}) \equiv v_s^\dagger(\boldsymbol{p})\beta \quad (\beta = \gamma^0)$$

We get

$$\overline{u}_+(\mathbf{0}) = \sqrt{m}(1,0,1,0)$$
$$\overline{u}_-(\mathbf{0}) = \sqrt{m}(0,1,0,1)$$
$$\overline{v}_+(\mathbf{0}) = \sqrt{m}(0,-1,0,1)$$
$$\overline{v}_-(\mathbf{0}) = \sqrt{m}(1,0,-1,0)$$

We can now find the spinors corresponding to an arbitrary three-momentum $\boldsymbol{p}$ by applying to $u_s(\mathbf{0})$ and $v_s(\mathbf{0})$ the matrix $D(\Lambda)$ that corresponds to an appropriate boost. This is given by

$$D(\Lambda) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$

where $\hat{\boldsymbol{p}}$ is a unit vector in the $\boldsymbol{p}$ direction, $D^i = \frac{i}{4}[\gamma^i,\gamma^0] = \frac{i}{2}\gamma^i\gamma^0$ is the boost matrix, and η is the rapidity. Thus we have

$$u_s(\boldsymbol{p}) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})u_s(\mathbf{0})$$

$$v_s(\boldsymbol{p}) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})v_s(\mathbf{0})$$

We also have

$$\overline{u}_s(\boldsymbol{p}) = \overline{u}_s(\mathbf{0})\exp(-i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$
$$\overline{v}_s(\boldsymbol{p}) = \overline{v}_s(\mathbf{0})\exp(-i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$

This follows from $\overline{K}^j = K^j$, where, for any general combination of gamma matrices,

$$\overline{A} \equiv \beta A^\dagger\beta$$

In particular, it turns out that

$$\gamma^\mu \quad S^{\mu\nu} \quad i\gamma^5 \quad \gamma^\mu\gamma^5 \quad i\gamma^5 S^{\mu\nu}$$

all satisfy $\overline{A} = A$.
The barred spinors satisfy the equations

$$\overline{u}_s(\boldsymbol{p})(\not{p} + m) = 0 \quad \overline{v}_s(\boldsymbol{p})(-\not{p} + m) = 0$$

Proposition 24.3

1.
$$\overline{u}_{s'}(\boldsymbol{p})u_s(\boldsymbol{p}) = 2m\delta_{ss'}$$
$$\overline{v}_{s'}(\boldsymbol{p})v_s(\boldsymbol{p}) = -2m\delta_{ss'}$$
$$\overline{u}_{s'}(\boldsymbol{p})v_s(\boldsymbol{p}) = 0$$
$$\overline{v}_{s'}(\boldsymbol{p})u_s(\boldsymbol{p}) = 0$$

2.
$$2m\overline{u}_{s'}(\boldsymbol{p}')\gamma^\mu u_s(\boldsymbol{p}) = \overline{u}_{s'}(\boldsymbol{p}')[(p'+p)^\mu - 2iS^{\mu\nu}(p'-p)_\nu]u_s(\boldsymbol{p})$$
$$-2m\overline{v}_{s'}(\boldsymbol{p}')\gamma^\mu v_s(\boldsymbol{p}) = \overline{v}_{s'}(\boldsymbol{p}')[(p'+p)^\mu - 2iS^{\mu\nu}(p'-p)_\nu]v_s(\boldsymbol{p})$$

3.
$$\overline{u}_{s'}(\boldsymbol{p})\gamma^\mu u_s(\boldsymbol{p}) = 2p^\mu\delta_{ss'}$$
$$\overline{v}_{s'}(\boldsymbol{p})\gamma^\mu v_s(\boldsymbol{p}) = 2p^\mu\delta_{ss'}$$
$$\overline{u}_{s'}(\boldsymbol{p})\gamma^0 v_s(-\boldsymbol{p}) = 0$$
$$\overline{v}_{s'}(\boldsymbol{p})\gamma^0 u_s(-\boldsymbol{p}) = 0$$

4.
$$\sum_{s=\pm} u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) = -\not{p} + m$$
$$\sum_{s=\pm} v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) = -\not{p} - m$$

5.
$$u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) = \frac{1}{2}(1 - s\gamma^5\not{z})(-\not{p} + m)$$
$$v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) = \frac{1}{2}(1 - s\gamma^5\not{z})(-\not{p} - m)$$

Here, z^μ is the boost of the vector $(0,0,0,1)$ from the frame with $\boldsymbol{p}' = 0$ to the frame with $\boldsymbol{p}' = \boldsymbol{p}$.

The proof can be found in chapter 38 from *Quantum field theory (Mark Srednicki)*

It is interesting to consider the extreme relativistic limit of this formula. Let us take the three-momentum to be in the z direction, so that it is parallel to the spin-quantization axis. The component of the spin in the direction of the three-momentum is called the helicity. A fermion with helicity $+1/2$ is said to be right-handed, and a fermion with helicity $-1/2$ is said to be left-handed. Note that the direction of spin of v_+ is in $-z$.
For rapidity η, we have

$$\frac{p^\mu}{m} = (\cosh\eta, 0, 0, \sinh\eta) \quad z^\mu = (\sinh\eta, 0, 0, \cosh\eta)$$

In the limit of large η,

$$z^\mu = \frac{p^\mu}{m} + \mathcal{O}(e^{-\eta})$$

Then the equations 3.3.5 can be written as

$$u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) \to \frac{1}{2}(1 + s\gamma^5)(-\not{p})$$

$$v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) \to \frac{1}{2}(1 - s\gamma^5)(-\not{p})$$

So, we can see the spinor corresponding to a right-handed fermion is $u_+(\boldsymbol{p})$ for a b-type particle and $v_-(\boldsymbol{p})$ for a d-type particle. The spinor corresponding to a left-handed fermion is $u_-(\boldsymbol{p})$ for a b-type particle and $v_+(\boldsymbol{p})$ for a d-type particle.

Note that $\beta u_s(\boldsymbol{0}) = +u_s(\boldsymbol{0})$ and $\beta v_s(\boldsymbol{0}) = -v_s(\boldsymbol{0})$. Also, $\beta D^j = -D^j\beta$. We then have

$$u_s(-\boldsymbol{p}) = \beta u_s(\boldsymbol{p}) \quad v_s(-\boldsymbol{p}) = -\beta v_s(\boldsymbol{p})$$

For charge conjugation matrix, note that that $\mathcal{C}\overline{u}_s^T(\boldsymbol{0}) = v_s(\boldsymbol{0})$, $\mathcal{C}\overline{v}_s^T(\boldsymbol{0}) = v_s(\boldsymbol{0})$, and $\mathcal{C}^{-1}D^j\mathcal{C} = -(D^j)^T$, we have

$$\mathcal{C}\overline{u}_s^T(\boldsymbol{p}) = v_s(\boldsymbol{p}) \quad \mathcal{C}\overline{v}_s^T(\boldsymbol{p}) = u_s(\boldsymbol{p})$$

Taking the complex conjugate of the above equation, we have

$$u_s^*(\boldsymbol{p}) = \mathcal{C}\beta v_s(\boldsymbol{p}) \quad v_s^*(\boldsymbol{p}) = \mathcal{C}\beta u_s(\boldsymbol{p})$$

Note that $\gamma^5 u_s(\boldsymbol{0}) = +sv_{-s}(\boldsymbol{0})$ and $\gamma^5 v_s(\boldsymbol{0}) = -su_{-s}(\boldsymbol{0})$, and that $\gamma^5 D^j = D^j\gamma^5$, we have

$$\gamma^5 u_s(\boldsymbol{p}) = +sv_{-s}(\boldsymbol{p}) \quad \gamma^5 v_s(\boldsymbol{p}) = -su_{-s}(\boldsymbol{p})$$

Combine the above equations, we can derive

$$u_{-s}^*(\boldsymbol{p}) = -s\mathcal{C}\gamma^5 u_s(\boldsymbol{p}) \quad v_{-s}^*(\boldsymbol{p}) = -s\mathcal{C}\gamma^5 v_s(\boldsymbol{p})$$

Fourier expansion

$$\Psi(x) = \sum_{s=\pm}\int \widetilde{dp}\left[b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ipx} + d_s^\dagger(\boldsymbol{p})v_s(\boldsymbol{p})e^{-ipx}\right]$$

$$\Pi(x) = i\sum_{s=\pm}\int \widetilde{dp}\left[b_s^\dagger(\boldsymbol{p})u_s^\dagger(\boldsymbol{p})e^{-ipx} + d_s(\boldsymbol{p})v_s^\dagger(\boldsymbol{p})e^{+ipx}\right]$$

$$b_s(\boldsymbol{p}) = \int d^3x e^{-ipx} \overline{u}_s(\boldsymbol{p}) \gamma^0 \Psi(x)$$

$$b_s^\dagger(\boldsymbol{p}) = \int d^3x e^{ipx} \overline{\Psi}(x) \gamma^0 u_s(\boldsymbol{p})$$

$$d_s(\boldsymbol{p}) = \int d^3x e^{-ipx} \overline{\Psi}(x) \gamma^0 v_s(\boldsymbol{p})$$

$$d_s^\dagger(\boldsymbol{p}) = \int d^3x e^{ipx} \overline{v}_s(\boldsymbol{p}) \gamma^0 \Psi(x)$$

We can get the anticommutation relation in terms of $b, b^\dagger, d, d^\dagger$. The only non-vanishing terms are

$$\{b_s(\boldsymbol{p}), b_{s'}^\dagger(\boldsymbol{p}')\} = (2\pi)^3 \delta^3(\boldsymbol{p} - \boldsymbol{p}') 2\omega \delta_{ss'}$$
$$\{d_s(\boldsymbol{p}), d_{s'}^\dagger(\boldsymbol{p}')\} = (2\pi)^3 \delta^3(\boldsymbol{p} - \boldsymbol{p}') 2\omega \delta_{ss'}$$

Operator represented by $b, b^\dagger, d, d^\dagger$

Define

$$N^+(\boldsymbol{p}, s) = b_s^\dagger(\boldsymbol{p}) b_s(\boldsymbol{p}) \quad N^-(\boldsymbol{p}, s) = d_s^\dagger(\boldsymbol{p}) d_s(\boldsymbol{p})$$

So we can derive

$$H = \sum_{s=\pm} \int \widetilde{dp}\, \omega \left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right] - 4\mathcal{E}_0 V$$

$$P^i = \sum_{s=\pm} \int \widetilde{dp}\, p^i \left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right]$$

$$S_3 = \sum_{s=\pm} \int \widetilde{dp}\, \frac{s}{2} \left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right]$$

$$Q = \sum_{s=\pm} \int \widetilde{dp}\, \left[N^+(\boldsymbol{p}, s) - N^-(\boldsymbol{p}, s)\right]$$

Causality

Firstly, we derive the anticommutation relation for field operators at any space-time.

$$\{\overline{\Psi}_a(x), \Psi_b(y)\} = (i\partial\!\!\!/_x + m)_{ab} i\Delta(x - y)$$

Here,

$$i\Delta(x - y) = \int \widetilde{dp}\, [e^{ip(x-y)} - e^{-ip(x-y)}]$$

For $(x - y)^2 > 0$ the anti-commutators vanish, because $\Delta(x - y)$ also vanishes. We then can verify that

$$[\overline{\Psi}_a(x)\Psi_b(x), \overline{\Psi}_c(y)\Psi_d(y)] = 0$$

for $(x - y)^2 > 0$.

In this way the microscopic causality is satisfied for the physical observables, such as the charge density or the momentum density.

The Dirac propagator

$$\langle 0|\Psi_a(x)\overline{\Psi}_b(y)|0\rangle = (i\partial\!\!\!/_x + m)_{ab}\int \widetilde{dp}\, e^{ip(x-y)}$$

$$\langle 0|\overline{\Psi}_b(y)\Psi_a(x)|0\rangle = -(i\partial\!\!\!/_x + m)_{ab}\int \widetilde{dp}\, e^{ip(y-x)}$$

Define retarded green function as

$$S_R(x-y)_{ab} \equiv \theta(x^0-y^0)\langle 0|\{\Psi_a(x)\overline{\Psi}_b(y)\}|0\rangle$$

It is easy to verify that

$$S_R(x-y) = (i\partial\!\!\!/_x + m)D_R(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i(p\!\!\!/ - m)}{p^2+m^2}e^{ip(x-y)}$$

and

$$(i\partial\!\!\!/_x - m)S_R(x-y) = i\delta(x-y)\cdot \mathbf{1}_{4\times 4}$$

Now, we define the time ordered product for fermion fields

$$T\eta(x)\eta(y) \equiv \theta(x^0-y^0)\eta(x)\eta(y) - \theta(y^0-x^0)\eta(y)\eta(x)$$

So,

$$S_F(x-y) \equiv \langle 0|T\Psi(x)\overline{\Psi}(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{i(p\!\!\!/ - m)}{p^2+m^2-i\epsilon}e^{ip(x-y)}$$

It is easy to verify that

$$\langle 0|T\overline{\Psi}_a(x)\Psi_b(y)|0\rangle = -\langle 0|T\Psi_b(y)\overline{\Psi}_a(x)|0\rangle = -S_F(y-x)_{ba}$$

24.6 Parity,time reversal and charge conjugation

Parity

We assume that

$$P^{-1}b_s(\boldsymbol{p})P = \eta_b b_s(-\boldsymbol{p})$$
$$P^{-1}d_s(\boldsymbol{p})P = \eta_d d_s(-\boldsymbol{p})$$

Here, η_b and η_d is the phase factor. Then, we can verify that

$$P^{-1}\boldsymbol{P}P = -\boldsymbol{P} \quad P^{-1}\boldsymbol{S}P = \boldsymbol{S}$$

So a parity transformation reverse the three momentum while leaving the spin direction unchanged. Further more, we have

$$P^{-1}\Psi(x)P = \sum_{s=\pm}\int \widetilde{dp}\ \left[\eta_b b_s(\boldsymbol{p})\beta u_s(\boldsymbol{p})e^{ip\mathcal{P}x} - \eta_d^* d_s^\dagger(\boldsymbol{p})\beta v_s(\boldsymbol{p})e^{-ip\mathcal{P}x}\right]$$

where,$\mathcal{P}^\mu{}_\nu = \mathrm{diag}(1,-1,-1,-1)$. So, we should demand that $\eta_b = -\eta_d^*$. And we can get

$$P^{-1}\Psi(x)P = \eta_b\beta\Psi(\mathcal{P}x) \quad P^{-1}\overline{\Psi}(x)P = \eta_b^*\overline{\Psi}(\mathcal{P}x)\beta$$

Generally, we have

$$P^{-1}(\overline{\Psi}A\Psi)P = \overline{\Psi}(\beta A\beta)\Psi$$

Time reversal

Note in quantum mechanics, we have shown that time reversal operator is antiunitary. Firstly, we assume that

$$T^{-1}b_s(\boldsymbol{p})T = \zeta_{b,s}b_{-s}(-\boldsymbol{p})$$

$$T^{-1}d_s(\boldsymbol{p})T = \zeta_{d,s}d_{-s}(-\boldsymbol{p})$$

Then we can verify that

$$T^{-1}\boldsymbol{P}T = -\boldsymbol{P} \quad T^{-1}\boldsymbol{S}T = -\boldsymbol{S}$$

So a parity transformation reverse the three momentum and the spin direction. Further more, we have

$$T^{-1}\Psi(x)T = \sum_{s=\pm}\int \widetilde{dp}\ -s\mathcal{C}\gamma^5\left[\zeta_{b,-s}b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ip\mathcal{T}x} + \zeta^*_{d,-s}d^\dagger_s(\boldsymbol{p})\beta v_s(\boldsymbol{p})e^{-ip\mathcal{T}x}\right]$$

where,$\mathcal{T}^\mu_{\ \nu} = \mathrm{diag}(-1,1,1,1)$. So, we should demand that $\zeta_{s,b} = s\zeta$ and $\zeta_{s,d} = s\zeta^*$. And we can get

$$T^{-1}\Psi(x)T = \zeta\mathcal{C}\gamma^5\Psi(\mathcal{T}x) \quad T^{-1}\overline{\Psi}(x)T = \zeta^*\overline{\Psi}(\mathcal{T}x)\gamma^5\mathcal{C}^{-1}$$

Generally, we have

$$T^{-1}(\overline{\Psi}A\Psi)T = \overline{\Psi}(\gamma^5\mathcal{C}^{-1}A\mathcal{C}\gamma^5)\Psi$$

Charge conjugation

We have already shown that

$$C^{-1}\Psi(x)C = \mathcal{C}\overline{\Psi}^T(x) \quad C^{-1}\overline{\Psi}(x)C = \Psi^T(x)\mathcal{C}$$

Generally, we have

$$C^{-1}(\overline{\Psi}A\Psi)C = \overline{\Psi}(\mathcal{C}^{-1}A^T\mathcal{C})\Psi$$

Summary

The transformation properties of the various fermion bilinears under C,P and T are summarized in the table below. Here we use the shorthand $(-1)^\mu \equiv 1$ for $\mu = 0$ and $(-1)^\mu \equiv -1$ for $\mu = 1,2,3$.

	$\bar\psi\psi$	$i\bar\psi\gamma^5\psi$	$\bar\psi\gamma^\mu\psi$	$\bar\psi\gamma^\mu\gamma^5\psi$	$\bar\psi\sigma^{\mu\nu}\psi$	∂_μ
P	$+1$	-1	$(-1)^\mu$	$-(-1)^\mu$	$(-1)^\mu(-1)^\nu$	$(-1)^\mu$
T	$+1$	-1	$(-1)^\mu$	$(-1)^\mu$	$-(-1)^\mu(-1)^\nu$	$-(-1)^\mu$
C	$+1$	$+1$	-1	$+1$	-1	$+1$
CPT	$+1$	$+1$	-1	-1	$+1$	-1

Figure 24.1: Transformation of fermion bilinears under CPT

Theorem 24.2 CPT theorem

Any hermitian combination of any set of fields (scalar, vector, Dirac, Majorana) and their derivatives that is a Lorentz scalar (and so carries no indices) is even under CPT. Since the lagrangian must be formed out of such combinations, we have $\mathcal{L}(x) \to \mathcal{L}(-x)$ under CPT, and so the action $S = \int d^4x\mathcal{L}$ is invariant. ♣

24.7 Perturbation theory for canonical quantization

We use Yukawa theory as an example.

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}M_0^2\phi^2 + i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m_0\overline{\Psi}\Psi - g_0\overline{\Psi}\Psi\phi$$

$$H = H_0 + H_{int} \quad H_{int} = \int d^3x\, g_0\overline{\Psi}\Psi\phi$$

Similar to ϕ^4 theory, the perturbation expansion of correlation functions is

$$\langle\Omega|T\{\Psi(x)\overline{\Psi}(y)\phi(z)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\Psi_I(x)\overline{\Psi}_I(y)\phi_I(z)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}$$

Before we state the Wick's theorem, we must note the following conventions:

1. The time-ordered product picks up one minus sign for each interchange of operators that is necessary to put the fields in time order.
2. The normal-ordered product picks up one minus sign for each interchange of operators that is necessary to put the fields in normal order.
3. Define contractions under the normal-ordering symbol to include minus signs for operator interchanges.

With these conventions, Wick's theorem takes the same form as before:

$$T\left\{\Psi_I(x_1)\overline{\Psi}_I(x_2)\Psi_I(x_3)\cdots\right\} = N\left\{\Psi_I(x_1)\overline{\Psi}_I(x_2)\Psi_I(x_3)\cdots + \text{ all possible contractions }\right\}$$

Example:

$$\begin{aligned}\langle 0|T\left\{\Psi_{Ia}(x_1)\overline{\Psi}_{Ib}(x_2)\Psi_{Ic}(x_3)\overline{\Psi}_{Id}(x_4)\right\}|0\rangle &= S_F(x_1-x_2)_{ab}S_F(x_3-x_4)_{cd} \\ &- S_F(x_1-x_4)_{ad}S_F(x_3-x_2)_{cb}\end{aligned}$$

Then we can derive the Feynman rule for Yukawa theory. Expand

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\phi_I(z)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle$$

to the first order of g_0

$$\begin{aligned}
& \langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\phi_I(z)(-ig_0)\int dw^4\overline{\Psi}_I(w)\Psi_I(w)\phi_I(w)\right\}|0\rangle \\
= & -(-ig_0)S_F(x-y)_{ab}\int d^4wD_F(z-w)\mathrm{Tr}[S_F(w-w)] \\
+ & (-ig_0)\int d^4w[S_F(x-w)S_F(w-y)]_{ab}D_F(w-z)
\end{aligned}$$

It can be represented by the so called Feynman diagram.

Figure 24.2: Feynman diagram representation of perturbation expansion

The Feynman rules for Yukama theory are:

1. For each Fermion propagator from y to x, $P = S_F(x-y)$
2. For each scalar propagator, $P = D_F(x-y)$
3. For each vertex, $V = (-ig_0)\int d^4w$
4. For each external point, $E = 1$
5. Divided by the symmetry factor

24.8 Path integral quantization

24.8.1 Grassmann numbers

Formal definition

Grassmann numbers are individual elements or points of the exterior algebra generated by a set of n Grassmann variables or Grassmann directions or supercharges $\{\theta_i\}$ with n possibly being infinite.
The Grassman variables are the basis vectors of a vector space (of dimension n) They form an algebra over a field, with the field usually being taken to be the complex numbers, although one could contemplate other fields, such as the reals. The algebra is a unital algebra, and the generators are anti-commuting:

$$\theta_i\theta_j = -\theta_j\theta_i$$

Since the θ_i form a vector space over the complex numbers, it is trivial that they commute with the complex numbers; this is by definition. That is, for complex x, one has

$$\theta_i x = x\theta_i$$

The squares of the generators vanish:

$$\theta_i\theta_i = -\theta_i\theta_i$$

In other words, a Grassmann variable is a non-zero square-root of zero.
Let V denote this n-dimensional vector space of Grassman variables. Note that it is independent of the choice of basis. The corresponding exterior algebra is defined as

$$\Lambda = \mathbb{C} \oplus V \oplus (V \wedge V) \oplus (V \wedge V \wedge V) \oplus \cdots$$

where $\wedge$ is the exterior product and $\oplus$ is the direct sum. The individual elements of this algebra are then called Grassman numbers. It is standard to completely omit the wedge symbol $\wedge$ when writing a Grassman number; it is used here only to clearly illustrate how the exterior algebra is built up out of the Grassman variables. Thus, a completely general Grassman number can be written as

$$z = \sum_{k=0}^{\infty} \sum_{i_1, i_2, \cdots, i_k} c_{i_1 i_2 \cdots i_k} \theta_{i_1} \theta_{i_2} \cdots \theta_{i_k}$$

where the c's are complex numbers, or, equivalently, $c_{i_1 i_2 \cdots i_k}$ is a complex-valued, completely antisymmetric tensor of rank k. Again, the θ_i can be clearly seen here to be playing the role of a basis vector of a vector space.
Observe that the Grassmann algebra generated by n linearly independent Grassmann variables has dimension $2n$; this follows from the binomial theorem applied to the above sum, and the fact that the $n+1$-fold product of variables must vanish, by the anti-commutation relations, above. In other words, for n variables, the sum terminates

$$\Lambda = \mathbb{C} \oplus \Lambda^1 V \oplus \Lambda^2 V \oplus \cdots \oplus \Lambda^n V$$

where $\Lambda^k V$ is the k-fold alternating product. The dimension of $\Lambda^k V$ is given by n choose k, the binomial coefficient. The special case of $n = 1$ is called a dual number, and was introduced by William Clifford in 1873.

Integral over Grassmann number

Single Variable:

$$\int d\theta (A + B\theta) \equiv B$$

Multi-variable:

$$\int d\theta d\eta \; \eta\theta \equiv 1$$

Complex Grassmann number:

$$(\theta\eta)^* \equiv \eta^*\theta^* = -\theta^*\eta^*$$

$$\int d\theta^* d\theta \; \theta\theta^* \equiv 1$$

Gaussian integral over a complex Grassmann number:

$$\int d\theta^* d\theta e^{-\theta^* b\theta} = b$$

$$\int d\theta^* d\theta\ \theta\theta^* e^{-\theta^* b\theta} = 1$$

Unitary transformation:
If $\theta_i' = U_{ij}\theta_j$ and U is unitary matrix, then we can derive

$$\prod_i \theta_i' = (\det U)(\prod_i \theta_i)$$

In a general integral

$$(\prod_i \int d\theta_i^* d\theta_i) f(\theta)$$

the only term of $f(\theta)$ that survives has exactly one factor of each θ_i and θ_i^*; it is proportional to $(\prod_i \theta_i)(\prod_i \theta_i^*)$. If we replace θ by $U\theta$, this term acquires a factor of $\det U \det U^* = 1$, so the integral is unchanged under the unitary transformation.

Gaussian integral over multiple complex Grassmann numbers:

$$(\prod_i \int d\theta_i^* d\theta_i) e^{-\theta_i^* B_{ij}\theta_j} = \det B$$

$$(\prod_i \int d\theta_i^* d\theta_i) \theta_k \theta_l^* e^{-\theta_i^* B_{ij}\theta_j} = B_{kl}^{-1} \det B$$

24.8.2 Path integral formulation for free Dirac field

A Grassmann field is a function of space-time whose values are Grassmann numbers. The classical Dirac field being used to evaluate path integral is a Grassmann field. The correlation function is given by

$$\langle\Omega|T\Psi_H(x_1)\overline{\Psi}_H(x_2)|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \exp\left[i\int_{-T}^{T} d^4x \overline{\Psi}(i\partial\!\!\!/ - m)\Psi\right] \Psi(x_1)\overline{\Psi}(x_2)}{\int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \exp\left[i\int_{-T}^{T} d^4x \overline{\Psi}(i\partial\!\!\!/ - m)\Psi\right]}$$

The generating function is

$$Z[\overline{\eta},\eta] = \int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \exp\left[i\int d^4x \overline{\Psi}(i\partial\!\!\!/ - m)\Psi + \overline{\eta}\Psi + \overline{\Psi}\eta\right]$$

Here, $\eta(x)$ is a Grassmann-valued source field. Define

$$\Psi'(x) \equiv \Psi(x) - i\int d^4y S_F(x-y)\eta(y)$$

Then we can derive that

$$\overline{\Psi}'(x) \equiv \overline{\Psi}(x) - i\int d^4y \overline{\eta}(y) S_F(y-x)$$

Recall that

$$(i\partial\!\!\!/_x - m)S_F(x-y) = i\delta(x-y)$$

and

$$S_F(y-x)(i\partial\!\!\!/_x + m) = -i\delta(x-y)$$

we can derive that

$$\int d^4x\overline{\Psi}(i\partial\!\!\!/ - m)\Psi + \overline{\eta}\Psi + \overline{\Psi}\eta = \int d^4x\overline{\Psi}'(i\partial\!\!\!/ - m)\Psi' + i\int d^4xd^4y\overline{\eta}(x)S_F(x-y)\eta(y)$$

After integration, we have

$$Z[\overline{\eta},\eta] = Z[0]\exp\left[-\int d^4xd^4y\ \overline{\eta}(x)S_F(x-y)\eta(y)\right]$$

If we adopt the convention that

$$\frac{d}{d\eta}\theta\eta = -\frac{d}{d\eta}\eta\theta = -\theta$$

the two point correlation functions are

$$\langle 0|T\Psi_H(x_1)\overline{\Psi}_H(x_2)|0\rangle = Z[0]^{-1}\left(-i\frac{\delta}{\delta\overline{\eta}(x_1)}\right)\left(i\frac{\delta}{\delta\eta(x_2)}\right)Z[\overline{\eta},\eta]|_{\overline{\eta},\eta=0} = S_F(x_1-x_2)$$

24.8.3 Perturbation theory for path integral quantization

We use Yukawa theory as an example.

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}M_0^2\phi^2 + i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m_0\overline{\Psi}\Psi - g_0\overline{\Psi}\Psi\phi \\
\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_1 \quad \mathcal{L}_1 = -g_0\overline{\Psi}\Psi\phi \\
Z[J] &= \int \mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+J\phi+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&= e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\int \mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4y[\mathcal{L}_0+J\phi+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\exp[-\int d^4yd^4z\frac{1}{2}J(y)D_F(y-z)J(z) + \overline{\eta}(y)S_F(y-z)\eta(z)] \\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[-ig_0\int d^4x(\frac{1}{i}\frac{\delta}{\delta J(x)}\cdot -\frac{1}{i}\frac{\delta}{\delta\eta(x)}\cdot\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)})]^V \\
&\times \sum_{P_1=0}^{\infty}\frac{1}{P_1!}[-\frac{1}{2}\int d^4y_1d^4z_1J(y_1)D_F(y_1-z_1)J(z_1)]^{P_1} \\
&\times \sum_{P_2=0}^{\infty}\frac{1}{P_2!}[-\int d^4y_2d^4z_2\overline{\eta}(y_2)S_F(y_2-z_2)\eta(z_2)]^{P_2}
\end{aligned}$$

If we focus on a term with particular values of V, P_1 and P_2, then the number of surviving scalar sources is $E_1 = 2P_1 - V$, the number of surviving fermion sources is $E_2 = 2P_2 - 2V$. We can introduce Feynman diagrams as in the ϕ^4 theory. In these diagrams, a dashed line segment stands for a scalar propagator $D_F(x-y)$, a line with an arrow pointing from y to x for a fermion propagator $S_F(x-y)$, a filled circle at one end of a dashed line segment for a scalar source $i\int d^4xJ(x)$, a filled circle at the start of a line with an arrow for a fermion source $i\int d^4x\eta(x)$, a filled circle at the end of a line with an arrow for a anti-fermion source $i\int d^4x\overline{\eta}(x)$, a vertex joining three line segments for $-ig_0\int d^4x$.

24.9 LSZ reduction formula

Similarly to the scalar field theory, we can derive the structure of the exact propagator of Dirac fermions in interaction theory.

$$\int d^4x\, e^{-ipx}\langle\Omega|T\Psi(x)\overline{\Psi}(0)|\Omega\rangle_C = \frac{iZ_2(\not{p}-m)}{p^2+m^2-i\epsilon}+\cdots$$

We eliminate the terms contributing none isolate poles for p^2. Here, m is the exact mass of the fermions. The constant Z_2 is the probability for the quantum field to create or annihilate an exact one-particle eigenstate of H:

$$\langle\Omega|\Psi(0)|p,s\rangle = \sqrt{Z_2}u^s(p)$$

The LSZ reduction formula for fermions would take the form as following.

Theorem 24.3 LSZ reduction formula

$$\begin{aligned}
&\langle \boldsymbol{p}_1\cdots\boldsymbol{p}_n\,\overline{\boldsymbol{p}}_1\cdots\overline{\boldsymbol{p}}_{\overline{n}}\,|S|\,\boldsymbol{k}_1\cdots\boldsymbol{k}_m\,\overline{\boldsymbol{k}}_1\cdots\overline{\boldsymbol{k}}_{\overline{n}}\rangle\\
=\;&\prod_1^n\int d^4x_ie^{-ip_ix_i}\prod_1^{\overline{n}}\int d^4\overline{x}_ie^{-i\overline{p}_i\overline{x}_i}\prod_1^m\int d^4y_je^{ik_jy_j}\prod_1^{\overline{m}}d^4\overline{y}_ie^{i\overline{k}_j\overline{y}_j}\\
\times\;&[\overline{u}_{s_1}(\boldsymbol{p}_1)(\not{p}_1+m)]\cdots[\overline{u}_{s_n}(\boldsymbol{p}_n)(\not{p}_n+m)]\\
\times\;&[\overline{v}_{\overline{r}_1}(\overline{\boldsymbol{k}}_1)(\overline{\not{k}}_1-m)]\cdots[\overline{v}_{\overline{r}_{\overline{m}}}(\overline{\boldsymbol{k}}_{\overline{m}})(\overline{\not{k}}_{\overline{m}}-m)]\\
\times\;&\langle\Omega|T\{\Psi(x_1)\cdots\Psi(x_n)\overline{\Psi}(\overline{x}_1)\cdots\overline{\Psi}(\overline{x}_{\overline{n}})\overline{\Psi}(y_1)\cdots\overline{\Psi}(y_m)\Psi(\overline{y}_1)\cdots\Psi(\overline{y}_{\overline{m}})\}|\Omega\rangle\\
\times\;&[(\not{k}_1+m)u_{r_1}(\boldsymbol{k}_1)]\cdots[(\not{k}_m+m)u_{r_m}(\boldsymbol{k}_m)]\\
\times\;&[(\overline{\not{p}}_1-m)v_{\overline{s}_1}(\overline{\boldsymbol{p}}_1)]\cdots[(\overline{\not{p}}_{\overline{n}}-m)v_{\overline{s}_{\overline{n}}}(\overline{\boldsymbol{p}}_{\overline{n}})]\\
\times\;&\left(\frac{i}{\sqrt{Z_2}}\right)^{m+\overline{m}+n+\overline{n}}
\end{aligned}$$

♣

From this equation, we can see that the scattering amplitude would vanish unless $n+\overline{m}=\overline{n}+m$, which implies the conservation of Q. The term e^{ipx} will impose the condition of momentum conservation, and the the term $\not{p}\pm m$ will remove the external legs.
Finally, we list the Feymann rules of Yukawa theory in momentum space as follows:

1. For each incoming electron, draw a solid line with an arrow pointed towards the vertex, and label it with the electron's four-momentum, k_i.

2. For each outgoing electron, draw a solid line with an arrow pointed away from the vertex, and label it with the electron's four-momentum, p_i.

3. For each incoming positron, draw a solid line with an arrow pointed away from the vertex, and label it with minus the positron's four-momentum, $-\bar{k}_i$.

4. For each outgoing positron, draw a solid line with an arrow pointed towards the vertex, and label it with minus the positron's four-momentum, $-\bar{p}_i$.

5. For each incoming scalar, draw a dashed line with an arrow pointed towards the vertex, and label it with the scalar's four-momentum, q_i.

6. For each outgoing scalar, draw a dashed line with an arrow pointed away from the vertex, and label it with the scalar's four-momentum, q'_i.

7. The only allowed vertex joins two solid lines, one with an arrow pointing towards it and one with an arrow pointing away from it, and one dashed line (whose arrow can point in either direction). Using this vertex, join up all the external lines, including extra internal lines as needed. In this way, draw all possible diagrams that are topologically inequivalent.

8. Assign each internal line its own four-momentum. Think of the four-momenta as flowing along the arrows, and conserve four-momentum at each vertex. For a tree diagram, this fixes the momenta on all the internal lines.

9. The value of a diagram consists of the following factors:

   - for each incoming or outgoing scalar, 1;
   - for each incoming electron, $u_r(\boldsymbol{k})$;
   - for each outgoing electron, $\overline{u}_s(\boldsymbol{p})$;
   - for each incoming positron, $\overline{v}_{\overline{r}}(\overline{\boldsymbol{k}})$;
   - for each outgoing positron, $v_{\overline{s}}(\overline{\boldsymbol{p}})$;
   - for each vertex, $-ig_0$;
   - for each internal scalar, $\frac{-i}{p^2+M^2-i\epsilon}$;
   - for each internal fermion, $\frac{i(\not{p}-m)}{p^2+m^2-i\epsilon}$

10. Spinor indices are contracted by starting at one end of a fermion line: specifically, the end that has the arrow pointing away from the vertex. The factor associated with the external line is either $\bar{u}$ or $\bar{v}$. Go along the complete fermion line, following the arrows backwards, and write down (in order from left to right) the factors associated with the vertices and propagators that you encounter. The last factor is either a u or v. Repeat this procedure for the other fermion lines, if any.

11. The overall sign of a tree diagram is determined by drawing all contributing diagrams in a standard form: all fermion lines horizontal, with their arrows pointing from left to right, and with the left endpoints labeled in the same fixed order (from top to bottom); if the ordering of the labels on the right endpoints of the fermion lines in a given diagram is an even (odd) permutation of an arbitrarily chosen fixed ordering, then the sign of that diagram is positive (negative).

12. Each closed fermion loop contributes an extra minus sign.

13. Value of $i\mathcal{M}$ is given by a sum over the values of the contributing diagrams.

14. $\langle f|S|i\rangle = (Z_1)^{\frac{n_s}{2}}(Z_2)^{\frac{n_f}{2}} i\mathcal{M}\delta(\sum p_f - \sum p_i)$

24.10 Functional determinant

We consider a theory of a Dirac fermion Ψ with

$$\mathcal{L} = i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m\overline{\Psi}\Psi - g\overline{\Psi}\Psi\phi$$

where ϕ is again a real scalar background field. We define the path integral

$$Z(\phi) = \int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x\mathcal{L}}$$

Recall that if we have N complex Grassmann variables θ_i, then we can evaluate gaussian integrals by the general formula

$$\int d^n\theta^* d\theta \exp[-i\theta_i^* M_{ij}\theta_j] \propto \det M$$

In the case of the functional integral,the matrix M becomes

$$M_{\alpha\beta}(x,y) = (-i\partial\!\!\!/_x + m + g\phi(x))_{\alpha\beta}\delta(x-y)$$

We define

$$M_{0\alpha\beta}(x,y) \equiv (-i\partial\!\!\!/_x + m)_{\alpha\beta}\delta(x-y)$$

$$M_{1\beta\gamma}(y,z) \equiv \delta_{\beta\gamma}\delta(y-z) + igS_F(y-z)_{\beta\gamma}\phi(z)$$

We can verify that

$$M_{\alpha\gamma}(x,z) = \int d^4y M_{0\alpha\beta}(x,y)M_{1\beta\gamma}(y,z)$$

Furthermore, we have $M_1 = I - G$, where I is the identity matrix and

$$G_{\alpha\beta}(x-y) = -igS_F(x-y)_{\alpha\beta}\phi(x)$$

So,

$$\log Z[\phi] = \log\det M_1 = -\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{Tr}G^n$$

where now

$$\mathrm{Tr}G^n = (-ig)^n\int d^4x_1\cdots d^4x_n \mathrm{tr}S_F(x_1-x_2)\phi(x_2)\cdots S_F(x_n-x_1)\phi(x_1)$$

and "tr" denotes a trace over spinor indices.

To better understand what it means, we will rederive it in a different way. Consider treating the

$-g\phi\overline{\Psi}\Psi$ term in $\mathcal{L}$ as an interaction. This leads to a vertex that connects two Ψ propagators; the associated vertex factor is $-ig\phi(x)$. So, we have

$$\log Z[\phi] = \sum_I C_I$$

where C_I represents connected diagram without external source. The only connected diagrams we can draw with these Feynman rules are fermion circles with n vertices. The has an n-fold cyclic symmetry, leading to a symmetry factor of $S = n$. The closed fermion loop implies a trace over the spinor indices. Thus the value of the n-vertex diagram is

$$\frac{1}{n}(-ig)^n \int d^4x_1 \cdots d^4x_n \mathrm{tr} S_F(x_1 - x_2)\phi(x_2) \cdots S_F(x_n - x_1)\phi(x_1)$$

Summing up these diagrams, we find that we are missing the overall minus sign. The appropriate conclusion is that we must associate an extra minus sign with each closed fermion loop.

Chapter 25
Vector Field

25.1 Vector field

Consider a vector field $A^\mu(x)$. Here the index μ is a vector index that takes on four possible values. Under a Lorentz transformation, we have

$$U(\Lambda)^{-1}A^\mu(x)U(\Lambda) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$$

For an infinitesimal transformation, we can write

$$\delta^\mu{}_\nu + \delta\omega^\mu{}_\nu = \delta^\mu{}_\nu + \frac{i}{2}\delta\omega_{\rho\sigma}(S_V^{\rho\sigma})^\mu{}_\nu$$

Here

$$(S_V^{\rho\sigma})^\mu{}_\nu = -i(\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu)$$

It is obvious that $A^{\dagger\mu}$ is also a vector field. We know that $\eta^{\mu\nu}$ is invariant under Lorentz transformation, i.e.

$$\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\eta^{\rho\sigma} = \eta^{\mu\nu}$$

We can use $\eta^{\mu\nu}$ and and its inverse $\eta_{\mu\nu}$ to raise and lower vector indices of the vector field,

$$A_\mu \equiv \eta_{\mu\nu}A^\nu$$

And we can verify the following equations

$$\Lambda^\mu{}_\nu\Lambda_\mu{}^\rho = \delta^\rho_\nu$$

$$A^\mu(x) = \eta^{\mu\nu}A_\nu(x)$$

$$\Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma\eta_{\rho\sigma} = \eta_{\mu\nu}$$

$$U(\Lambda)^{-1}A_\mu(x)U(\Lambda) = \Lambda_\mu{}^\nu A_\nu(\Lambda^{-1}x)$$

Define $C_i \equiv \frac{1}{2}\epsilon_{ijk}S_V^{jk}$,$D_i \equiv S_V^{i0}$. For example, we have

$$(C_3)_\mu{}^\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvectors of C_3 are

$$[(-1, [(0,\ 1,\ -i,\ 0)], 1), (1, [(0,\ 1,\ i,\ 0)], 1), (0, [(1,\ 0,\ 0,\ 0), (0,\ 0,\ 0,\ 1)], 2)]$$

We further define $N_i \equiv \frac{1}{2}(C_i - iD_i)$ and $N_i^\dagger \equiv \frac{1}{2}(C_i + iD_i)$. For example, we have

$$(N_1)_\mu{}^\nu = \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}i \\ 0 & 0 & \frac{1}{2}i & 0 \end{pmatrix}$$

The eigenvectors of N_1 are

$$\left[\left(-\frac{1}{2}, [(1,\ 1,\ 0,\ 0), (0,\ 0,\ 1,\ -i)], 2\right), \left(\frac{1}{2}, [(1,\ -1,\ 0,\ 0), (0,\ 0,\ 1,\ i)], 2\right)\right]$$

And we can conclude that vector is in the $(2, 2)$ representation of the Lie algebra of the Lorentz group.

25.2 Electromagnetic field and gauge invariance

The Lagrangian of EM field is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

Here,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad \text{and} \quad A^\mu = (\phi, \boldsymbol{A})$$

So,

$$F_{0i} = \dot{A}^i + \nabla_i \phi \equiv -E^i \quad \text{and} \quad F_{ij} = \nabla_i A^j - \nabla_j A^i \equiv \epsilon_{ijk} B^k$$

We can derive the equation of motion of the EM field by variation method,

$$\partial_\mu F^{\mu\nu} = 0$$

It can be rewritten in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$, i.e. Maxwell equations:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E} &= 0 \quad & \frac{\partial \boldsymbol{E}}{\partial t} &= \boldsymbol{\nabla} \times \boldsymbol{B} \\ \boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \quad & \frac{\partial \boldsymbol{B}}{\partial t} &= -\boldsymbol{\nabla} \times \boldsymbol{E} \end{aligned}$$

The massless vector field A_μ has 4 components, which would naively seem to tell us that the gauge field has 4 degrees of freedom.But there are two related comments which will ensure that quantizing the gauge field A_μ gives rise to 2 degrees of freedom, rather than 4.

- The field A_0 has no kinetic term $\dot{A}_0$ in the Lagrangian: it is not dynamical. This means that if we are given some initial data A_i and $\dot{A}_i$ at a time t_0, then the field A_0 is fully determined by the equation of motion $\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0$,which, expanding out, reads

  $$\nabla^2 A_0 = \boldsymbol{\nabla} \cdot \frac{\partial \boldsymbol{A}}{\partial t}$$

  So A_0 is not independent: we don't get to specify A_0 on the initial time slice.

- If we transform the EM field as

$$A_\mu \to A_\mu + \partial_\mu \lambda(x)$$

we can derive that

$$F_{\mu\nu} \to F_{\mu\nu} \quad \mathcal{L} \to \mathcal{L}$$

The seemed infinite number of symmetries, one for each function $\lambda(x)$, is to be viewed as a redundancy in our description. That is, two states related by a gauge symmetry are to be identified: they are the same physical state. One way to see that this interpretation is necessary is to notice that Maxwell's equations are not sufficient to specify the evolution of A_μ.The equations read,

$$(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)A^\nu = 0$$

But the operator $(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)$ is not invertible: it annihilates any function of the form $\partial_\mu\lambda$. This means that given any initial data, we have no way to uniquely determine A_μ at a later time since we can't distinguish between A_μ and $A_\mu + \partial_\mu\lambda$. This would be problematic if we thought that A_μ is a physical object. However, if we're happy to identify A_μ and $A_\mu + \partial_\mu\lambda$ as corresponding to the same physical state, then our problems disappear.

The picture that emerges for the theory of electromagnetism is of an enlarged phase space, foliated by gauge orbits. All states that lie along a given gauge orbit can be reached by a gauge transformation and are identified. To make progress, we pick a representative from each gauge orbit. It doesn't matter which representative we pick after all, they're all physically equivalent. But we should make sure that we pick a "good" gauge, in which we cut the orbits. Here we'll look at two different gauges:

- Coulomb Gauge: $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$
  We can make use of the residual gauge transformations in Lorentz gauge to pick $\boldsymbol{\nabla} \cdot \dot{\boldsymbol{A}} = 0$. We have as a consequence $A_0 = 0$. Coulomb gauge is sometimes called radiation gauge.

- Lorentz Gauge: $\partial^\mu A_\mu = 0$
  In fact this condition doesn't pick a unique representative from the gauge orbit. We're always free to make further gauge transformations with $\partial^\mu\partial_\mu\lambda = 0$, which also has non-trivial solutions. As the name suggests, the Lorentz gauge has the advantage that it is Lorentz invariant.

25.3 Canonical quantization of EM field

25.3.1 Canonical quantization in Coulomb gauge

Canonical momentum and Hamiltonian

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = 0 \quad \pi^i = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = \dot{A}^i + \nabla_i\phi = -E^i$$

$$\mathcal{H} = \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2) + (\boldsymbol{\pi} \cdot \boldsymbol{\nabla})A_0$$

Integration by parts can give

$$H = \int d^3x \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2)$$

Momentum and angular momentum

$$P^0 = H \quad \vec{P} = \int -\boldsymbol{\pi}\vec{\nabla}\boldsymbol{A} d^3x$$

$$\vec{J} = -\int \boldsymbol{\pi}(\vec{x} \times \vec{\nabla} + i\vec{C})\boldsymbol{A}\ d^3x \quad \vec{S} = -i\int \boldsymbol{\pi}\vec{C}\boldsymbol{A}\ d^3x$$

Canonical quantization

In Coulomb gauge, we have

$$A_0 = \pi^0 = 0 \quad \pi^i = \dot{A}^i$$

Three pairs of A_i and π^i are not independent from each other. They must satisfy the constraint equations

$$\nabla \cdot \boldsymbol{A} = 0 \quad \nabla \cdot \boldsymbol{\pi} = 0$$

A reasonable quantization condition can be written as

$$[A_i(\boldsymbol{x}, t), A_j(\boldsymbol{x}', t)] = 0 \quad [\pi^i(\boldsymbol{x}, t), \pi^j(\boldsymbol{x}', t)] = 0$$

$$[A_i(\boldsymbol{x}, t), \pi^j(\boldsymbol{x}', t)] = i\left(\delta_i^j - \frac{\partial_i \partial^j}{\nabla^2}\right)\delta(\boldsymbol{x} - \boldsymbol{x}') \equiv i\int \frac{d^3k}{(2\pi)^3}\ (\delta_i^j - \frac{k_i k^j}{\boldsymbol{k}^2})e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}$$

In this case, we can verify that

$$\dot{A}_i = -i[A_i(\boldsymbol{x}, t), H] = \pi_i(\boldsymbol{x}, t)$$

$$\dot{\pi}^i = -i[\pi^i(\boldsymbol{x}, t), H] = \nabla^2 A^i(\boldsymbol{x}, t)$$

It is constant with the field equation we derive from Euler-Lagrange equation.

Fourier expansion

$$\boldsymbol{A}(x) = \sum_{r=\pm} \int \widetilde{dp}[a_r(\boldsymbol{p})\boldsymbol{\epsilon}_r(\boldsymbol{p})e^{ipx} + a_r^\dagger(\boldsymbol{p})\boldsymbol{\epsilon}_r^*(\boldsymbol{p})e^{-ipx}]$$

And we can derive from constraint condition that

$$\boldsymbol{\epsilon} \cdot \boldsymbol{p} = 0$$

We will choose $\boldsymbol{\epsilon}$ to satisfy that

$$\boldsymbol{\epsilon}_r \cdot \boldsymbol{\epsilon}_s^* = \delta_{rs}$$

So, the completeness relation for the polarization vectors is

$$\sum_{r=\pm} \epsilon_r^i(\boldsymbol{p})\epsilon_r^{*j}(\boldsymbol{p}) = \delta^{ij} - \frac{p^i p^j}{|\boldsymbol{p}|^2}$$

Example: If $\boldsymbol{p} = (0, 0, p)$, we usually choose

$$\boldsymbol{\epsilon}_+ = \frac{1}{\sqrt{2}}(1, i, 0) \quad \boldsymbol{\epsilon}_- = \frac{1}{\sqrt{2}}(1, -i, 0)$$

$\boldsymbol{\epsilon}_+$ corresponds to anticlockwise rotation and it is the eigenvectors of the space-part of C_3 with eigenvalue $+1$. $\boldsymbol{\epsilon}_-$ corresponds to clockwise rotation and it is eigenvector of the space-part of C_3 with eigenvalue -1.

We can further derive from above discussion that

$$\boldsymbol{\pi}(x) = -i \sum_{r=\pm} \int \widetilde{dp} \omega [a_r(\boldsymbol{p}) \boldsymbol{\epsilon}_r(\boldsymbol{p}) e^{ipx} - a_r^\dagger(\boldsymbol{p}) \boldsymbol{\epsilon}_r^*(\boldsymbol{p}) e^{-ipx}]$$

$$a_r(\boldsymbol{p}) = \boldsymbol{\epsilon}_r^* \int d^3x e^{-ikx} (i\boldsymbol{\pi} + \omega \boldsymbol{A})$$

$$a_r^\dagger(\boldsymbol{p}) = \boldsymbol{\epsilon}_r \int d^3x e^{ikx} (-i\boldsymbol{\pi} + \omega \boldsymbol{A})$$

$$[a_r(\boldsymbol{p}), a_{r'}(\boldsymbol{p'})] = 0 \quad [a_r^\dagger(\boldsymbol{p}), a_{r'}^\dagger(\boldsymbol{p'})] = 0 \quad [a_r(\boldsymbol{p}), a_{r'}^\dagger(\boldsymbol{p'})] = (2\pi)^3 2\omega \delta_{rr'} \delta(\boldsymbol{p} - \boldsymbol{p'})$$

Operator represented by a and $a^\dagger$

Define that

$$N(\boldsymbol{p}, r) \equiv a_r^\dagger(\boldsymbol{p}) a_r(\boldsymbol{p})$$

So, we can derive

$$H = \sum_{r=\pm} \int \widetilde{dp}\, \omega N(\boldsymbol{p}, r) + 2\mathcal{E}_0 V$$

$$\vec{P} = \sum_{r=\pm} \int \widetilde{dp}\, \vec{p} N(\boldsymbol{p}, r)$$

$$\vec{S} = \sum_{r,s=\pm} \int \widetilde{dp}\, \frac{1}{2} (\boldsymbol{\epsilon}_s^* \vec{C} \boldsymbol{\epsilon}_r - \boldsymbol{\epsilon}_r \vec{C} \boldsymbol{\epsilon}_s^*) a_s^\dagger(\boldsymbol{p}) a_r(\boldsymbol{p})$$

From above equation, we can say that $a_r^\dagger(\boldsymbol{p})$ create an photon with energy ω, momentum $\boldsymbol{p}$ and spin angular momentum along the direction of momentum r.

Propagator

$$G_F(x-y)_{ij} \equiv \langle 0 | T A_i(x) A_j(y) | 0 \rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2 - i\epsilon} \left(\delta_{ij} - \frac{p_i p_j}{|\boldsymbol{p}|^2} \right) e^{ip(x-y)}$$

25.3.2 Canonical quantization in Lorentz gauge

Undefined metric formalism

Modify the Maxwell Lagrangian introducing a new term

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial_\mu A^\mu)^2$$

The equations of motion are now

$$\partial^2 A_\mu - (1 - \frac{1}{\xi})\partial^\mu(\partial \cdot A) = 0$$

Canonical momentums are

$$\pi^0 = \frac{1}{\xi}\partial \cdot A = \frac{1}{\xi}(-\dot{A}_0 + \partial_i A^i) \quad \pi^i = \dot{A}^i + \nabla^i A^0 = -E^i$$

Hamiltonian is

$$\mathcal{H} = \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2 - \xi\pi^0\pi^0) + (\boldsymbol{\pi} \cdot \boldsymbol{\nabla})A_0 + \pi^0(\boldsymbol{\nabla} \cdot \boldsymbol{A})$$

$$H = \int \left[\frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2 - \xi\pi^0\pi^0) - A_0(\boldsymbol{\nabla} \cdot \boldsymbol{\pi}) + \pi^0(\boldsymbol{\nabla} \cdot \boldsymbol{A})\right] d^3x$$

We remark that the above Lagrangian and the equations of motion, reduce to Maxwell theory in the gauge $\partial \cdot A = 0$. This why we say that our choice corresponds to a class of Lorenz gauges with parameter ξ. With this abuse of language (in fact we are not setting $\partial \cdot A = 0$, otherwise the problems would come back) the value of $\xi = 1$ is known as the Feynman gauge and $\xi = 0$ as the Landau gauge. From now on we will take the case of the so-called Feynman gauge, where $\xi = 1$. Then the equation of motion coincide with the Maxwell theory in the Lorenz gauge. In Feymann gauge,the canonical quantization conditions can be written as

$$[A_\mu(\boldsymbol{x}, t), A_\nu(\boldsymbol{x}', t)] = 0 \quad [\pi^\mu(\boldsymbol{x}, t), \pi^\nu(\boldsymbol{x}', t)] = 0 \quad [A_\mu(\boldsymbol{x}, t), \pi^\nu(\boldsymbol{x}', t)] = i\delta_\mu^\nu \delta(\boldsymbol{x} - \boldsymbol{x}')$$

we can also derive that

$$[\dot{A}_\mu(\boldsymbol{x}, t), \dot{A}_\nu(\boldsymbol{x}', t)] = 0 \quad [A_\mu(\boldsymbol{x}, t), \dot{A}_\nu(\boldsymbol{x}', t)] = i\eta_{\mu\nu}\delta(\boldsymbol{x} - \boldsymbol{x}')$$

Fourier expansion

$$A(x) = \sum_{\lambda=0}^{3} \int \widetilde{dp}[a_\lambda(\boldsymbol{p})\epsilon_\lambda(\boldsymbol{p})e^{ipx} + a_\lambda^\dagger(\boldsymbol{p})\epsilon_\lambda^*(\boldsymbol{p})e^{-ipx}]$$

where $\epsilon_{\lambda\mu}$ are a set of four independent 4-vectors. We will now make a choice for these 4-vectors. We choose $\epsilon_{1\mu}$ and $\epsilon_{2\mu}$ orthogonal to k^μ and n^μ, such that

$$\epsilon_{\lambda\mu}\epsilon_{\lambda'}^{*\mu} = \delta_{\lambda\lambda'} \quad \lambda, \lambda' = 1, 2$$

After, we choose $\epsilon_{3\mu}$ in the plane (k^μ, n^μ) and perpendicular to n^μ such that

$$\epsilon_{3\mu}n^\mu = 0 \quad \epsilon_{3\mu}\epsilon_3^{*\mu} = 1$$

Finally we choose $\epsilon_{0\mu} = n_\mu$. The vectors $\epsilon_{1\mu}$ and $\epsilon_{2\mu}$ are called transverse polarizations, while $\epsilon_{3\mu}$ and $\epsilon_{0\mu}$ longitudinal and scalar polarizations, respectively.
In general we can show that

$$\epsilon_\lambda \cdot \epsilon_{\lambda'}^* = \eta_{\lambda\lambda'} \quad \eta^{\lambda\lambda'}\epsilon_{\lambda\mu}\epsilon_{\lambda'\nu}^* = \eta_{\mu\nu}$$

We can further derive from above discussion that

$$\dot{A}(x) = -i\sum_{\lambda=0}^{3}\int \widetilde{dp}\omega[a_\lambda(\boldsymbol{p})\epsilon_\lambda(\boldsymbol{p})e^{ipx} - a_\lambda^\dagger(\boldsymbol{p})\epsilon_\lambda^*(\boldsymbol{p})e^{-ipx}]$$

$$a_\lambda(\boldsymbol{p}) = \eta_{\lambda\lambda'}\epsilon_{\lambda'}^* \cdot \int d^3x e^{-ipx}(i\dot{A} + \omega A)$$

$$a_\lambda^\dagger(\boldsymbol{p}) = \eta_{\lambda\lambda'}\epsilon_{\lambda'} \cdot \int d^3x e^{ipx}(-i\dot{A} + \omega A)$$

$$[a_\lambda(\boldsymbol{p}), a_{\lambda'}(\boldsymbol{p'})] = 0 \quad [a_\lambda^\dagger(\boldsymbol{p}), a_{\lambda'}^\dagger(\boldsymbol{p'})] = 0 \quad [a_\lambda(\boldsymbol{p}), a_{\lambda'}^\dagger(\boldsymbol{p'})] = (2\pi)^3 2\omega\eta_{\lambda\lambda'}\delta(\boldsymbol{p} - \boldsymbol{p'})$$

Indefinite metric problem

We Introduce the vacuum state defined by

$$a_\lambda(\boldsymbol{p})|0\rangle = 0$$

To see the problem with the sign we construct the one-particle state with scalar polarization, that is

$$|1\rangle = \int \widetilde{dp} f(p) a_0^\dagger(\boldsymbol{p})|0\rangle$$

and calculate its norm

$$\langle 1|1\rangle = -\langle 0|0\rangle \int \widetilde{dp}|f(p)|^2$$

The state $|1\rangle$ has a negative norm.
To solve this problem we note that we are not working anymore with the classical Maxwell theory because we modified the Lagrangian. What we would like to do is to impose the condition $\partial \cdot A = 0$, but that is impossible as an equation for operators. We can, however, require that condition on a weaker form, as a condition only to be verified by the physical states.
More specifically, we require that the part of $\partial \cdot A$ that contains the annihilation operator (positive frequencies) annihilates the physical states,

$$\partial^\mu A_\mu^+|\psi\rangle = 0$$

The states $|\psi\rangle$ can be written in the form

$$|\psi\rangle = |\psi_T\rangle|\phi\rangle$$

where $|\psi_T\rangle$ is obtained from the vacuum with creation operators with transverse polarization and $|\phi\rangle$ with scalar and longitudinal polarization.
$\partial^\mu A_\mu^+$ contains only scalar and longitudinal polarizations

$$\partial^\mu A_\mu^+ = i\sum_{\lambda=0,3}\int \widetilde{dp} a_\lambda(\boldsymbol{p})(p \cdot \epsilon_\lambda(\boldsymbol{p}))e^{ipx}$$

Therefore the previous condition becomes

$$i\sum_{\lambda=0,3}(p \cdot \epsilon_\lambda(\boldsymbol{p}))a_\lambda(\boldsymbol{p})|\phi\rangle = 0$$

The condition is equivalent to,

$$(a_0(\boldsymbol{p}) - a_3(\boldsymbol{p}))|\phi\rangle = 0$$

We can construct $|\phi\rangle$ as a linear combination of states $|\phi\rangle$ with n scalar or longitudinal photons:

$$|\phi\rangle = C_0|\phi_0\rangle + C_1|\phi\rangle + \cdots \quad \text{Here,} |\phi_0\rangle \equiv |0\rangle$$

The states $|\phi_n\rangle$ are eigenstates of the operator number for scalar or longitudinal photons

$$N'|\phi_n\rangle = n|\phi_n\rangle$$

where,

$$N' = \int \widetilde{dp}[a_3^\dagger(\boldsymbol{p})a_3(\boldsymbol{p}) - a_0^\dagger(\boldsymbol{p})a_0(\boldsymbol{p})]$$

Then

$$n\langle\phi_n|\phi_n\rangle = \langle\phi_n|N'|\phi_n\rangle = 0$$

This means that

$$\langle\phi_n|\phi_n\rangle = \delta_{n0}$$

that is, for $n \neq 0$, the state $|\phi_n\rangle$ has zero norm. We have then for the general state $|\phi\rangle$,

$$\langle\phi|\phi\rangle = |C_0|^2 \geq 0$$

and the coefficients $C_i(i = 1, 2, \cdots)$ are arbitrary.

Operator represented by a and $a^\dagger$

Define that

$$N'(\boldsymbol{p}) \equiv a_3^\dagger(\boldsymbol{p})a_3(\boldsymbol{p}) - a_0^\dagger(\boldsymbol{p})a_0(\boldsymbol{p})$$

$$N(\boldsymbol{p}, 1) \equiv a_1^\dagger(\boldsymbol{p})a_1(\boldsymbol{p}) \quad N(\boldsymbol{p}, 2) \equiv a_2^\dagger(\boldsymbol{p})a_2(\boldsymbol{p}) \quad N_T(\boldsymbol{p}) \equiv N(\boldsymbol{p}, 1) + N(\boldsymbol{p}, 2)$$

We have that

$$\langle\psi|N'(\boldsymbol{p})|\psi\rangle = 0 \quad \langle\psi|N_T(\boldsymbol{p})|\psi\rangle = \langle\psi_T|N_T(\boldsymbol{p})|\psi_T\rangle$$

We can derive

$$H = \int \widetilde{dp}\, \omega[N'(\boldsymbol{p}) + N_T(\boldsymbol{p})] + 2\mathcal{E}_0 V$$

$$\vec{P} = \int \widetilde{dp}\, \vec{p}[N'(\boldsymbol{p}) + N_T(\boldsymbol{p})]$$

So, the arbitrariness of $C_i(i = 1, 2, \cdots)$ does not affect the physical observables. Only the physical transverse polarizations contribute to the result. Two states that differ only in their timelike and longitudinal photon content, $|\phi_n\rangle$ with $n \geq 1$ are said to be physically equivalent. We can think of the gauge symmetry of the classical theory as descending to the Hilbert space of the quantum theory.

It is important to note that although for the average values of the physical observables only the transverse polarizations contribute, the scalar and longitudinal polarizations are necessary for the consistency of the theory. In particular they show up when we consider complete sums over the intermediate states.

Propagator

$$G_F(x-y)_{\mu\nu} \equiv \langle 0|TA_\mu(x)A_\nu(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i\eta_{\mu\nu}}{p^2-i\epsilon}e^{ip(x-y)}$$

It is easy to verify that $G_F(x-y)_{\mu\nu}$ is the Green's function of the equation of motion, that for $\xi = 1$ is the wave equation, that is

$$\partial^2 G_F(x-y)_{\mu\nu} = i\eta_{\mu\nu}\delta(x-y)$$

For the general case, $\xi \neq 0$, the equal times commutation relations are more complicated. And the propagator will be

$$G_F(x-y)_{\mu\nu} = \int \frac{d^4p}{(2\pi)^4}\left[\frac{-i\eta_{\mu\nu}}{p^2-i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2-i\epsilon)^2}\right]e^{ip(x-y)}$$

25.4 Perturbation theory for canonical quantization

25.4.1 Lagrangian of QED

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi + e_0 j^\mu A_\mu,$$

where $j^\mu \equiv \overline{\Psi}\gamma^\mu\Psi$. Usually, we also define a covariant derivative,

$$D_\mu\Psi \equiv \partial_\mu\Psi - ie_0A_\mu\Psi$$

So, the Lagrangian can also be written as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(iD\!\!\!/ - m_0)\Psi$$

The Lagrangian is invariant under the gauge transformation

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e_0}\partial_\mu\alpha(x) \quad \Psi(x) \to e^{i\alpha(x)}\Psi(x)$$

25.4.2 Coulomb gauge

The constraint equations are

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0 \quad \boldsymbol{\nabla}^2 A_0 = e_0 j^0$$

The solution for A_0 is

$$A_0(\boldsymbol{x},t) = -e_0\int d^3x'\frac{j^0(\boldsymbol{x}',t)}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}$$

Finally, we can derive that

$$H = H_D + H_M + H_{\text{int}}$$

where,

$$H_D = \int d^3x \ -\Pi(\vec{\alpha}\cdot\vec{\nabla} + i\beta m)\Psi \quad H_M = \int d^3x \ \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2)$$

$$H_{\text{int}} = \int d^3x \ \left[-e_0\boldsymbol{j}\cdot\boldsymbol{A} + \frac{e_0^2}{2}\int d^3x' \frac{j^0(\boldsymbol{x})j^0(\boldsymbol{x}')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\right]$$

The perturbation expansion of correlation functions is

$$\langle\Omega|T\{\Psi(x)\overline{\Psi}(y)A(z)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\Psi_I(x)\overline{\Psi}_I(y)A_I(z)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}$$

Here, $\int_T^T dtH_I$ can be written as

$$\left[-\int d^4x e_0\overline{\Psi}_I\boldsymbol{\gamma}\Psi_I\cdot\boldsymbol{A}_I\right] + \left[\int d^4x\int d^4x' \frac{e_0^2\delta(t-t')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\ \frac{1}{2}\overline{\Psi}_I(\boldsymbol{x},t)\gamma^0\Psi_I(\boldsymbol{x},t)\overline{\Psi}_I(\boldsymbol{x}',t')\gamma^0\Psi_I(\boldsymbol{x}',t')\right]$$

Wick' s theorem for photons takes is similar to that in ϕ^4 theory:

$$T\left\{A_I(x_1)A_I(x_2)A_I(x_3)\cdots\right\} = N\left\{A_I(x_1)A_I(x_2)A_I(x_3)\cdots + \text{ all possible contractions }\right\}$$

Example:

$$\begin{aligned} & \langle 0|T\left\{A_{Ii}(x_1)A_{Ij}(x_2)A_{Ik}(x_3)A_{Il}(x_4)\right\}|0\rangle \\ = \ & G_F(x_1-x_2)_{ij}G_F(x_3-x_4)_{kl} + G_F(x_1-x_3)_{ik}G_F(x_2-x_4)_{jl} + G_F(x_1-x_4)_{il}G_F(x_2-x_3)_{jk} \end{aligned}$$

Now we can derive the Feynman rule for QED theory. Firstly, we evaluate this term,

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)A_{Ii}(z)(ie_0)\int dw^4\overline{\Psi}_I(w)\boldsymbol{\gamma}\Psi_I(w)\cdot A_I(w)\right\}|0\rangle$$

After contraction, it can be written as

$$\begin{aligned} - \quad & (ie_0)S_F(x-y)_{ab}\int d^4wG_F(z-w)_{ik}\mathrm{Tr}[\gamma^k S_F(w-w)] \\ + \quad & (ie_0)\int d^4wG_F(w-z)_{ik}[S_F(x-w)\gamma^k S_F(w-y)]_{ab} \end{aligned}$$

It can be represented by the following Feynman diagram.

Figure 25.1: Feynman diagram representation of perturbation expansion

Secondly, we evaluate this term,

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\int d^4w\int d^4w' \frac{-ie_0^2\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}\ \frac{1}{2}\overline{\Psi}_I(w)\gamma^0\Psi_I(w)\overline{\Psi}_I(w')\gamma^0\Psi_I(w')\right\}|0\rangle$$

After contraction, it can be written as

$$
\begin{aligned}
- \quad & (ie_0)^2 \int d^4w d^4w' \frac{i\delta(w^0 - w'^0)}{4\pi|\boldsymbol{w} - \boldsymbol{w}'|} [S_F(x-w)\gamma^0 S_F(w-y)]_{ab} \mathrm{Tr}[\gamma^0 S_F(w'-w')] \\
+ \quad & (ie_0)^2 \int d^4w d^4w' \frac{i\delta(w^0 - w'^0)}{4\pi|\boldsymbol{w} - \boldsymbol{w}'|} [S_F(x-w)\gamma^0 S_F(w-w')\gamma^0 S_F(w'-y)]_{ab} \\
+ \quad & \frac{1}{2}(ie_0)^2 S_F(x-y)_{ab} \int d^4w d^4w' \frac{i\delta(w^0 - w'^0)}{4\pi|\boldsymbol{w} - \boldsymbol{w}'|} \mathrm{Tr}[\gamma^0 S_F(w-w)] \mathrm{Tr}[\gamma^0 S_F(w'-w')] \\
- \quad & \frac{1}{2}(ie_0)^2 S_F(x-y)_{ab} \int d^4w d^4w' \frac{i\delta(w^0 - w'^0)}{4\pi|\boldsymbol{w} - \boldsymbol{w}'|} \mathrm{Tr}[\gamma^0 S_F(w-w')\gamma^0 S_F(w'-w)]
\end{aligned}
$$

It can be represented by the following Feynman diagram.

Figure 25.2: Feynman diagram representation of perturbation expansion

Before we write down Feymann rules, we notice that the offending non-local interaction comes from the A_0 component of the gauge field, we could try to redefine the propagator to include a $G_F(x-y)_{00}$ piece which will capture this term. We can verify that

$$
\frac{i\delta(w^0 - w'^0)}{4\pi|\boldsymbol{w} - \boldsymbol{w}'|} = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{ip(w-w')}}{|\boldsymbol{p}|^2}
$$

So we can combine the non-local interaction with the transverse photon propagator by defining a new photon propagator

$$
G_F(p)_{\mu\nu} \equiv \begin{cases} \frac{i}{|\boldsymbol{p}|^2} & \mu, \nu = 0 \\ \frac{-i}{p^2 - i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\boldsymbol{p}|^2}\right) & \mu = i \neq 0, \nu = j \neq 0 \\ 0 & \text{otherwise} \end{cases}
$$

With this propagator, the wavy photon line now carries a $\mu\nu = 0, 1, 2, 3$ index, with the extra $\mu = 0$ component taking care of the instantaneous interaction.
The Feynman rules for QED are:

1. For each Fermion propagator from y to x, $P = S_F(x-y)$
2. For each vector propagator, $P = G_F(x-y)$
3. For each vertex, $V = (ie_0\gamma^\mu) \int d^4w$
4. For each external point, $E = 1$
5. Divided by the symmetry factor

25.4.3 Lorentz Gauge

In Lorentz gauge,

$$\mathcal{H}_{int} = -e_0 \overline{\Psi} \gamma^\mu \Psi A_\mu$$

The Feynman rules for QED in Lorentz gauge will be the same as that in Coulomb gauge expect for that the vector propagator will be

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2 - i\epsilon)^2}$$

Especially, for Feynman gauge $\xi = 1$, we have

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon}$$

25.5 Path integral quantization

25.5.1 Path integral formulation for free EM field

The correlation function is given by

$$\langle \Omega | T A_H(x_1) A_H(x_2) | \Omega \rangle = \lim_{T \to \infty(1-i\epsilon)} \frac{\int \mathcal{D}A \exp\left[i \int_{-T}^{T} d^4x (-\frac{1}{4} F^{\mu\nu} F_{\mu\nu})\right] A(x_1) A(x_2)}{\int \mathcal{D}A \exp\left[i \int_{-T}^{T} d^4x (-\frac{1}{4} F^{\mu\nu} F_{\mu\nu})\right]}$$

The generating function is

$$Z[J] = \int \mathcal{D}A \exp\left[i \int d^4x (-\frac{1}{4} F^{\mu\nu} F_{\mu\nu}) + J^\mu A_\mu\right]$$

We can verify that

$$S = \int d^4x (-\frac{1}{4} F^{\mu\nu} F_{\mu\nu}) = \frac{1}{2} \int d^4x A_\mu(x) (\partial^2 \eta^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu(x)$$

Notice that $(\partial^2 \eta^{\mu\nu} - \partial^\mu \partial^\nu)$ is singular, since for any $\alpha(x)$,

$$(\partial^2 \eta^{\mu\nu} - \partial^\mu \partial^\nu) \partial_\mu \alpha(x) = 0$$

This difficulty is due to gauge invariance: $\alpha(x)$ is gauge equivalent to 0. The functional is badly defined because we are redundantly integrating over a continuous infinity of physically equivalent field configurations. To fix the problem, we would like to isolate the interesting part of the functional integral, which counts each physical configuration only once.
Let $G(A)$ be some function that we wish to set equal zero as a gauge-fixing condition. We could constrain the functional integral to cover only the configurations with $G(A) = 0$ by inserting a functional delta function, $\delta(G(A))$. To do so, we insert 1 in the path integral:

$$1 = \int \mathcal{D}\alpha(x) \delta(G(A(\alpha))) \det\left(\frac{\delta G}{\delta \alpha}\right)$$

where,

$$A_\mu(\alpha(x)) = A_\mu(x) + \frac{1}{e_0}\partial_\mu\alpha(x)$$

We set the gauge fixing function as $G(A) = \partial^\mu A_\mu - \omega(x)$, so $G(A(\alpha)) = \partial^\mu A_\mu + \frac{1}{e_0}\partial^2\alpha - \omega(x)$. It is obvious that $\det\left(\frac{\delta G}{\delta\alpha}\right)$ is equivalent to $\det(\partial^2)/e_0$, which is independent of A. So,

$$Z[0] = \det\left(\frac{\delta G}{\delta\alpha}\right)\int \mathcal{D}\alpha \int \mathcal{D}A e^{iS[A]}\delta(G(A(\alpha)))$$

Now change variables from A to $A(\alpha)$. This is a simple shift, so $\mathcal{D}A = \mathcal{D}A(\alpha)$. Also, by gauge invariance, $S[A] = S[A(\alpha)]$. Since $A(\alpha)$ is now just a dummy integration variable, we can rename it bake to A, so

$$Z[0] = \det\left(\frac{\delta G}{\delta\alpha}\right)\int \mathcal{D}\alpha \int \mathcal{D}A e^{iS[A]}\delta(\partial^\mu A_\mu - \omega(x))$$

Since the above equation is hold for any $\omega(x)$,so we have

$$\begin{aligned}
Z[0] &= N(\xi)\int \mathcal{D}\omega \exp\left[-i\int d^4x\frac{\omega^2}{2\xi}\right]\det(\partial^2)\int \mathcal{D}\alpha \int \mathcal{D}A e^{iS[A]}\delta(\partial^\mu A_\mu - \omega(x)) \\
&= N(\xi)\frac{\det(\partial^2)}{e_0}\int \mathcal{D}\alpha \int \mathcal{D}A e^{iS[A]}\exp\left[-i\int d^4x\frac{1}{2\xi}(\partial^\mu A_\mu)^2\right] \\
&= W(\xi)\int \mathcal{D}A \exp\left[\frac{i}{2}\int d^4x A_\mu(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)A_\nu\right]
\end{aligned}$$

And we rewrite the generating function as

$$Z[J] = W(\xi)\int \mathcal{D}A \exp\left[\frac{i}{2}\int d^4x A_\mu(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)A_\nu + J^\mu A_\mu\right]$$

Define

$$A'(x) = A(x) - i\int d^4y G_F(x-y)J(y)$$

Recall that

$$(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)G_F(x-y)_{\nu\rho} = i\delta^\mu_\rho\delta(x-y)$$

We can derive that

$$Z[J] = Z[0]\exp\left[-\frac{1}{2}\int d^4x d^4y J^\mu(x)G_F(x-y)_{\mu\nu}J^\nu(y)\right]$$

The two point correlation functions are

$$\langle 0|TA_H(x_1)A_H(x_2)|0\rangle = Z[0]^{-1}\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right)Z[J]\Big|_{J=0} = G_F(x_1 - x_2)$$

25.5.2 Perturbation theory for path integral quantization

We use QED theory as an example.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi + e_0\overline{\Psi}\gamma^\mu\Psi A_\mu$$

As we stated in section 4.1, the Lagrangian of QED is also invariant under a general gauge transformation. And we also notice that the measure $\mathcal{D}\Psi\mathcal{D}\overline{\Psi}$ is invariant under gauge transformation. By the similar method, we can show that

$$Z[0] \equiv \int \mathcal{D}A\mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS[A,\Psi,\overline{\Psi}]} = W(\xi)\int \mathcal{D}A\mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS[A,\Psi,\overline{\Psi}]}\exp\left[-i\int d^4x\frac{1}{2\xi}(\partial^\mu A_\mu)^2\right]$$

So,define

$$\mathcal{L}_0 \equiv -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2 + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi \quad \mathcal{L}_1 \equiv e_0\overline{\Psi}\gamma^\mu\Psi A_\mu$$

$$\begin{aligned}
Z[J] &= W(\xi)\int \mathcal{D}A\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+JA+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&= W(\xi)e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\int \mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}Ae^{i\int d^4y[\mathcal{L}_0+JA+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\exp[-\int d^4yd^4z\frac{1}{2}J(y)G_F(y-z)J(z) + \overline{\eta}(y)S_F(y-z)\eta(z)] \\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[ie_0\int d^4x(\frac{1}{i}\frac{\delta}{\delta J^\mu(x)}\cdot -\frac{1}{i}\frac{\delta}{\delta\eta(x)}\cdot\gamma^\mu\cdot\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)})]^V \\
&\times \sum_{P_1=0}^{\infty}\frac{1}{P_1!}[-\frac{1}{2}\int d^4y_1d^4z_1J(y_1)G_F(y_1-z_1)J(z_1)]^{P_1} \\
&\times \sum_{P_2=0}^{\infty}\frac{1}{P_2!}[-\int d^4y_2d^4z_2\overline{\eta}(y_2)S_F(y_2-z_2)\eta(z_2)]^{P_2}
\end{aligned}$$

If we focus on a term with particular values of V, P_1 and P_2, then the number of surviving scalar sources is $E_1 = 2P_1 - V$, the number of surviving fermion sources is $E_2 = 2P_2 - 2V$. We can introduce Feynman diagrams as in the ϕ^4 theory. In these diagrams, a wavy line segment stands for a vector propagator $G_F(x-y)$, a line with an arrow pointing from y to x for a fermion propagator $S_F(x-y)$, a filled circle at one end of a wavy line segment for a vector source $i\int d^4xJ(x)$, a filled circle at the start of a line with an arrow for a fermion source $i\int d^4x\eta(x)$, a filled circle at the end of a line with an arrow for a anti-fermion source $i\int d^4x\overline{\eta}(x)$, a vertex joining three line segments for $ie_0\gamma^\mu\int d^4x$.

25.5.3 Ward-Takahahsi identity (1)

The Noether current of the symmetry $\Psi \to e^{i\alpha}\Psi$ is $j^\mu = \overline{\Psi}\gamma^\mu\Psi$. Recall the conservation law in functional formalism

$$\langle\partial_\mu j^\mu(x)\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n}\langle\phi(x_1)\cdots(i\Delta\phi(x_i)\delta(x-x_i))\cdots\phi(x_n)\rangle$$

So, we can write the charge conservation law as

$$ie_0\partial_\mu\langle\Omega|Tj^\mu\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle = -ie_0\delta(x-x_1)\langle\Omega|T\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle+ie_0\delta(x-x_2)\langle\Omega|T\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle$$

Notice that $ie_0\langle\Omega|Tj^\mu\psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle$ can be represented by

Figure 25.3: Feynman diagram representation of correlation function

From the diagram, we have

$$\langle A_\nu(y)\rangle = \int d^4x G_F(x-y)_{\mu\nu} ie_0\langle j^\mu(x)\rangle = \int \frac{d^4p}{(2\pi)^4} e^{-ipy} G_F(p)_{\mu\nu} \int d^4x e^{ipx} ie_0\langle j^\mu(x)\rangle$$

So,

$$\int d^4y\langle A_\nu(y)\rangle e^{iky} = G_F(k)_{\mu\nu}\int d^4x e^{ikx} ie_0\langle j^\mu(x)\rangle$$

Compute the Fourier transformation of the equation of charge conservation by integrating

$$\int d^4x e^{ikx}\int d^4x_1 e^{-iqx_1}\int d^4x_2 e^{ipx_2}$$

We can get

Figure 25.4: Feynman diagram representation of Ward identity

Note that in the diagram above, the external leg of photon will be cut-off, but external leg of fermion will remain. The above equation can be generated to the diagram with n external fermions. Another proof of Ward-Takahashi identity by calculating the Feynman diagram directly can be found in chapter 7.4 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

25.6 Exact propagator of photon

25.6.1 Photon self-energy

The exact propagator of photon is

$$\mathcal{G}(x)_{\mu\nu} = \langle\Omega|T\{A_\mu(x)A_\nu(0)|\Omega\rangle_C$$

Figure 25.5: Feynman diagram representation of exact propagator of photon

Its Fourier transformation can be represented by the following diagram.
Let us define $i\Pi^{\mu\nu}$ to be the sum of all 1-particle-irreducible insertions into the photon propagator. So, we have

$$\mathcal{G}(k) = G_F(k) + G_F(k)(i\Pi(k))G_F(k) + \cdots = G_F(k)\frac{1}{1 - i\Pi(k)G_F(k)}$$

Hence,

$$(i\mathcal{G})^{-1} = (iG_F)^{-1} - \Pi$$

Recall that

$$iG_F(p)_{\mu\nu} = \frac{\eta_{\mu\nu}}{k^2 - i\epsilon} - (1-\xi)\frac{k_\mu k_\nu}{(k^2 - i\epsilon)^2} = \frac{1}{k^2 - i\epsilon}(P^T_{\mu\nu} + \xi P^L_{\mu\nu})$$

Here,

$$P^T_{\mu\nu} \equiv \eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \quad P^L_{\mu\nu} \equiv \frac{k_\mu k_\nu}{k^2}$$

We can derive that

$$(iG_F)^{-1}_{\mu\nu} = k^2(P^T_{\mu\nu} + \frac{1}{\xi}P^L_{\mu\nu})$$

We may also expand $i\Pi^{\mu\nu}$ as

$$\Pi^{\mu\nu} = P^{\mu\nu}_T f_T(k^2) + P^{\mu\nu}_L f_L(k^2) = \eta^{\mu\nu} f_T + \frac{k^\mu k^\nu}{k^2}(f_L - f_T)$$

Therefore,

$$(i\mathcal{G})^{-1}_{\mu\nu} = (k^2 - f_T(k^2))P^T_{\mu\nu} + (\frac{k^2}{\xi} - f_L(k^2))P^L_{\mu\nu}$$

$$\mathcal{G}(k)_{\mu\nu} = \frac{-i}{k^2 - f_T(k^2)}P^T_{\mu\nu} + \frac{-i}{\frac{k^2}{\xi} - f_L(k^2)}P^L_{\mu\nu}$$

We observe that if $f_{T,L}(k^2 = 0) \neq 0$, a mass will be generated for the photon. Because $\Pi(k)$ comes from 1PI diagrams, it should not be singular at $k^2 = 0$, and so $f_L - f_T = O(k^2)$, as $k \to 0$. We will show that gauge invariance ensures that no mass is generated from the loop corrections

25.6.2 Ward identities(2)

We define the generating functional for connected diagrams

$$Z[J, \eta, \overline{\eta}] = e^{-iE[J,\eta,\overline{\eta}]}$$

For example,

$$\mathcal{G}(x-y)_{\mu\nu} = -i\frac{\delta^2 E[J,\eta,\overline{\eta}]}{\delta J^\mu(x)\delta J^\nu(y)}\bigg|_{J,\eta,\overline{\eta}=0}$$

Recall, for infinitesimal gauge transformations, $\delta A_\mu = \partial_\mu \lambda$, $\delta \Psi = ie_0 \lambda \Psi$ and $\delta \overline{\Psi} = -ie_0 \lambda \overline{\Psi}$. For a change of variables in the path integral, $Z[J, \eta, \overline{\eta}]$ will remain the same. Recall that

$$Z[J, \eta, \overline{\eta}] = W(\xi) \int \mathcal{D}A \mathcal{D}\overline{\Psi} \mathcal{D}\Psi e^{i \int d^4x [\mathcal{L}_0 + \mathcal{L}_1 + JA + \overline{\eta}\Psi + \overline{\Psi}\eta]}$$

The change of action is

$$\delta S = -\frac{1}{\xi} \int d^4x \partial_\mu A^\mu \partial^2 \lambda + \int d^4x J^\mu \partial_\mu \lambda + ie_0 \overline{\eta} \Psi \lambda - ie_0 \overline{\Psi} \eta \lambda$$

Hence, we must have

$$\int d^4x \lambda(x) W(\xi) \int \mathcal{D}A \mathcal{D}\overline{\Psi} \mathcal{D}\Psi e^{iS} \left[-\frac{1}{\xi} \partial^2 \partial_\mu A^\mu - \partial_\mu J^\mu + ie_0 (\overline{\eta}\Psi - \overline{\Psi}\eta) \right] = 0$$

Since

$$\langle A_\mu(x) \rangle_{J,\eta,\overline{\eta}} = -\frac{\delta E}{\delta J^\mu} \quad \langle \Psi(x) \rangle_{J,\eta,\overline{\eta}} = -\frac{\delta E}{\delta \overline{\eta}} \quad \langle \overline{\Psi}(x) \rangle_{J,\eta,\overline{\eta}} = \frac{\delta E}{\delta \eta}$$

The above equation can be written as

$$\frac{1}{\xi} \partial^2 \partial^\mu \frac{\delta E}{\delta J^\mu} - \partial_\mu J^\mu - ie_0 \left[\overline{\eta} \frac{\delta E}{\delta \overline{\eta}} + \frac{\delta E}{\delta \eta} \eta \right] = 0$$

We can derive that

$$\frac{1}{\xi} \partial^2 \partial^\mu \frac{\delta^2 E[J, \eta, \overline{\eta}]}{\delta J^\mu(x) \delta J^\nu(y)} \bigg|_{J,\eta,\overline{\eta}=0} - \partial_\nu \delta(x-y) = 0$$

that is,

$$\frac{i}{\xi} \partial^2 \partial^\mu \mathcal{G}(x-y)_{\mu\nu} + \partial_\nu \delta(x-y) = 0$$

or, written in momentum-space,

$$-\frac{i}{\xi} k^2 k^\mu \mathcal{G}(k)_{\mu\nu} + k_\nu = 0$$

So

$$-\frac{k^2}{k^2 - \xi f_L(k^2)} k_\nu + k_\nu = 0$$

Which means $f_L(k^2) = 0$ and so, we have $f_T(k^2) \to O(k^2)$ as $k^2 \to 0$. The exact propagator of photon is

$$\mathcal{G}(k)_{\mu\nu} = \frac{-i}{k^2(1 - \pi(k^2))} P^T_{\mu\nu} + \frac{-i\xi}{k^2} P^L_{\mu\nu}$$

where $\pi(k^2) \equiv \frac{f_T(k^2)}{k^2}$

25.7 LSZ reduction formula

25.7.1 LSZ reduction formula and Feynman rules

Suppose that the probability for the quantum field to create or annihilate an exact one-particle eigenstate of H is Z_3, i.e.

$$\langle\Omega|A(0)|p,\lambda\rangle = \sqrt{Z_3}\epsilon_\lambda(p)$$

In Feynman gauge, because the norm of $|p,0\rangle$ is negative, the expansion of orthogonal complete set will be written as

$$\frac{d^3q}{(2\pi)^3}\frac{1}{2E_q}\eta^{\lambda\lambda'}|p,\lambda\rangle\langle p,\lambda'|$$

We have demonstrated that photon will remain massless when interacting with charged fermions. And recall that $\sum_\lambda \xi_\lambda(p)\xi^*_\lambda(p) = \eta_{\mu\nu}$. So, we can derive by the similar method in ϕ^4 theory that

$$\int d^4x\, e^{-ipx}\langle\Omega|TA_\mu(x)A_\nu(0)|\Omega\rangle_C = \frac{-iZ_3\eta_{\mu\nu}}{p^2-i\epsilon}+\cdots$$

The LSZ reduction formula for fermions would take the form as

Theorem 25.1 LSZ reduction formula

$$\begin{aligned}
&\langle \boldsymbol{p}_1\cdots\boldsymbol{p}_n\,|S|\,\boldsymbol{k}_1\cdots\boldsymbol{k}_m\,\rangle\\
=\;&\prod_1^n\int d^4x_ie^{-ip_ix_i}\prod_1^m\int d^4y_je^{ik_jy_j}\\
\times\;&\left(\frac{i}{\sqrt{Z_3}}\right)^{m+n}[p_1^2\epsilon^{*\mu_1}_{\lambda_1}(p_1)]\cdots[p_n^2\epsilon^{*\mu_n}_{\lambda_n}(p_n)][k_1^2\epsilon^{\nu_1}_{\lambda'_1}(k_1)]\cdots[k_m^2\epsilon^{\nu_m}_{\lambda'_m}(p_m)]\\
\times\;&\langle\Omega|T\{A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n)A_{\nu_1}(y_1)\cdots A_{\nu_m}(y_m)\}|\Omega\rangle
\end{aligned}$$

The LSZ reduction formula in other gauge would give similar procedure for calculating scattering amplitude: Fourier transform the Green function in position space to momentum space, cut-off the external legs and multiply the polarization vector of asymptotic states. (Note there is still an extra factor $\sqrt{Z_3}^{m+n}$ to multiply, similar to the ϕ^4 theory).
Finally, we list the Feymann rules of QED in momentum space as follows:

1. For each incoming electron, draw a solid line with an arrow pointed towards the vertex, and label it with the electron's four-momentum, p_i.
2. For each outgoing electron, draw a solid line with an arrow pointed away from the vertex, and label it with the electron's four-momentum, p'_i.
3. For each incoming positron, draw a solid line with an arrow pointed away from the vertex, and label it with minus the positron's four-momentum, $-p_i$.
4. For each outgoing positron, draw a solid line with an arrow pointed towards the vertex, and label it with minus the positron's four-momentum, $-p'_i$.

5. For each incoming photon, draw a wavy line with an arrow pointed towards the vertex, and label it with the photon's four-momentum,k_i.

6. For each outgoing photon, draw a wavy line with an arrow pointed away from the vertex, and label it with the photon's four-momentum, k_i'.

7. The only allowed vertex joins two solid lines, one with an arrow pointing towards it and one with an arrow pointing away from it, and one wavy line. Using this vertex, join up all the external lines, including extra internal lines as needed. In this way, draw all possible diagrams that are topologically inequivalent.

8. Assign each internal line its own four-momentum. Think of the four-momenta as flowing along the arrows, and conserve four-momentum at each vertex.

9. The value of a diagram consists of the following factors:

   - for each incoming photon, $\epsilon_\lambda^\mu(k)$; for each outgoing photon, $\epsilon_\lambda^{*\mu}(k)$;
   - for each incoming electron, $u_r(k)$; for each outgoing electron, $\overline{u}_s(p)$;
   - for each incoming positron, $\overline{v}_{\overline{r}}(\overline{k})$; for each outgoing positron, $v_{\overline{s}}(\overline{p})$;
   - for each vertex, $ie_0\gamma^\mu$; for each internal photon, $G_F(p)$; for each internal fermion, $S_F(p)$.

10. Spinor indices are contracted by starting at one end of a fermion line: specifically, the end that has the arrow pointing away from the vertex. The factor associated with the external line is either $\bar{u}$ or $\bar{v}$. Go along the complete fermion line, following the arrows backwards, and write down (in order from left to right) the factors associated with the vertices and propagators that you encounter. The last factor is either a u or v. Repeat this procedure for the other fermion lines, if any. The vector index on each vertex is contracted with the vector index on either the photon propagator (if the attached photon line is internal) or the photon polarization vector (if the attached photon line is external).

11. The overall sign of a tree diagram is determined by drawing all contributing diagrams in a standard form: all fermion lines horizontal, with their arrows pointing from left to right, and with the left endpoints labeled in the same fixed order (from top to bottom); if the ordering of the labels on the right endpoints of the fermion lines in a given diagram is an even (odd) permutation of an arbitrarily chosen fixed ordering, then the sign of that diagram is positive (negative)

12. Each closed fermion loop contributes an extra minus sign.

13. Value of $i\mathcal{M}$ is given by a sum over the values of the contributing diagrams.

14. $\langle f|S|i\rangle = (Z_2)^{\frac{n_f}{2}}(Z_3)^{\frac{n_p}{2}} i\mathcal{M}\delta(\sum p_f - \sum p_i)$

25.7.2 Ward Takahashi identity (3)

Suppose the invariant matrix element for a process is $\mathcal{M}$, if we replace the polarization state vector ϵ_λ^μ (or $\epsilon_\lambda^{*\mu}$) of one incoming (or outgoing) photon with its momentum vector k^μ, we have

$$k^\mu \mathcal{M}_\mu = 0$$

☞ **Proof:** Without losing generality, we can consider a physical process with a single incoming and outgoing fermion lines respectively. So, the ward identities states that

$$-ik_\mu F^\mu(k;p,q) = ie_0\left[F_0(p+k,q) - F_0(p,q-k)\right]$$

Here, F keeps the external fermion legs but cuts external photon lines. According to the LSZ reduction formula, from each diagram a contribution to an S matrix element by taking the coefficient of the product of poles

$$\left(\frac{-i}{\not{p}+m}\right)\left(\frac{-i}{\not{q}+m}\right)$$

But the terms on the right hand side contain one of these poles, but neither contains both poles. So they contribute nothing to S-matrix. So, we can have

$$k^\mu \mathcal{M}_\mu = 0$$

□

We also note that when calculating invariant matrix element, the main difference between different gauge are photon propagator. In Coulomb gauge, we have

$$G_F(p)_{\mu\nu} \equiv \begin{cases} \frac{i}{|\boldsymbol{p}|^2} & \mu,\nu = 0 \\ \frac{-i}{p^2-i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\boldsymbol{p}|^2}\right) & \mu = i \neq 0, \nu = j \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

In Lorentz gauge, we have

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2-i\epsilon)^2}$$

We now argue that the lead to the same $\mathcal{M}$ element. For a general process, it can be represented as follows.

Figure 25.6: Feynman diagram representation of a QED process

The value of the diagram is

$$\mathcal{M}_1^\mu G_F(k)_{\mu\nu}\mathcal{M}_2^\nu$$

and $k_\mu \mathcal{M}_1^\mu = 0$, $k_\nu \mathcal{M}_2^\nu = 0$. So the factor ξ in Lorentz gauge is irrelevant to the value of $\mathcal{M}$. As for Coulomb gauge, denote $\mathcal{M}_1^\mu$ as α^μ, $\mathcal{M}_2^\mu$ as β^μ, so

$$\alpha^\mu G_F(k)_{\mu\nu}\beta^\nu = i\left(-\frac{\boldsymbol{\alpha}\cdot\boldsymbol{\beta}}{k^2} + \frac{(\boldsymbol{\alpha}\cdot\boldsymbol{k})(\boldsymbol{\beta}\cdot\boldsymbol{k})}{k^2\boldsymbol{k}^2} + \frac{\alpha^0\beta^0}{\boldsymbol{k}^2}\right)$$

Using the fact that $\boldsymbol{\alpha}\cdot\boldsymbol{k} + \alpha^0 k_0 = 0$, we can verify that

$$\alpha^\mu G_F(k)_{\mu\nu}\beta^\nu = \alpha^\mu\left(-\frac{i\eta_{\mu\nu}}{k^2}\right)\beta^\nu$$

So, the invariant matrix element is gauge invariant.

25.8 Renormalization

25.8.1 Renormalized quantum electrodynamics

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi + e_0 j^\mu A_\mu,$$

Suppose the number of external photons is N_γ, the number of external fermions is N_e, the number of vertex is V. The superficial divergence is

$$D = 4 - N_\gamma - \frac{3}{2}N_e$$

Recall that C is a symmetry of QED, so $C|\Omega\rangle = |\Omega\rangle$. But j^μ changes sign under charge conjugation, $Cj^\mu(x)C^\dagger = -j^\mu(x)$, so its vacuum expectation value must vanish:

$$\langle\Omega|Tj^\mu(x)|\Omega\rangle = \langle\Omega|TC^\dagger Cj^\mu(x)C^\dagger C|\Omega\rangle = -\langle\Omega|Tj^\mu(x)|\Omega\rangle = 0$$

So, the amplitude with $N_\gamma = 1, N_e = 0$ will vanish. Similarly, we can verify that the diagram with $N_\gamma = 3, N_e = 0$ will also vanish.
As for the amplitude with $N_\gamma = 4$, the superficial divergence $D = 0$. So the only divergence must be of the form $\log\Lambda$. But the ward identity implies that

$$k^\mu \mathcal{M}_{\mu\nu\sigma\rho} = 0$$

So, the divergent terms must vanish.
If we neglect the vacuum term with $N_\gamma = 0, N_e = 0$, there are only three divergent amplitude terms left.
We need four counter terms to eliminate all the divergence. The Lagrangian can be written as

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_{ct}$$

Here,

$$\mathcal{L}_1 = -\frac{1}{4}F_r^{\mu\nu}F_{r\mu\nu} + i\overline{\Psi}_r\gamma^\mu\partial_\mu\Psi_r - m\overline{\Psi}_r\Psi_r + e\overline{\Psi}_r\gamma^\mu\Psi_r A_{r\mu}$$

Figure 25.7: Feynman diagram representation of divergent amplitude in QED

$$\mathcal{L}_{ct} = -\frac{1}{4}\delta_3 F_r^{\mu\nu} F_{r\mu\nu} + i\delta_2 \overline{\Psi}_r \gamma^\mu \partial_\mu \Psi_r - \delta_m \overline{\Psi}_r \Psi_r + e\delta_1 \overline{\Psi}_r \gamma^\mu \Psi_r A_{r\mu}$$

$$A = \sqrt{Z_3} A_r \quad \Phi = \sqrt{Z_2}\Phi_r \quad \delta_3 = Z_3 - 1 \quad \delta_2 = Z_2 - 1$$

$$\delta_m = Z_2 m_0 - m \quad \delta_1 = Z_1 - 1 = (e_0/e) Z_2 Z_3^{\frac{1}{2}} - 1$$

The Feynman rules of counter terms are:

$$-\frac{1}{4}\delta_3 F_r^{\mu\nu} F_{r\mu\nu} \qquad -i(\eta^{\mu\nu}q^2 - q^\mu q^\nu)\delta_3$$

$$i\overline{\Psi}_r(\delta_2 \not{\partial} - \delta_m)\Psi_r \qquad -i(\delta_2 \not{p} + \delta_m)$$

$$e\delta_1 \overline{\Psi}_r \gamma^\mu \Psi_r A_{r\mu} \qquad ie\gamma^\mu \delta_1$$

We also denote the renormalized 1PI component of exact propagator of photon as $i(\eta^{\mu\nu}q^2 - q^\mu q^\nu)\Pi_r(q^2)$, the renormalized 1PI component of exact propagator of fermion as $-i\Sigma_r(\not{p})$, the renormalized exact amputated photon-fermion-antifermion vertex as $ie\Gamma_r^\mu(p, p')$.
So, renormalized exact propagator of photon is

$$\mathcal{G}_r(q)_{\mu\nu} = \frac{-i}{q^2(1 - \Pi_r(q^2))} P^T_{\mu\nu}$$

renormalized exact propagator of fermion is

$$\mathcal{S}_r(p) = \frac{-i}{\not{p} + m + \Sigma_r(\not{p})}$$

The on-shell renormalization conditions are

$$\begin{aligned}
\Sigma_r(\not{p} = -m\gamma^0) &= 0 \\
\frac{d}{d\not{p}}\Sigma_r(\not{p})\Big|_{\not{p}=-m\gamma^0} &= 0 \\
\Pi_r(q^2 = 0) &= 0 \\
ie\Gamma_r^\mu(p = p', p^2 = -m^2) &= ie\gamma^\mu
\end{aligned}$$

Recall the ward identity, we have

$$\begin{aligned} & ieZ_2\partial_\mu\langle\Omega|Tj_r^\mu\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega\rangle \\ = & -ie\delta(x-x_1)\langle\Omega|T\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega\rangle + ie\delta(x-x_2)\langle\Omega|T\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega\rangle \end{aligned}$$

In momentum space, we have

$$-k_\mu Z_2 Z_1^{-1}\mathcal{S}_r(p+k)[ie\Gamma_r^\mu(p+k,p)]\mathcal{S}_r(p)\frac{1}{1-\Pi_r(k^2)} = e(\mathcal{S}_r(p+k)-\mathcal{S}_r(p))$$

So,

$$Z_2 Z_1^{-1}k_\mu[\Gamma_r^\mu(p+k,p)]\frac{1}{1-\Pi_r(k^2)} = \not{k} + \Sigma_r(\not{k}+\not{p}) - \Sigma_r(\not{p})$$

Since Γ_r, Π_r and Σ_r are all finite by renormalization, so Z_1/Z_2 must be finite. In $\overline{MS}$ renormalization scheme, we immediately get

$$Z_1 = Z_2$$

In on-shell renormalization scheme, taking the limit of $k \to 0$ and $p^2 = -m^2$, we can also get that

$$Z_1 = Z_2$$

So, we know

$$e = \sqrt{Z_3}e_0$$

This means that the relation between the bare and renormalized electric charge depends only on the photon field strength renormalization, not on quantities particular to the fermions, leading to a universal electric charge which has the same value for all species.
In the following subsection, we would omit the subscript r unless it is necessary to emphasis the difference of bare field and renormalized fields.

25.8.2 One loop structure of QED

Photon propagator

Figure 25.8: The one-loop and counterterm corrections to the photon propagator

$$\Pi(k^2) = -\frac{e^2}{\pi^2}\int_0^1 dx x(1-x)\left[\frac{1}{\epsilon} - \frac{1}{2}\ln(\frac{D}{\mu^2})\right] - \delta_3 + O(e^4)$$

where $D = x(1-x)k^2 + m^2 - i\epsilon$ and $\mu^2 = 4\pi e^{-\gamma}\tilde{\mu}^2$. Impose the OS renormalization condition $\Pi(0) = 0$, we have

$$\delta_3 = -\frac{e^2}{6\pi^2}\left[\frac{1}{\epsilon} - \ln(\frac{m}{\mu})\right] + O(e^4)$$

$$\Pi(k^2) = \frac{e^2}{2\pi^2}\int_0^1 dx x(1-x)\ln(\frac{D}{m^2}) + O(e^4)$$

Fermion propagator

The exact renormalized fermion propagator in OS renormalization can be written in Lehmann–Kallen form as

$$i\mathcal{S}(\not{p}) = \frac{1}{\not{p}+m-i\epsilon} + \int_{m_{th}}^{\infty} ds \frac{\rho_\Psi(s)}{\not{p}+\sqrt{s}-i\epsilon}$$

We see that the first term has a pole at $\not{p} = -m$ with residue one. This residue corresponds to the field normalization that is needed for the validity of the LSZ formula.

There is a problem, however: in quantum electrodynamics, the threshold mass m_{th} is m, corresponding to the contribution of a fermion and a zero energy photon. Thus the second term has a branch point at $\not{p} = -m$. The pole in the first term is therefore not isolated, and its residue is ill defined.

This is a reflection of an underlying infrared divergence, associated with the massless photon. To deal with it, we must impose an infrared cutoff that moves the branch point away from the pole. The most direct method is to change the denominator of the photon propagator from k^2 to $k^2 + m_\gamma^2$, where m_γ is a fictitious photon mass. Ultimately, we must deal with this issue by computing cross sections that take into account detector inefficiencies. In quantum electrodynamics, we must specify the lowest photon energy ω_{min} that can be detected. Only after computing cross sections with extra undetectable photons, and then summing over them, is it safe to take the limit $m_\gamma \to 0$.

Figure 25.9: The one-loop and counterterm corrections to the fermion propagator

$$\Sigma(\not{p}) = \frac{e^2}{8\pi^2}\int_0^1 dx\left((2-\epsilon)(1-x)\not{p} + (4-\epsilon)m\right)\left[\frac{1}{\epsilon} - \frac{1}{2}\ln(\frac{D}{\mu^2})\right] + \delta_2\not{p} + \delta_m + O(e^4)$$

where $D = x(1-x)p^2 + xm^2 + (1-x)m_\gamma^2$. The fitness of $\Sigma(\not{p})$ requires that

$$\delta_2 = -\frac{e^2}{8\pi^2}\left(\frac{1}{\epsilon} + \text{ finite }\right) + O(e^4)$$

$$\delta_m/m = -\frac{e^2}{2\pi^2}\left(\frac{1}{\epsilon} + \text{ finite }\right) + O(e^4)$$

Impose the OS renormalization condition $\Sigma(-m) = 0$ and $\Sigma'(-m) = 0$, we have

$$\Sigma(\not{p}) = -\frac{e^2}{8\pi^2}\int_0^1 dx\left((1-x)\not{p} + 2m\right)\ln(\frac{D}{D_0}) + \kappa_2(\not{p}+m) + O(e^4)$$

where $D_0 = x^2m^2 + (1-x)m_\gamma^2$ and $\kappa_2 = -2\ln(m/m_\gamma) + 1$.

Figure 25.10: The one-loop correction to the photon-fermion-fermion vertex

Vertex

$$\Gamma^\mu(p,p') = (1+\delta_1)\gamma^\mu + \frac{e^2}{8\pi^2}\left[\left(\frac{1}{\epsilon} - 1 - \frac{1}{2}\int dF_3 \ln(D/\mu^2)\right)\gamma^\mu + \frac{1}{4}\int dF_3 \frac{\tilde{N}^\mu}{D}\right] + O(e^4)$$

where

$$\int dF_3 = 2\int_0^1 dx_1 dx_2 dx_3 \delta(x_1+x_2+x_3-1)$$

$$D = x_1(1-x_1)p^2 + x_2(1-x_2)p'^2 - 2x_1x_2 p\cdot p' + (x_1+x_2)m^2 + x_3 m_\gamma^2$$

$$\tilde{N}^\mu = \gamma_\nu[x_1\not{p} - (1-x_2)\not{p}' + m]\gamma^\mu[-(1-x_1)\not{p} + x_2\not{p}' + m]\gamma^\nu$$

Fitness of Γ^μ requires that

$$\delta_1 = -\frac{e^2}{8\pi^2}(\frac{1}{\epsilon} + \text{ finite }) + O(e^4)$$

Impose the OS renormalization condition $\Gamma_r^\mu(p=p', p^2=-m^2) = \gamma^\mu$, we have

$$\Gamma^\mu(p,p') = \gamma^\mu - \frac{e^2}{16\pi^2}\int dF_3\left[(\ln(D/D_0) + 2\kappa_1)\gamma^\mu - \frac{\tilde{N}^\mu}{2D}\right] + O(e^4)$$

where

$$D_0 = (1-x_3)^2 m^2 + x_3 m_\gamma^2 \quad \kappa_1 = -2\ln(m/m_\gamma) + \frac{5}{2}$$

Vertex function

Consider the process of electron-electron scattering In order to compute the contribution of

Figure 25.11: Diagrams for the exact electron-electron scattering amplitude. The vertices and photon propagator are exact; external lines stand for the usual u and $\bar{u}$ spinor factors, times the unit residue of the pole at $p^2 = -m^2$.

these diagrams, we must evaluate $\overline{u}_{s'}(\boldsymbol{p}')\Gamma^\mu(p',p)u_s(\boldsymbol{p})$ with $p^2 = p'^2 = -m^2$, but with $q^2 = (p-p')^2$ arbitrary. Finally,we have

$$\overline{u}_{s'}(\boldsymbol{p}')\Gamma^\mu(p',p)u_s(\boldsymbol{p}) = \overline{u}'_{s'}(\boldsymbol{p}')\left[F_1(q^2)\gamma^\mu - \frac{i}{m}F_2(q^2)S^{\mu\nu}q_\nu\right]u_s(\boldsymbol{p})$$

where we have defined the form factors

$$\begin{aligned}
F_1(q^2) &= 1 - \frac{e^2}{16\pi^2}\int dF_3 \ln\left(1 + \frac{x_1x_2q^2/m^2}{(1-x_3)^2}\right) \\
&- \frac{e^2}{16\pi^2}\int dF_3 \frac{1-4x_3+x_3^2}{(1-x_3)^2 + x_3m_\gamma^2/m^2} + \frac{(x_3+x_1x_2)q^2/m^2 - (1-4x_3+x_3^2)}{x_1x_2q^2/m^2 + (1-x_3)^2 + x_3m_\gamma^2/m^2} + O(e^4) \\
F_2(q^2) &= \frac{e^2}{8\pi^2}\int_0^1 dy \frac{1}{1+y(1-y)q^2/m^2} + O(e^4)
\end{aligned}$$

and

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$$

In on-shell renormalization scheme, we have $F_1(0) = 1$ and $F_2(0) = \frac{\alpha}{2\pi} + O(\alpha^2)$.

25.8.3 Renormalization group

$\overline{MS}$ renormalization scheme

$$\Pi(k^2) = \frac{e^2}{2\pi^2}\int_0^1 dx\, x(1-x)\ln\left(\frac{x(1-x)k^2+m^2}{\mu^2}\right) + O(e^4)$$

$$Z_3 = 1 + \delta_3 = 1 - \frac{e^2}{6\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$\Sigma(\not{p}) = -\frac{e^2}{8\pi^2}\int_0^1 dx \ln\left(\frac{x(1-x)p^2 + xm^2 + (1-x)m_\gamma^2}{\mu^2}\right)\left((1-x)\not{p} + 2m\right) + \frac{1}{2}\not{p} + m + O(e^4)$$

$$Z_2 = 1 + \delta_2 = 1 - \frac{e^2}{8\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$Z_m = 1 + \delta_m/m = 1 - \frac{e^2}{2\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$\Gamma^\mu(p,p') = \gamma^\mu + \frac{e^2}{8\pi^2}\left[-\left(1 + \frac{1}{2}\int dF_3 \ln(D/\mu^2)\right)\gamma^\mu + \frac{1}{4}\int dF_3 \frac{\tilde{N}^\mu}{D}\right] + O(e^4)$$

$$D = x_1(1-x_1)p^2 + x_2(1-x_2)p'^2 - 2x_1x_2p\cdot p' + (x_1+x_2)m^2 + x_3m_\gamma^2$$

$$Z_1 = 1 + \delta_1 = 1 - \frac{e^2}{8\pi^2}\frac{1}{\epsilon} + O(e^4)$$

Renormalization group

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{0\mu\nu}F_0^{\mu\nu} + \overline{\Psi}_0(i\partial\!\!\!/ - m_0)\Psi_0 + e_0\overline{\Psi}_0\gamma^\mu\Psi_0 A_{0\mu},$$

It can be written as

$$\mathcal{L} = -\frac{1}{4}Z_3F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(iZ_2\partial\!\!\!/ - Z_m m)\Psi + Z_1 e\overline{\Psi}\gamma^\mu\Psi A_\mu,$$

So,

$$\Psi_0 = Z_2^{1/2}\phi \quad m_0 = Z_2^{-1}Z_m m \quad A_0 = Z_3^{1/2}A \quad e_0 = Z_2^{-1}Z_1Z_3^{-1/2}e\tilde{\mu}^{\epsilon/2} = Z_3^{-1/2}e\tilde{\mu}^{\epsilon/2}$$

After using dimensional regularization, the infinities coming from loop integrals take the form of inverse powers of ϵ. In the $\overline{\text{MS}}$ renormalization scheme, we choose the Zs to cancel off these powers of $1/\epsilon$, and nothing more. Therefore the Zs can be written as

$$Z_1 = 1 + \sum_{n=1}^{\infty}\frac{a_n(\lambda)}{\epsilon^n} \quad Z_2 = 1 + \sum_{n=1}^{\infty}\frac{b_n(\lambda)}{\epsilon^n}$$
$$Z_3 = 1 + \sum_{n=1}^{\infty}\frac{c_n(\lambda)}{\epsilon^n} \quad Z_m = 1 + \sum_{n=1}^{\infty}\frac{d_n(\lambda)}{\epsilon^n}$$

In QED, $a_1 = -\frac{e^2}{8\pi^2} + O(e^4)$, $b_1 = -\frac{e^2}{8\pi^2} + O(e^4)$,$c_1 = -\frac{e^2}{6\pi^2} + O(e^4)$,$d_1 = -\frac{e^2}{2\pi^2} + O(e^4)$
Remember that bare fields and parameters must be independent of μ.Define

$$E(e,\epsilon) \equiv \ln(Z_3^{-1/2}) = \sum_{n=1}^{\infty}\frac{E_n(e)}{\epsilon^n}$$

We can calculate $E_1 = -\frac{1}{2}c_1 = \frac{e^2}{12\pi^2} + O(e^4)$. As $\ln e_0 = E + \ln e + \frac{\epsilon}{2}\ln\tilde{\mu}$. From the independence of e_0, we can derive

$$\left(1 + \frac{eE_1'}{\epsilon} + \cdots\right)\frac{de}{d\ln\mu} + \frac{\epsilon}{2}e = 0$$

In a renormalizable theory, we should have

$$\frac{de}{d\ln\mu} = -\frac{\epsilon}{2}e + \beta(e)$$

So

$$\beta(e) = \frac{e^2}{2}E_1'(\lambda)$$

In QED, we have

$$\beta(e) = \frac{e^3}{12\pi^2} + O(e^5)$$

Define

$$M(e,\epsilon) \equiv \ln(Z_m Z_2^{-1}) = \sum_{n=1}^{\infty}\frac{M_n(\lambda)}{\epsilon^n}$$

We can calculate $M_1 = d_1 - b_1 = -\frac{3e^2}{8\pi^2} + O(e^4)$. As $\ln m_0 = M + \ln m$, define the anomalous dimension of the mass

$$\gamma_m(e) \equiv \frac{1}{m}\frac{dm}{d\ln\mu}$$

From the independence of m_0, we can derive

$$\gamma_m(e) = \frac{e}{2}M_1'$$

In QED theory, we have

$$\gamma_m(e) = -\frac{3e^2}{8\pi^2} + O(e^4)$$

Expand $\ln Z_2$ as

$$\ln Z_2 = \frac{b_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the fermion field

$$\gamma_2(e) \equiv \frac{1}{2}\frac{d\ln Z_2}{d\ln\mu}$$

We can derive

$$\gamma_2(e) = -\frac{1}{4}eb_1'$$

In QED theory, we have

$$\gamma_2(e) = \frac{e^2}{16\pi^2} + O(e^4)$$

Expand $\ln Z_3$ as

$$\ln Z_3 = \frac{c_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the EM field

$$\gamma_3(e) \equiv \frac{1}{2}\frac{d\ln Z_3}{d\ln\mu}$$

We can derive

$$\gamma_3(e) = -\frac{1}{4}ec_1'$$

In QED theory, we have

$$\gamma_3(e) = \frac{e^2}{12\pi^2} + O(e^4)$$

25.8.4 The magnetic moment of the electron

Suppose there is a constant classical electromagnetic background field $A_\mu^{\rm cl}(\boldsymbol{x})$. Then we have a new term

$$e\gamma^\mu A_\mu^{\rm cl}(\boldsymbol{x})\overline{\Psi}\Psi$$

in the Lagrangian. In Feynman diagram, this term joins two fermion lines. The value of the vertex is

$$ie\gamma^\mu \int d^4x e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} e^{i(E'-E)t} A_\mu^{\rm cl}(\boldsymbol{x}) = ie\gamma^\mu \tilde{A}_\mu^{\rm cl}(\boldsymbol{q})(2\pi)\delta(E'-E)$$

where

$$\boldsymbol{q} = \boldsymbol{p}' - \boldsymbol{p} \quad \tilde{A}^{\mathrm{cl}}_{\mu}(\boldsymbol{q}) = \int d^3x e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} A^{\mathrm{cl}}_{\mu}(\boldsymbol{x})$$

If an electron scatters in the background field and its momentum changes from p to p', then the scattering matrix is

$$\langle \boldsymbol{p}'|S|\boldsymbol{p}\rangle = ie\overline{u}_{s'}(\boldsymbol{p}')\Gamma^{\mu}(p',p)u_s(\boldsymbol{p})\tilde{A}^{\mathrm{cl}}_{\mu}(\boldsymbol{q})(2\pi)\delta(E'-E)$$

where

$$F_1(q^2)\gamma^{\mu} - \frac{i}{m}F_2(q^2)S^{\mu\nu}q_{\nu}$$

If there is only magnetic field, we have

$$A^{\mathrm{cl}}_{\mu}(x) = (0, \boldsymbol{A}^{\mathrm{cl}}_i) \quad \boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}^{\mathrm{cl}}$$

In momentum space, we have

$$\tilde{\boldsymbol{B}}^{\mathrm{cl}}(\boldsymbol{q}) = i\boldsymbol{q}\times\tilde{\boldsymbol{A}}^{\mathrm{cl}}(\boldsymbol{q})$$

Recall that

$$u(\boldsymbol{p}) = D(\Lambda)u(\boldsymbol{0})$$

where Λ boost the the four momentum $(m, \boldsymbol{0})$ to $(E, \boldsymbol{p})$. In the non-relativistic limit, we keep only the first order of momentum. So the rapidity for the boost is $\boldsymbol{\beta} = \frac{\boldsymbol{p}}{m}$. And so

$$D(\Lambda) = I + \frac{i}{2}\delta\omega_{\mu\nu}S^{\mu\nu} = I + i\beta^i S^{i0} = I - \frac{1}{2m}\begin{bmatrix} p^i\sigma^i & 0 \\ 0 & -p^i\sigma^i \end{bmatrix}$$

So,

$$u(\boldsymbol{p}) = \sqrt{m}\begin{pmatrix} (1-\boldsymbol{p}\cdot\boldsymbol{\sigma}/2m)\xi \\ (1+\boldsymbol{p}\cdot\boldsymbol{\sigma}/2m)\xi \end{pmatrix}$$

where ξ is the state of electron spin. We can get

$$\bar{u}(\boldsymbol{p}')\gamma^i u(\boldsymbol{p}) = 2m\xi'^{\dagger}\left(\frac{\boldsymbol{p}'\cdot\boldsymbol{\sigma}}{2m}\sigma^i + \sigma^i\frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}}{2m}\right)\xi = 2m\xi'^{\dagger}\left(\frac{p_i+p'_i}{2m} - \frac{i}{2m}\epsilon_{ijk}q_j\sigma_k\right)\xi$$

The first term in the bracket is spin-independent. It is the contribution of the operator $[\boldsymbol{p}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{p}]$ in the standard kinetic energy term of non-relativistic quantum mechanics. The second is the magnetic moment interaction we are seeking. So we retain only the latter term in the following discussion.
We can also get

$$\bar{u}(\boldsymbol{p}')\left(-\frac{i}{m}S^{i\nu}q_{\nu}\right)u(\boldsymbol{p}) = 2m\xi'^{\dagger}\left(-\frac{i}{2m}\epsilon_{ijk}q_j\sigma_k\right)\xi$$

Finally, we have

$$\begin{aligned}\langle \boldsymbol{p}'|S|\boldsymbol{p}\rangle &= 2ime\xi'^{\dagger}\left(-\frac{i}{2m}\epsilon_{ijk}q_j\sigma_k[F_1(0)+F_2(0)]\right)\xi\tilde{A}^{\mathrm{cl}}_i(\boldsymbol{q})(2\pi)\delta(E-E') \\ &= i(2m)\xi'^{\dagger}\left(\frac{e}{m}[F_1(0)+F_2(0)]\frac{\sigma_k}{2}\right)\xi\tilde{B}^{\mathrm{cl}}_k(\boldsymbol{q})(2\pi)\delta(E-E')\end{aligned}$$

Recall in quantum mechanics, the time dependent perturbation theory gives the transition matrix in first order as

$$\langle \boldsymbol{p}'|T|\boldsymbol{p}\rangle = -i\int_0^t dt e^{i(E'-E)t}\langle \boldsymbol{p}'|V(\boldsymbol{x})|\boldsymbol{p}\rangle \to 2\pi\delta(E'-E)V(\boldsymbol{q})$$

where $V(\boldsymbol{x})$ is perturbed potential we applied to free particles. Note that in quantum field theory, we normalize the momentum eigenvector as

$$\langle \boldsymbol{p}'|\boldsymbol{p}\rangle = (2\pi)^3 2E\delta(\boldsymbol{p}-\boldsymbol{p}')$$

rather than

$$\langle \boldsymbol{p}'|\boldsymbol{p}\rangle = (2\pi)^3\delta(\boldsymbol{p}-\boldsymbol{p}')$$

So there is a extra factor $2m$ in the equation we derived by field theory. By comparison, we conclude that

$$V(\boldsymbol{q}) = -\xi'^\dagger\left(\frac{e}{m}[F_1(0)+F_2(0)]\frac{\sigma_k}{2}\right)\xi\tilde{B}_k^{\text{cl}}(\boldsymbol{q})$$

So, the magnetic moment of electron is

$$\boldsymbol{\mu} = \frac{e}{2m}[2F_1(0)+2F_2(0)]\boldsymbol{S}$$

The g-factor is

$$g = 2F_0 + 2F_2(0) = 2 + \frac{\alpha}{2\pi}$$

25.8.5 Lamb shift

Let us consider the scattering of distinguishable fermions, for example electron and proton, in the non-relativistic limit.

Figure 25.12: Coulomb potential scattering

Up to tree level, we have

$$i\mathcal{M} = \bar{u}(k')(-ie\gamma^\mu)u(k)\frac{-i\eta_{\mu\nu}}{q^2}\bar{u}(p')(ie\gamma^\nu)u(p) = -i\frac{-e^2}{\boldsymbol{q}^2}2m\delta_{ss'}2m\delta_{rr'}$$

By comparison to Born approximation in quantum mechanics, we have

$$\tilde{V}(\boldsymbol{q}) = \frac{-e^2}{\boldsymbol{q}^2}$$

Transform the potential to position space, we can get

$$V(r) = \frac{-e^2}{4\pi r}$$

It is the Coulomb potential of electron and proton.

Next let us examine how $\Pi(q^2)$ modifies the electromagnetic interaction. We have the modified potential

$$\tilde{V}(\boldsymbol{q}) = \frac{-e^2}{\boldsymbol{q}^2(1-\Pi(q^2))}$$

where

$$\Pi(q^2) = -\frac{2\alpha}{\pi}\int_0^1 dx x(1-x)\ln\left(\frac{m^2}{m^2+x(1-x)q^2}\right)$$

In the limit of $|q^2| < m^2$, we have

$$\Pi(q^2) = \frac{\alpha}{15\pi}\frac{\boldsymbol{q}^2}{m^2} \quad \tilde{V}(\boldsymbol{q}) = -\frac{4\pi\alpha}{\boldsymbol{q}^2} - \frac{4\alpha^2}{15m^2}$$

Transform the potential to position space, we can get

$$V(r) = -\frac{\alpha}{r} - \frac{4\alpha^2}{15m^2}\delta(\boldsymbol{x})$$

The correction term indicates that the electromagnetic force becomes much stronger at small distances. This effect can be measured in the hydrogen atom, where the energy levels are shifted by

$$\Delta E = -\frac{4\alpha^2}{15m^2}|\psi(0)|^2$$

The wave function is non-zero at the origin only for s -wave states. For the 2S state, the shift is about -1.123×10^{-7} eV. Note that the energy level of fine structure is E_{nj}. This modified potential causes a split for degenerate levels of different l. This is a (small) part of the Lamb shift splitting.
A more precise correction is given by Uehling potential

$$\delta V(r) = -\frac{\alpha^2}{4\sqrt{\pi}r}\frac{e^{-2mr}}{(mr)^{3/2}}$$

Thus the range of the correction term is roughly the electron Compton wavelength, $\frac{1}{m}$. Since hydrogen wave functions are nearly constant on this scale, the delta function was a good approximation.
We can interpret the correction as being due to screening. At $r > \frac{1}{m}$, virtual e^+e^- pairs make the vacuum a dielectric medium in which the apparent charge is less than the true charge. At smaller distances we begin to penetrate the polarization cloud and see the bare charge. This phenomenon is known as vacuum polarization.

Chapter 26
Gauge Field

26.1 Nonabelian gauge theory

26.1.1 Nonabelian symmetries

Consider the theory of N real scalar fields ϕ_i

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - \frac{1}{2}m^2\phi_i\phi_i - \frac{1}{16}\lambda(\phi_i\phi_i)^2$$

This Lagrangian is clearly invariant under the $SO(N)$ transformation

$$\phi_i(x) \to R_{ij}\phi_j(x)$$

where R is an orthogonal matrix with a positive determinant: $R^T = R^{-1}$ and $\det R = +1$. Consider an infinitesimal $SO(N)$ transformation

$$R_{ij} = \delta_{ij} + \theta_{ij} + O(\theta^2)$$

Orthogonality of R_{ij} implies that θ_{ij} is real and antisymmetric. It is convenient to express θ_{ij} in terms of a basis set of hermitian matrices T^a_{ij}. The index a runs from 1 to $\frac{1}{2}N(N-1)$, the number of linearly independent, hermitian, antisymmetric, $N \times N$ matrices. Commonly, for $SO(N)$ group, we demand these matrices obey the normalization condition

$$\mathrm{Tr}(T^aT^b) = 2\delta^{ab}$$

In terms of them, we can write

$$\theta_{ij} = -i\theta^a T^a_{ij}$$

The T^as are the generator matrices of $SO(N)$. The product of any two $SO(N)$ transformations is another $SO(N)$ transformation; this implies that the commutator of any two generator matrices must be a linear combination of generator matrices,

$$[T^a, T^b] = if^{abc}T^c$$

The numerical factors f^{abc} are the structure coefficients of the group. If $f^{abc} = 0$, the group is abelian. Otherwise, it is nonabelian. Under our normalization condition, we have

$$f^{abc} = -\frac{i}{2}\mathrm{Tr}\left([T^a, T^b]T^c\right)$$

Using the cyclic property of the trace, we find that f^{abc} must be completely antisymmetric. Taking the complex conjugate of equation above, we find that f^{abc} must be real.

Example: The simplest nonabelian group is $SO(3)$. In this case, we can choose $T^a_{ij} = \epsilon^{aij}$. The commutation relations become

$$[T^a, T^b] = i\epsilon^{abc}T^c$$

Consider now the theory of N complex scalar fields ϕ_i

$$\mathcal{L} = -\partial_\mu\phi_i^\dagger\partial^\mu\phi_i - m^2\phi_i^\dagger\phi_i - \frac{1}{4}\lambda(\phi_i^\dagger\phi_i)^2$$

This Lagrangian is clearly invariant under the $U(N)$ transformation

$$\phi_i(x) \to U_{ij}\phi_j(x)$$

where U is a unitary matrix, $R^\dagger = R^{-1}$. We can write $U_{ij} = e^{-i\theta}\widetilde{U}_{ij}$, where θ is a real parameter and $\det\widetilde{U} = 1$. $\widetilde{U}_{ij}$ is called a special unitary matrix. Clearly the product of two special unitary matrices is another special unitary matrix; the $N \times N$ special unitary matrices form the group $SU(N)$. The group $U(N)$ is the direct product of the group $U(1)$ and the group $SU(N)$.
Consider an infinitesimal $SU(N)$ transformation

$$\widetilde{U}_{ij} = \delta_{ij} - i\theta^a T^a_{ij} + O(\theta^2)$$

where θ^a is a set of real, infinitesimal parameters. Unitarity of $\widetilde{U}$ implies that the generator matrices T are hermitian, and $\det\widetilde{U} = 1$ implies that each T is traceless. The index a runs from 1 to $N^2 - 1$, the number of linearly independent, hermitian, traceless, $N \times N$ matrices. Commonly, for $SU(N)$ group, we demand these matrices obey the normalization condition

$$\mathrm{Tr}(T^aT^b) = \frac{1}{2}\delta^{ab}$$

Example: For $SU(2)$, we can choose $T^a_{ij} = \frac{1}{2}\sigma^a_{ij}$. The commutation relations become

$$[T^a, T^b] = i\epsilon^{abc}T^c$$

26.1.2 Nonabelian gauge theory

Consider a Lagrangian with N scalar or spinor fields $\phi^i(x)$ that is invariant under a continuous $SU(N)$ symmetry,

$$\phi_i(x) = U_{ij}\phi_j(x)$$

It is called a global symmetry transformation, because the matrix U does not depend on the space-time label x.
If we want to generalize the symmetry of Lagrangian to local transformation

$$\phi_i(x) = U_{ij}(x)\phi_j(x)$$

terms with derivatives, such as $\partial^\mu \psi^\dagger \partial_\mu \phi_i$, will not remain invariant under local transformation. So we must include a traceless hermitian $N \times N$ gauge field $A_\mu(x)$, and promote ordinary derivatives ∂_μ to covariant derivatives $D_\mu = \partial_\mu - igA_\mu$ to ensure that

$$D_\mu \phi \to U D_\mu \phi$$

As a result, the gauge field must transform as

$$A_\mu(x) \to U(x) A_\mu(x) U^\dagger(x) + \frac{i}{g} U(x) \partial_\mu U^\dagger(x)$$

Replacing all ordinary derivatives in $\mathcal{L}$ with covariant derivatives renders $\mathcal{L}$ gauge invariant (assuming, of course, that $\mathcal{L}$ originally had a global $SU(N)$ symmetry).
We can write $U(x)$ in terms of the generator matrices as $\exp[-ig\Gamma(x)T^a]$. If the structure constant $f^{abc} \neq 0$, we have a nonabelian gauge theory.

We still need a kinetic term for $A_\mu(x)$. Let us define the field strength

$$F_{\mu\nu}(x) \equiv \frac{i}{g}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$$

We can verify that the field strength transform as

$$F_{\mu\nu}(x) \to U(x) F_{\mu\nu}(x) U^\dagger(x)$$

Therefore, a reasonable kinetic term is

$$\mathcal{L}_{\text{kin}} = -\frac{1}{2}\text{Tr}(F^{\mu\nu} F_{\mu\nu})$$

Since we have taken A_μ to be hermitian and traceless, we can expand it in terms of the generator matrices:

$$A_\mu(x) = A^a_\mu(x) T^a$$

Similarly, we have

$$F_{\mu\nu}(x) = F^a_{\mu\nu}(x) T^a$$

We can get

$$F^c_{\mu\nu} = \partial_\mu A^c_\nu - \partial_\nu A^c_\mu + g f^{abc} A^a_\mu A^b_\nu$$

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} F^{c\mu\nu} F_{c\mu\nu}$$

Everything we have just said about $SU(N)$ also goes through for $SO(N)$, with unitary replaced by orthogonal, and traceless replaced by antisymmetric. There is also another class of compact nonabelian groups called $Sp(2N)$, and five exceptional compact groups: $G(2)$, $F(4)$, $E(6)$, $E(7)$ and $E(8)$. Compact means that $\text{Tr}(T^a T^b)$ is a positive definite matrix. Nonabelian gauge theory must be based on a compact group, because otherwise some of the terms in $\mathcal{L}_{\text{kin}}$ would have the wrong sign, leading to a Hamiltonian that is unbounded below.

As a specific example, let us consider quantum chromodynamics, or QCD, which is based on

the gauge group $SU(3)$. There are several Dirac fields corresponding to quarks. Each quark comes in three colors; these are the values of the $SU(3)$ index. There are also six flavours: up, down, strange, charm, bottom, and top. Thus we consider the Dirac field $\Psi_{iI}(x)$, where i is the color index and I is the flavour index. The Lagrangian is

$$\mathcal{L} = i\overline{\Psi}_{iI}\not{D}_{ij}\Psi_{jI} - m_I\overline{\Psi}_I\Psi_I - \frac{1}{2}\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu})$$

The different quark flavours have different masses, ranging from a few MeV for the up and down quarks to 174 GeV for the top quark. The covariant derivative is

$$D_{\mu ij} = \delta_{ij}\partial_\mu - igA^a_\mu T^a_{ij}$$

The index a on A^a_μ runs from 1 to 8, and the corresponding massless spin-one particles are the eight gluons.

In a nonabelian gauge theory in general, we can consider scalar or spinor fields in different representations of the group. A representation of a compact nonabelian group is a set of finite-dimensional hermitian matrices T^a_R that obey the same commutation relations as the original generator matrices T^a. Given such a set of $D(R) \times D(R)$ matrices, and a field $\phi(x)$ with $D(R)$ components, we can write its covariant derivative as $D_\mu = \partial_\mu - igA^a_\mu T^a_R$. Under a gauge transformation, $\phi(x) \to U_R(x)\phi(x)$. The theory will be gauge invariant provided that

$$A^c_\mu \to A^c_\mu + g\theta^a A^b_\mu f^{abc} - \partial_\mu\theta^c$$

under infinitesimal transformation, which is independent of representation.

26.1.3 Group representations

Given the structure coefficients f^{abc} of a compact nonabelian group, a representation of that group is specified by a set of $D(R) \times D(R)$ traceless hermitian matrices T^a_R that obey the same commutation relations as the original generators matrices T^a. The number $D(R)$ is the dimension of the representation. The original T^as correspond to the fundamental or defining representation.
If T^a_R is a representation of the group, then we can verify that $-(T^a_R)^*$ is also a representation.

- If $T^a_R = -(T^a_R)$, or if we can find a unitary transformation $T^a_R \to U^{-1}T^a_R U$ that makes $-(T^a_R)^* = T^a_R$ for every a, then the representation R is real.
- If such a unitary transformation does not exist, but we can find a unitary matrix $V \neq I$ such that $V^{-1}T^a_R V = -(T^a_R)^*$ for every a, then the representation R is pseudoreal.
- If such a unitary matrix also does not exist, then the representation R is complex.
- One way to prove that a representation is complex is to show that at least one generator matrix T^a_R (or a real linear combination of them) has eigenvalues that do not come in plus-minus pairs.

Example:

- The fundamental representation of $SO(N)$ is real
- The fundamental representation for $SU(2)$ is pseudoreal
- The fundamental representation for $SU(N)$ with $N \geq 3$ is complex

We note that

$$\mathrm{Tr}T^e\left([[T^a,T^b],T^c]+[[T^c,T^a],T^b]+[[T^b,T^c],T^a]\right)=0$$

by which we can derive the Jacobian identity

$$f^{abd}f^{dce}+f^{bcd}f^{dae}+f^{cad}f^{dbe}=0$$

It can be rearranged as

$$(-if^{abd})(-if^{cde})-(-if^{cbd})(-if^{ade})=if^{acd}(-if^{dbe})$$

Define

$$(T_A^a)^{bc}\equiv -if^{abc}$$

we can get a new representation of the group. It is called adjoint representation. The dimension of adjoint representation is equal to the number of the generators. And it is easy to show that adjoint representation is real.

It can be demonstrated that once the fundamental representation is normalized so that $\mathrm{Tr}T^aT^b$ is proportion to the δ^{ab}, so is it with any irreducible representation of the group. Then we can define the index of a representation $T(R)$ as

$$Tr(T_R^aT_R^b)=T(R)\delta^{ab}$$

We can verify that the matrix $T_R^aT_R^a$ commutes with every generator, and so must be a number times the identity matrix. We define quadratic Casimir $C(R)$ by

$$T_R^aT_R^a=C^(R)I$$

It is easy to show that

$$T(R)D(A)=C(R)D(R)$$

For $SU(N)$, the standard normalization for fundamental representation is $T(N)=\frac{1}{2}$, so we have $C(N)=\frac{N^2-1}{2N}$. We will show later that $T(A)=C(A)=N$.
For $SO(N)$, the standard normalization for fundamental representation is $T(N)=2$, so we have $C(N)=N-1$. We will show later that $T(A)=C(A)=2N-4$.

A representation R is reducible if there is a unitary transformation $T_R^a\to U^{-1}T_R^aU$ that puts all the nonzero entries into the same diagonal blocks for each a; otherwise it is irreducible. Consider a reducible representation R whose generators can be put into two blocks, with the blocks forming the generators of the irreducible representations R_1 and R_2 . Then R is the direct sum representation $R=R_1\oplus R_2$, and we have

$$D(R_1\oplus R_2)=D(R_1)+D(R_2)$$

$$T(R_1 \oplus R_2) = T(R_1) + T(R_2)$$

Note: *We define the $T(R)$ of reducible representation by taking the trace of the product of two identical generator.*

Suppose we have a field ϕ_{iI} that carries two group indices, one for the representation R_1 and one for the representation R_2, denoted by i and I respectively. This field is in the direct product representation $R_1 \otimes R_2$. The corresponding generator matrix is

$$(T^a_{R_1 \otimes R_2})_{iI,jJ} = (T^a_{R_1})\delta_{IJ} + \delta_{ij}(T^a_{R_2})_{IJ}$$

We then have

$$D(R_1 \otimes R_2) = D(R_1)D(R_2)$$

$$T(R_1 \otimes R_2) = T(R_1)D(R_2) + T(R_2)D(R_1)$$

Consider a field ϕ in the complex representation R. We will adopt the convention that such a field carries a down index: ϕ_i, and its Hermitian conjugation carries a up index: $\phi^{\dagger i}$, where $i = 1, \cdots, D(R)$.
Since

$$\phi_i \to \phi_i - i\theta_a (T^a_R)_i{}^j \phi_j$$

we have

$$\phi^{\dagger i} \to \phi^{\dagger j} - i\theta_a(-T^{a*}_R)_i{}^j \phi^{\dagger j} = \phi^{\dagger j} - i\theta_a(-T^a_R)_j{}^i \phi^{\dagger j} \equiv \phi^{\dagger j} - i\theta_a (T^a_{\overline{R}})^i{}_j \phi^{\dagger j}$$

i.e.

$$(T^a_{\overline{R}})^i{}_j = (-T^a_R)_j{}^i$$

Consider the Kronecker delta symbol with one index down and one up: $\delta_i{}^j$. Under a group transformation, we have

$$\delta_i{}^j \to (1 + i\theta^a T^a_R)_i{}^k (1 + i\theta^a T^a_{\overline{R}})^j{}_l \delta_k{}^l = \delta_i{}^j + O(\theta^2)$$

$\delta_i{}^j$ is an invariant symbol of the group. This existence of this invariant symbol, which carries one index for R and one for $\overline{R}$, tells us that the product of the representations R and $\overline{R}$ must contain the singlet representation 1, specified by $T^a_1 = 0$. We therefore can write

$$R \otimes \overline{R} = 1 \oplus \cdots$$

We can further verify that the generator matrix $(T^a_R)_i{}^j$, which carries one index for R, one for $\overline{R}$, and one for the adjoint representation A, is also an invariant symbol. This implies that

$$R \times \overline{R} \otimes A = 1 \oplus \cdots$$

If we now multiply both sides by A, and use $A \otimes A = 1 \oplus \cdots$ (note A is real), we find $R \times \overline{R} = A \oplus \cdots$. Finally, we have

$$R \otimes \overline{R} = 1 \oplus A \oplus \cdots$$

So the product of a representation with its complex conjugate is always reducible into a sum that includes at least the singlet and adjoint representations. For the fundamental representation of $SU(N)$, we have

$$N \otimes \overline{N} = 1 \oplus A$$

by dimension analysis. We can derive $T(A) = N$ from this equation.

Consider now a real representation R, we have

$$R \times R \otimes A = 1 \oplus \cdots$$

The singlet on the right-hand side implies the existence of an invariant symbol with two R indices; this symbol is the Kronecker delta δ_{ij} . It is invariant because

$$\delta_{ij} \to (1 + i\theta^a T^a_R)_i{}^k (1 + i\theta^a T^a_{\overline{R}})_j{}^l \delta_{kl} = \delta_{ij} - i\theta^a[(T^a_R)_{ij} + (T^a_R)_{ji}] + O(\theta^2)$$

The term in square brackets vanishes by hermiticity and $(T^a_{\overline{R}})^i{}_j = (-T^a_R)_j{}^i$. The fact that $\delta_{ij} = \delta_{ji}$ implies that the singlet on the right-hand side of the equation above appears in the symmetric part of this product of two identical representations.
The fundamental representation of $SO(N)$ is real, and we have

$$N \times N = 1_S \oplus A_A \oplus S_S$$

The representation S corresponds to a field with a symmetric traceless pair of fundamental indices.

Consider now a pseudoreal representation R. Since R is equivalent to its complex conjugate, up to a change of basis, $R \times R \otimes A = 1 \oplus \cdots$ still holds. However, we cannot identify δ_{ij} as the corresponding invariant symbol, because then R would have to be real, rather than pseudoreal. From the perspective of the direct product, the only alternative is to have the singlet appear in the antisymmetric part of the product, rather than the symmetric part. The corresponding invariant symbol must then be antisymmetric on exchange of its two R indices. An example is the fundamental representation of $SU(2)$, which is discussed in spinor field theory.

The structure constants f^{abc} are another invariant symbol. This follows from $(T^a_A)^{bc} = -if^{abc}$, since we have seen that generator matrices are invariant.
Alternatively, given the generator matrices in a representation R, we can write

$$T(R)f^{abc} = -i\mathrm{Tr}(T^a_R[T^b_R, T^c_R])$$

Since the right-hand side is invariant, the left-hand side must be as well.
If we use an anticommutator in place of the commutator, we get another invariant symbol,

$$A(R)d^{abc} = \frac{1}{2}\mathrm{Tr}(T^a_R\{T^b_R, T^c_R\})$$

where $A(R)$ is the anomaly coefficient of the representation. The cyclic property of the trace implies that d^{abc} is symmetric on exchange of any pair of indices. Using $(T^a_{\overline{R}})^i{}_j = (-T^a_R)_j{}^i$, we can see that

$$A(\overline{R}) = -A(R)$$

Thus, if R is real or pseudoreal, $A(R) = 0$. We also have

$$A(R_1 \oplus R_2) = A(R_1) + A(R_2)$$

$$A(R_1 \otimes R_2) = A(R_1)D(R_2) + A(R_2)D(R_1)$$

We normalize the anomaly coefficient so that it equals one for the smallest complex representation. In particular, for $SU(3)$ with $N \geq 3$, the smallest complex representation is the fundamental, and $A(N) = 1$. For $SU(2)$, all representations are real or pseudoreal, and $A(R) = 0$ for all of them.

26.2 Quantization of nonabelian gauge theory

26.2.1 The path integral for nonabelian gauge theory

We wish to evaluate the path integral for nonabelian gauge theory

$$Z[0] = \int \mathcal{D}A e^{iS_{YM}}$$

$$S_{YM} = \int d^4x \left[-\frac{1}{4} F^{a\mu\nu} F^a_{\mu\nu} \right]$$

Recall under an infinitesimal gauge transformation, we have

$$A^a_\mu \to A^a_\mu - (\delta^{ac}\partial_\mu - igA^b_\mu (T^b_A)^{ac})\theta^c$$

We can define a covariant derivative for adjoint representation

$$D^{ac}_\mu \equiv \delta^{ac}\partial_\mu - igA^b_\mu (T^b_A)^{ac}$$

so we have

$$A^a_\mu \to A^a_\mu - D^{ac}_\mu \theta^c$$

Similar to the case in electromagnetic field, we introduce a gauge fixing function

$$G^a(A(\theta)) \equiv \partial^\mu A^a_\mu(\theta) - \omega^a(x)$$

Now, we insert 1 in the path integral

$$1 = \int \mathcal{D}\theta \delta(G) \det\left(\frac{\delta G}{\delta \theta}\right)$$

Here,

$$\frac{\delta G^a(x)}{\delta \theta^b(y)} = -\partial^\mu D^{ab}_\mu(\theta)\delta^4(x-y)$$

A functional determinant can be written as a path integral over complex Grassmann variables. So let us introduce the complex Grassmann field $c^a(x)$, and its hermitian conjugate $\overline{c}^a(x)$. These fields are called Faddeev–Popov ghosts. Then we can write

$$\det\left(\frac{\delta G}{\delta \theta}\right) \propto \int \mathcal{D}c\mathcal{D}\overline{c} e^{iS_{gh}[A(\theta)]}$$

where

$$\mathcal{L}_{gh}[A(\theta)] = \overline{c}^a \partial^\mu D^{ab}_\mu(\theta) c^b = -\partial^\mu \overline{c}^a \partial_\mu c^a + g f^{abc} A^c_\mu(\theta) \partial^\mu \overline{c}^a c^b$$

We see that $c^a(x)$ has the standard kinetic term for a complex scalar field. The ghost field is also a Grassmann field, and so a closed loop of ghost lines in a Feynman diagram carries an extra factor of minus one. Since the particles associated with the ghost field do not in fact exist (and would violate the spin-statistics theorem if they did), it must be that the amplitude to produce them in any scattering process is zero.

Now change variables from A to $A(\theta)$. This is a simple shift, so $\mathcal{D}A = \mathcal{D}A(\theta)$. Also, by gauge invariance, $S_{YM}[A] = S_{YM}[A(\theta)]$. Since $A(\theta)$ is now just a dummy integration variable, we can rename it bake to A, so

$$Z[0] \propto \left(\int \mathcal{D}\theta\right) \int \mathcal{D}c\mathcal{D}\overline{c}\mathcal{D}A e^{iS_{YM}+iS_{gh}} \delta(\partial^\mu A^a_\mu - \omega^a(x))$$

Since the above equation is hold for any $\omega(x)$,so we have

$$\begin{aligned} Z[0] &\propto \int \mathcal{D}\omega \exp\left[-i\int d^4x \frac{\omega^a\omega^a}{2\xi}\right] \int \mathcal{D}c\mathcal{D}\overline{c}\mathcal{D}A e^{iS_{YM}+iS_{gh}} \delta(\partial^\mu A^a_\mu - \omega^a(x)) \\ &= \int \mathcal{D}c\mathcal{D}\overline{c}\mathcal{D}A e^{iS_{YM}+iS_{gh}+iS_{gf}} \end{aligned} \tag{26.1}$$

where

$$\mathcal{L}_{gf} = -\frac{1}{2\xi}\partial^\mu A^a_\mu \partial^\nu A^a_\nu$$

26.2.2 The Feynman rules for nonabelian gauge theory

$$\mathcal{L}_{YM} = -\frac{1}{2}\partial^\mu A^{a\nu}\partial_\mu A^a_\nu + \frac{1}{2}\partial^\mu A^{a\nu}\partial_\nu A^a_\mu - g f^{abe} A^{a\mu} A^{b\nu} \partial_\mu A^e_\nu - \frac{1}{4} g^2 f^{abe} f^{cde} A^{a\mu} A^{b\nu} A^c_\mu A^d_\nu$$

Add the gauge-fixing term for R_ξ gauge, and do some integrations-by-parts in the quadratic terms, we find

$$\mathcal{L}_{YM}+\mathcal{L}_{gf} = \frac{1}{2}A^{e\mu}(\eta_{\mu\nu}\partial^2 - (1-\frac{1}{\xi})\partial_\mu\partial_\nu)A^{e\nu} - g f^{abe} A^{a\mu} A^{b\nu}\partial_\mu A^e_\nu - \frac{1}{4} g^2 f^{abe} f^{cde} A^{a\mu} A^{b\nu} A^c_\mu A^d_\nu$$

So, the gluon propagator in R_ξ gauge is

$$G_F(k)^{ab}_{\mu\nu} = \frac{-i\delta^{ab}}{k^2 - i\epsilon}\left(\eta_{\mu\nu} - (1-\frac{1}{\xi})\frac{k_\mu k_\nu}{k^2}\right)$$

The three-gluon vertex factor is

$$\begin{aligned} iV^{abc}_{\mu\nu\rho}(p,q,r) &= i(-g f^{abc})(-i r_\mu g_{\nu\rho}) + [5 \text{ permutations of } (a,\mu,p),(b,\nu,q),(c,\rho,r)] \\ &= g f^{abc}[(q-r)_\mu g_{\nu\rho} + (r-p)_\nu g_{\rho\mu} + (p-q)_\rho g_{\mu\nu}] \end{aligned}$$

The four-gluon vertex factor is

$$\begin{aligned} iV^{abcd}_{\mu\nu\rho\sigma} &= -i g^2 f^{abe} f^{cde} g_{\mu\rho} g_{\nu\sigma} + [5 \text{ permutations of } (b,\nu),(c,\rho),(d,\sigma)] \\ &= -i g^2 [f^{abe} f^{cde}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}) + f^{ace} f^{dbe}(g_{\mu\sigma}g_{\rho\nu} - g_{\mu\nu}g_{\rho\sigma}) + f^{ade} f^{bce}(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\rho}g_{\sigma\nu})] \end{aligned}$$

Figure 26.1: The three-gluon and four-gluon vertices in nonabelian gauge theory

For loop calculations, we need to include the ghosts. The ghost Lagrangian is

$$\mathcal{L}_{gh} = -\partial^\mu \overline{c}^c \partial_\mu c^c + g f^{abc} A^a_\mu \partial^\mu \overline{c}^b c^c$$

The ghost propagator is

$$D_F(k) = \frac{-i\delta^{ab}}{k^2 - i\epsilon}$$

Because the ghosts are complex scalars, their propagators carry a charge arrow.

Figure 26.2: The ghost-ghost-gluon vertex in nonabelian gauge theory

The ghost–ghost–gluon vertex factor is

$$iV^{abc}_\mu(-iq_\mu) = g f^{abc} a_\mu$$

If we include a quark coupled to the gluons, we have the quark Lagrangian

$$\mathcal{L}_q = i\overline{\Psi}_i \not{D}_{ij} \Psi_j - m\overline{\Psi}_i \Psi_i = i\overline{\Psi}_i \not{\partial} \Psi_i - m\overline{\Psi}_i \Psi_i + g A^a_\mu \overline{\Psi}_i \gamma^\mu T^a_{ij} \Psi_j$$

Figure 26.3: The quark-quark-gluon vertex in nonabelian gauge theory

The quark propagator is

$$S_F(p)_{ij} = \frac{i(\not{p} - m)\delta_{ij}}{p^2 + m^2 - i\epsilon}$$

The quark–quark–gluon vertex factor is

$$iV_{ij}^{\mu a} = ig\gamma^\mu T_{ij}^a$$

26.3 Renormalization of nonabelian gauge theory

We now rescale the fields to the renormalized field strengths by extracting the factors Z_2, Z_3 and Z_2^c for the fermions, gauge bosons, and ghosts, and shift the coupling to the renormalized coupling g. The counterterm Lagrangian then takes the form

$$\begin{aligned}
\mathcal{L}_{ct} &= -\frac{1}{4}\delta_3 F^{a\mu\nu}F^a_{\mu\nu} + \overline{\Psi}(i\delta_2\not{\partial} - \delta_m)\Psi + \delta_2^c\overline{c}^a\partial^2 c^a \\
&+ g\delta_1 A^a_\mu\overline{\Psi}\gamma^\mu T^a\Psi - g\delta_1^{3g} f^{abe}A^{a\mu}A^{b\nu}\partial_\mu A^e_\nu - \frac{1}{4}g^2\delta_1^{4g} f^{abe}f^{cde}A^{a\mu}A^{b\nu}A^c_\mu A^d_\nu \\
&+ g\delta_1^c f^{abc}A^a_\mu\partial^\mu\overline{c}^b c^c
\end{aligned}$$

with the counterterms de ned by

$$\delta_2 = Z_2 - 1 \quad \delta_3 = Z_3 - 1 \quad \delta_2^c = Z_2^c - 1 \quad \delta_m = Z_2 m_0 - m$$

$$\delta_1 = \frac{g_0}{g}Z_2(Z_3)^{1/2} - 1 \quad \delta_1^{3g} = \frac{g_0}{g}(Z_3)^{3/2} - 1 \quad \delta_1^{4g} = \frac{g_0^2}{g^2}(Z_3)^2 - 1 \quad \delta_1^c = \frac{g_0}{g}Z_2^c(Z_3)^{1/2} - 1$$

Notice that these eight counterterms depend on five underlying parameters; thus, there are three relations among them. The situation is very similar to that for the scalar theories with spontaneously broken symmetry that we studied before. The underlying symmetry of the theory - local gauge invariance - implies relations among the divergent amplitudes of the theory and among the counterterms required to cancel them. In the present case, a set of five renormalization conditions uniquely specifies all of the counterterms in a way that removes all divergences from the theory. The rigorous proof will be omitted here.

Quark propagator

Figure 26.4: The one-loop and counterterm corrections to the quark propagator in quantum chromodynamics

In $\overline{MS}$ renormalization scheme and Feynman gauge, we have

$$Z_2 = 1 - C(R)\frac{g^2}{8\pi^2}\frac{1}{\epsilon} + O(g^4)$$

$$Z_m = 1 - C(R)\frac{g^2}{2\pi^2}\frac{1}{\epsilon} + O(g^4)$$

Quark-quark-gluon vertex

Figure 26.5: The one-loop corrections to the quark–quark–gluon vertex in quantum chromodynamics.

$$Z_1 = 1 - [C(R) + T(A)]\frac{g^2}{8\pi^2}\frac{1}{\epsilon} + O(g^4)$$

Gluon propagator

Figure 26.6: The one-loop and counterterm corrections to the gluon propagator in quantum chromodynamics.

$$Z_3 = 1 + \left[\frac{5}{3}T(A) - \frac{4}{3}n_F T(R)\right]\frac{g^2}{8\pi^2}\frac{1}{\epsilon} + O(g^4)$$

Beta function

We define

$$\alpha \equiv \frac{g^2}{4\pi}$$

Then we have

$$\alpha_0 = \frac{Z_1^2}{Z_2^2 Z_3}\alpha\tilde{\mu}^\epsilon$$

Let us write

$$\ln\left(Z_3^{-1}Z_2^{-2}Z_1^2\right) = \sum_{n=1}^{\infty}\frac{G_n(\alpha)}{\epsilon^n}$$

Then we have

$$\ln\alpha_0 = \sum_{n=1}^{\infty}\frac{G_n(\alpha)}{\epsilon^n} + \ln\alpha + \epsilon\ln\tilde{\mu}$$

And we can get

$$G_1(\alpha) = -\left[\frac{11}{3}T(A) - \frac{4}{3}n_F T(R)\right]\frac{\alpha}{2\pi} + O(\alpha^2)$$

Since the bare coefficients is independent of μ, we then can get

$$\beta(\alpha) = \alpha^2 G_1(\alpha) = -\left[\frac{11}{3}T(A) - \frac{4}{3}n_F T(R)\right]\frac{\alpha^2}{2\pi} + O(\alpha^3)$$

In quantum chromodynamics, the gauge group is $SU(3)$, and the quarks are in the fundamental representation. Thus $T(A) = 3$ and $T(R) = \frac{1}{2}$, and the factor in square brackets is $11 - \frac{2}{3}n_F$. So for $n_F \leq 16$, the beta function is negative: the gauge coupling in quantum chromodynamics gets weaker at high energies, and stronger at low energies.
This has dramatic physical consequences. Perturbation theory cannot serve as a reliable guide to the low-energy physics. And indeed, in nature we do not see isolated quarks or gluons. (Quarks, in particular, have fractional electric charges and would be easy to discover.) The appropriate conclusion is that color is confined : all finite-energy states are invariant under a global $SU(3)$ transformation. This has not yet been rigorously proven, but it is the only hypothesis that is consistent with all of the available theoretical and experimental information.

26.4 Chiral gauge theories and anomalies

26.4.1 Chiral gauge theories

Recall that a Dirac field Ψ can be written in terms of two left-handed Weyl ields χ and ξ as

$$\Psi = \begin{pmatrix}\chi \\ \xi^\dagger\end{pmatrix}$$

If Ψ is in a representation R of the gauge group, then χ and $\xi^\dagger$ must be as well. Equivalently, χ must be in the representation R, and ξ must be in the complex conjugate representation $\overline{R}$. Now suppose that we have a single left-handed Weyl field ψ in a complex representation R.

Such a gauge theory is automatically parity violating (because the right-handed hermitian conjugate of the left-handed Weyl field is in an inequivalent representation of the gauge group), and is said to be chiral. The Lagrangian is

$$\mathcal{L} = i\psi^\dagger \overline{\sigma}^\mu D_\mu \psi - \frac{1}{4} F^{a\mu\nu} F^a_{\mu\nu}$$

where $D_\mu = \partial_\mu - igA^a_\mu T^a_R$.
Since T^a_R is a hermitian matrix, $i\psi^\dagger \overline{\sigma}^\mu D_\mu \psi$ is hermitian (up to a total divergence). We cannot include a mass term for ψ, though, because $\psi\psi$ transforms as $R \otimes R$, and $R \otimes R$ does not contain a singlet if R is complex. Thus, $\psi\psi$ is not gauge invariant.

Now we focus on the simplest possible example: a $U(1)$ theory with a single Weyl field ψ with charge $+1$. The Lagrangian is

$$\mathcal{L} = i\psi^\dagger \overline{\sigma}^\mu (\partial_\mu - igA_\mu)\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

We note that

$$P_L \Psi = \begin{pmatrix} \psi \\ 0 \end{pmatrix}$$

Then we can write

$$\mathcal{L} = i\overline{\Psi}\gamma^\mu (\partial_\mu - igA_\mu) P_L \Psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

To better understand the physical consequences of the projection operator, consider the case of a free field. The mode expansion is

$$P_L \Psi(x) = \sum_{s=\pm} \int \tilde{dp} [b_s(\boldsymbol{p}) P_L u_s(\boldsymbol{p}) e^{ipx} + d^\dagger_s(\boldsymbol{p}) P_L v_s(\boldsymbol{p}) e^{-ipx}]$$

For a massless field, we have $P_L u_+(\boldsymbol{p}) = 0$ and $P_L v_-(\boldsymbol{p}) = 0$. Thus we can write as

$$P_L \Psi(x) = \int \tilde{dp} [b_-(\boldsymbol{p}) u_-(\boldsymbol{p}) e^{ipx} + d^\dagger_+(\boldsymbol{p}) v_+(\boldsymbol{p}) e^{-ipx}]$$

We can easily read the Feynman rules from the Lagrangian. In particular, the fermion propagator in momentum space is

$$\frac{iP_L \not{p}}{p^2 - i\epsilon}$$

and the fermion-fermion-photon vertex is

$$ig\gamma^\mu P_L$$

Now consider the correction to the photon propagator. Calculation shows that at the one-loop level, the contribution to $\Pi^{\mu\nu}(k)$ of a single charged Weyl field is half that of a Dirac field. This is physically reasonable, since a Dirac field is equivalent to two charged Weyl fields.
Nothing interesting happens in the one-loop corrections to the fermion propagator, or the fermion-fermion-photon vertex. There is simply an extra factor of P_L along the fermion line,

which can be moved to the far right. Except for this factor, the results exactly duplicate those of spinor electrodynamics.
All of this implies that a single Weyl field makes half the contribution of a Dirac field to the leading term in the beta function for the gauge coupling.

Next we turn to diagrams with three external photons, and no external fermions. In spinor electrodynamics, the fact that the vector potential is odd under charge conjugation implies that the sum of these diagrams must vanish. For the present case of a single Weyl field, there is no charge-conjugation symmetry, and so we must evaluate these diagrams.

Figure 26.7: One-loop contributions to the three-photon vertex

Detailed calculation shows that

$$r_\rho V^{\mu\nu\rho}(p,q,r) = 0$$

However,

$$p_\mu V^{\mu\nu\rho}(p,q,r) = \frac{ig^3}{8\pi^2}\epsilon^{\alpha\nu\beta\rho}p_\alpha q_\beta$$

$$q_\nu V^{\mu\nu\rho}(p,q,r) = \frac{ig^3}{8\pi^2}\epsilon^{\alpha\rho\beta\mu}q_\alpha p_\beta$$

We have therefore failed to construct a gauge-invariant $U(1)$ theory with a single charged Weyl field.

Consider now a $U(1)$ gauge theory with several left-handed Weyl fields ψ_i, with charges Q_i, so that the covariant derivative of ψ_i is $\partial_\mu - igQ_iA_\mu$. Then each of these fields circulates in the loop in figure above, and each vertex has an extra factor of Q_i. The right-hand sides of equation above are now multiplied by $\sum_i Q_i^3$. And if $\sum_i Q_i^3$ happens to be zero, then gauge invariance is restored.
The simplest possibility is to have the ψs come in pairs with equal and opposite charges. But there are other possibilities as well: for example, one field with charge $+2$ and eight with charge -1. Such a gauge theory is still chiral, but it is anomaly free.

All of this has a straightforward generalization to nonabelian gauge theories. Suppose we have a single Weyl field in a (possibly reducible) representation R of the gauge group. Then we must attach an extra factor of $\mathrm{Tr}(T_R^aT_R^bT_R^c)$ to the first diagram, and a factor of $\mathrm{Tr}(T_R^aT_R^cT_R^b)$ to the second; here the group indices a, b, c go along with the momenta p, q, r, respectively. Repeating our analysis shows that the diagrams with $P_L \to \frac{1}{2}$ come with an extra factor of $\frac{1}{2}\mathrm{Tr}([T_R^a, T_R^b]T_R^c)$; these contribute to the renormalization of the tree-level three-gluon vertex.

Diagrams with $P_L \to -\frac{1}{2}\gamma_5$ come with an extra factor of

$$\frac{1}{2}\mathrm{Tr}(\{T_R^a, T_R^b\}T_R^c) = A(R)d^{abc}$$

Here d^{abc} is a completely symmetric tensor that is independent of the representation, and $A(R)$ is the anomaly coefficient of R. In order for this theory to exist, we must have $A(R) = 0$. Since $A(R) + A(\overline{R}) = 0$; thus a theory whose left-handed Weyl fields come in $R \oplus \overline{R}$ pairs is automatically anomaly free (as is one whose Weyl fields are all in real representations). Otherwise, we must arrange the cancellation by hand.
For $SU(2)$ and $SO(N)$, $N \neq 2, 6$ all representations have $A(R) = 0$. For $SU(N)$ with $N \geq 3$, the fundamental representation has $A(N) = 1$, and most complex $SU(N)$ representations R have $A(R) \neq 0$. So the cancellation is nontrivial.

26.4.2 Anomalies in global symmetries

Let us consider electrodynamics with a massless Dirac field. The Lagrangian is

$$\mathcal{L} = i\overline{\Psi}\not{D}\Psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}$$

We can write Ψ in terms of two left-handed Weyl fields χ and ξ via

$$\Psi = \begin{pmatrix} \chi \\ \xi^\dagger \end{pmatrix}$$

In terms of χ and ξ, the Lagrangian is

$$\mathcal{L} = i\chi^\dagger\overline{\sigma}^\mu(\partial_\mu - igA_\mu)\chi + i\xi^\dagger\overline{\sigma}^\mu(\partial_\mu + igA_\mu)\xi$$

Because the fermion field is massless, the Lagrangian is invariant under a global symmetry in which χ and ξ transform with the same phase

$$\chi(x) \to e^{i\alpha}\chi(x) \quad \xi(x) \to e^{i\alpha}\xi(x)$$

In terms of Ψ, this is

$$\Psi(x) \to e^{-i\alpha\gamma_5}\Psi(x) \quad \overline{\Psi}(x) \to \overline{\Psi}e^{-i\alpha\gamma_5}$$

This is called axial $U(1)$ symmetry, because the associated Noether current

$$j_A^\mu = \overline{\Psi}(x)\gamma^\mu\gamma_5\Psi(x)$$

is an axial vector. Noether's theorem leads us to expect that this current is conserved. However, we will show that the axial current actually has an anomalous divergence.

Consider the matrix element $\langle p, q|j_A^\rho(z)|0\rangle$, where $\langle p, q|$ is a state of two outgoing photons with four-momenta p and q, and polarization vectors ϵ_μ and ϵ'_ν, respectively. Using the LSZ formula for photons, we have

$$\langle p, q|j_A^\rho(z)|0\rangle = (ig)^2\epsilon_\mu\epsilon'_\nu \int d^4x d^4y e^{-i(px+qy)}\langle 0|Tj^\mu(x)j^\nu(y)j_A^\rho(z)|0\rangle$$

where

$$j^\mu = \overline{\Psi}\gamma^\mu\Psi$$

Let us define $C^{\mu\nu\rho}(p,q,r)$ via

$$(2\pi)^4\delta(p+q+r)C^{\mu\nu\rho}(p,q,r) \equiv \int d^4x d^4y d^4z e^{-i(px+qy+rz)}\langle 0|Tj^\mu(x)j^\nu(y)j_A^\rho(z)|0\rangle$$

Then we have

$$\langle p,q|j_A^\rho(z)|0\rangle = -g^2\epsilon_\mu\epsilon'_\nu\, C^{\mu\nu\rho}(p,q,r)e^{irz}\big|_{r=-q-p}$$

Taking the divergence of the current yields

$$\langle p,q|\partial_\rho j_A^\rho(z)|0\rangle = -ig^2\epsilon_\mu\epsilon'_\nu r_\rho\, C^{\mu\nu\rho}(p,q,r)e^{irz}\big|_{r=-q-p}$$

We compute $C^{\mu\nu\rho}(p,q,r)$ with Feynman diagrams. At the one-loop level, the contributing diagrams are exactly those we computed in previous subsection, except that the three vertex factors are now γ^μ, γ^ν and $\gamma^\rho\gamma_5$ instead of $ig\gamma^\mu P_L$, $ig\gamma^\nu P_L$ and $ig\gamma^\rho P_L$. But, as we saw, the three P_Ls can be combined into just one at the last vertex, and then this one can be replaced by $-\frac{1}{2}\gamma_5$. Thus, the vertex function $iV^{\mu\nu\rho}(p,q,r)$ is related to $C^{\mu\nu\rho}(p,q,r)$ by

$$iV^{\mu\nu\rho}(p,q,r) = -\frac{1}{2}(ig)^3C^{\mu\nu\rho}(p,q,r) + O(g^5)$$

We could choose a regularization scheme that preserved $p_\mu C^{\mu\nu\rho}(p,q,r) = 0$ and $q_\nu C^{\mu\nu\rho}(p,q,r) = 0$ but not also $r_\rho C^{\mu\nu\rho}(p,q,r) = 0$. This results in

$$r_\rho C^{\mu\nu\rho}(p,q,r) = -\frac{i}{2\pi^2}\epsilon^{\mu\nu\alpha\beta}p_\alpha q_\beta + O(g^2)$$

So, we have

$$\langle p,q|j_A^\rho(z)|0\rangle = -\frac{g^2}{2\pi^2}\epsilon^{\mu\nu\alpha\beta}p_\alpha q_\beta\epsilon_\mu\epsilon'_\nu e^{-i(p+q)z} + O(g^4)$$

The right-hand side of the equation is exactly what we get in free-field theory for the matrix element of

$$-\frac{g^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} + O(g^4)$$

In the next section, we will see the relation is exact.

26.4.3 Anomalies and the path integral for fermions

We begin with the path integral over the Dirac field, with the gauge field treated as a fixed background, to be integrated later. We have

$$Z(A) \equiv \int \mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS(A)}$$

where

$$S(A) \equiv \int d^4x\overline{\Psi}i\not{D}\Psi$$

is the Dirac action, $i\not{D} = i\gamma^\mu D_\mu$ is the Dirac operator, and

$$D_\mu = \partial_\mu - igA_\mu$$

is the covariant derivative. Here A_μ is either the $U(1)$ gauge field, or the matrix-valued non-abelian gauge field, depending on the theory under consideration.
Now consider an axial $U(1)$ transformation of the Dirac field, but with a space-time dependent parameter $\alpha(x)$

$$\Psi(x) \to e^{-i\alpha(x)\gamma_5}\Psi(x) \quad \overline{\Psi}(x) \to \overline{\Psi}e^{-i\alpha(x)\gamma_5}$$

The corresponding change in the action is

$$S(A) \to S(A) + \int d^4x j_A^\mu(x)\partial_\mu\alpha(x) = S(A) - \int d^4x\alpha(x)\partial_\mu j_A^\mu(x)$$

If we assume that the measure $\mathcal{D}\Psi\mathcal{D}\overline{\Psi}$ is invariant under the axial $U(1)$ transformation, then we have

$$Z(A) = \int \mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS(A)}e^{-i\int d^4x\alpha(x)\partial_\mu j_A^\mu(x)}$$

This implies that $\partial_\mu j_A^\mu(x) = 0$ holds inside quantum correlation functions, up to contact terms.

However, the assumption that the measure $\mathcal{D}\Psi\mathcal{D}\overline{\Psi}$ is invariant under the axial $U(1)$ transformation must be examined more closely. The change of variable is implemented by the functional matrix

$$J(x,y) = \delta(x-y)e^{-i\alpha(x)\gamma_5}$$

Because the path integral is over fermionic variables (rather than bosonic), we get a jacobian factor of $(\det J)^{-1}$ (rather than $\det J$) for each of the transformations, so that we have

$$\mathcal{D}\Psi\mathcal{D}\overline{\Psi} \to (\det J)^{-2}\mathcal{D}\Psi\mathcal{D}\overline{\Psi}$$

Using $\log\det J = \mathrm{Tr}\log J$, we can write

$$(\det J)^{-2} = \exp\left[2i\int d^4x\alpha(x)\mathrm{Tr}\delta(x-x)\gamma_5\right]$$

where the explicit trace is over spin and group indices. We regularize the delta function by

$$\delta(x-y) \to e^{(i\not{D}_x)^2/M^2}\delta(x-y)$$

The following calculation can be found in section 77 of *Quantum Field Theory (Mark Srednichi)*. Here, we list the final result

$$\mathrm{Tr}\delta(x-y)\gamma_5 \to -\frac{g^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}\mathrm{Tr}F_{\mu\nu}F_{\rho\sigma}$$

We then have

$$Z(A) = \int \mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS(A)}e^{-i\int d^4x\alpha(x)[\frac{g^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}\mathrm{Tr}F_{\mu\nu}F_{\rho\sigma}\partial_\mu j_A^\mu(x)]}$$

So

$$\partial_\mu j_A^\mu = -\frac{g^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}\mathrm{Tr}F_{\mu\nu}F_{\rho\sigma}$$

is exact. This result is known as the Adler–Bardeen theorem.

26.5 Spontaneous breaking of gauge symmetries

Consider scalar electrodynamics, specified by the Lagrangian

$$\mathcal{L} = -(D_\mu\phi)^\dagger D_\mu\phi - V(\phi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}$$

where ϕ is a complex scalar field, $D_\mu = \partial_\mu - ieA_\mu$, and

$$V(\phi) = m^2\phi^\dagger\phi + \frac{1}{4}\lambda(\phi^\dagger\phi)^2$$

Let us consider $m^2 < 0$. Classically, the field has a nonzero vacuum expectation value (VEV for short), given by

$$\langle 0|\phi(x)|0\rangle = \frac{1}{\sqrt{2}}v$$

where we have made a global $U(1)$ transformation to set the phase of the VEV to zero, and

$$v = \sqrt{\frac{4|m|^2}{\lambda}}$$

We therefore write

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x))e^{-i\chi(x)/v}$$

where $\rho(x)$ and $\chi(x)$ are real scalar fields. The scalar potential depends only on ρ, and is given by

$$V = \frac{1}{4}\lambda v^2\rho^2 + \frac{1}{4}\lambda v\rho^3 + \frac{1}{16}\lambda\rho^4$$

Since χ does not appear in the potential, it is massless; it is the Goldstone boson for the spontaneously broken $U(1)$ symmetry. We can make a gauge transformation that shifts the phase of $\phi(x)$ by an arbitrary space-time function. We can use this gauge freedom to set $\chi(x) = 0$; this choice is called unitary gauge. We have

$$-(D_\mu\phi)^\dagger D_\mu\phi = -\frac{1}{2}\partial^\mu\rho\partial_\mu\rho - \frac{1}{2}g^2(v+\rho)^2A^\mu A_\mu$$

We see that the gauge field now has a mass

$$M = gv$$

This is the Higgs mechanism: the Goldstone boson disappears, and the gauge field acquires a mass. Note that this leaves the counting of particle spin states unchanged: a massless spin-one particle has two spin states, but a massive one has three. The Goldstone boson has become the third or longitudinal state of the now-massive gauge field. A scalar field whose VEV breaks a gauge symmetry is generically called a Higgs field.

This generalizes in a straightforward way to a nonabelian gauge theory. Consider a complex scalar field ϕ in a representation R of the gauge group. The kinetic term for φ is $-(D_\mu\phi)^\dagger D_\mu\phi$,

where the covariant derivative is $D_\mu\phi_i = \partial_\mu\phi_i - igA^a_\mu(T^a_R)_i{}^j\phi_j$, and the indices i and j run from 1 to $d(R)$. We assume that ϕ acquires a VEV

$$\langle 0|\phi_i(x)|0\rangle = \frac{1}{\sqrt{2}}v_i$$

If we replace ϕ by its VEV in $-(D_\mu\phi)^\dagger D_\mu\phi$, we find a mass term for the gauge fields

$$\mathcal{L}_{mass} = -\frac{1}{2}(M^2)^{ab}A^{a\mu}A^b_\mu$$

where the mass-squared matrix is

$$(M^2)^{ab} = g^2 v_i^*(T^a_R T^b_R)_{ij}v_j$$

If the field ϕ is real rather than complex (which is possible only if R is a real representation), then we remove the factor of root-two from the right-hand side of vacuum expectation value of ϕ, but this is compensated by an extra factor of one-half from the kinetic term for a real scalar field; thus the form of mass-squared matrix remains invariant.
We say a generator T^a is spontaneously broken if $(T^a_R)_{ij}v_j \neq 0$. We see that gauge fields corresponding to broken generators get a mass, while those corresponding to unbroken generators do not. The unbroken generators (if any) form a gauge group with massless gauge fields. The massive gauge fields (and all other fields) form representations of this unbroken group.

Consider the gauge group $SU(N)$, with a complex scalar field ϕ in the fundamental representation. We can make a global $SU(N)$ transformation to bring the VEV entirely into the last component, and furthermore make it real. Any generator $(T^a)_i$ that does not have a nonzero entry in the last column will remain unbroken. These generators form an unbroken $SU(N-1)$ gauge group. There are three classes of broken generators: those with $(T^a)_i = \frac{1}{2}$ for $i \neq N$ (there are $N-1$ of these); those with $(T^a)_i = -\frac{i}{2}$ for $i \neq N$ (there are also $N-1$ of these), and finally the single generator $T^{N^2-1} = [2N(N-1)]^{-1/2}\mathrm{diag}(1, 1, \cdots, -N-1)$.
The gauge fields corresponding to the generators in the first two classes get a mass $M = \frac{1}{2}gv$. we can group them into a complex vector field that transforms in the fundamental representation of the unbroken $SU(N-1)$ subgroup. The gauge field corresponding to T^{N^2-1} gets a mass $M = [2N(N-1)]^{1/2}gv$; it is a singlet of $SU(N-1)$.

Consider the gauge group $SO(N)$, with a real scalar field in the fundamental representation. We can make a global $SO(N)$ transformation to bring the VEV entirely into the last component. Any generator $(T^a)_i$ that does not have a nonzero entry in the last column will remain unbroken. These generators form an unbroken $SO(N-1)$ subgroup. There are $N-1$ broken generators, those with $(T^a)_i = -i$ for $i \neq N$. The corresponding gauge fields get a mass $M = gv$; they form a fundamental representation of the unbroken $SO(N-1)$ subgroup.

Consider the gauge group $SU(N)$, with a real scalar field Φ a in the adjoint representation. It will prove more convenient to work with the $N \times N$ matrix-valued field $\Phi = \phi^a T^a$; the covariant derivative of Φ is $D_\mu\Phi = \partial_\mu\Phi - igA^a_\mu[T^a, \Phi]$, and the VEV of ϕ is a traceless hermitian $N \times N$ matrix V. Thus the mass squared matrix for the gauge fields is

$$-2g^2\mathrm{Tr}\left([T^a, V][T^b, V]\right)$$

We can make a global $SU(N)$ transformation to bring V into diagonal form. Suppose the diagonal entries consist of N_1 v_1s, followed by N_2 v_2s, etc., where $v_1 < v_2 < \cdots$ and $\sum_i N_i v_i = 0$. Then all generators whose nonzero entries lie entirely within the ith block commute with V, and hence form an unbroken $SU(N_i)$ subgroup. Furthermore, generators that is proportional to V also commutes with V, and forms a $U(1)$ subgroup. Thus the unbroken gauge group is $SU(N_1) \times SU(N_2) \times \cdots \times U(1)$. The gauge coupling constants for the different groups are all the same, and equal to the original $SU(N)$ gauge coupling constant.

Example: Consider the case of $SU(5)$, which has 24 generators. Let the diagonal entries of V be given by $(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{2}, \frac{1}{2})$. The unbroken subgroup is then $SU(3) \times SU(2) \times U(1)$. The number of broken generators is $24 - 8 - 3 - 1 = 12$.

26.6 Quantization of spontaneously broken gauge theory

26.6.1 Spontaneously broken abelian gauge theory

Consider scalar electrodynamics, specified by the Lagrangian

$$\mathcal{L} = -(D_\mu \phi)^\dagger D_\mu \phi - V(\phi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

where ϕ is a complex scalar field, $D_\mu = \partial_\mu - ieA_\mu$. We choose

$$V(\phi) = \frac{1}{4}\lambda(\phi^\dagger \phi - \frac{1}{2}v^2)^2$$

which yields a nonzero VEV for ϕ. We therefore write

$$\phi = \frac{1}{\sqrt{2}}(v + h + ib)$$

where h and b are real scalar fields. In terms of h and b, the potential is

$$V = \frac{1}{4}\lambda v^2 h^2 + \frac{1}{4}\lambda v h(h^2 + b^2) + \frac{1}{16}\lambda(h^2 + b^2)^2$$

The covariant derivative term can be expanded as

$$\begin{aligned} -(D_\mu \phi)^\dagger D_\mu = \; & - \frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{2}\partial_\mu b \partial^\mu b - \frac{1}{2} g^2 v^2 A^\mu A_\mu + g v A^\mu \partial_\mu b \\ & + g A^\mu (h \partial_\mu b - b \partial_\mu h) - g v h A^\mu A_\mu - \frac{1}{2} g^2 (h^2 + b^2) A^\mu A_\mu \end{aligned}$$

The first line contains all the terms that are quadratic in the fields. The first two are the kinetic terms for the h and b fields. The third is the mass term for the vector field. The fourth is an annoying cross term between the vector field and the derivative of b.

In abelian gauge theory, in the absence of spontaneous symmetry breaking, we fix gauge by adding to $\mathcal{L}$ the gauge-fixing and ghost terms

$$\mathcal{L}_{gf} + \mathcal{L}_{gh} = -\frac{G^2}{2\xi} - \bar{c}\frac{\delta G}{\delta\theta}c$$

where $G = \partial^\mu A_\mu$, and $\theta(x)$ parameterizes an infinitesimal gauge transformation

$$A_\mu \to A_\mu - \partial_\mu\theta \quad \phi \to ig\theta\phi$$

Since $\delta G/\delta\theta = -\partial^2$, the ghost fields have no interactions, and can be ignored.

In the presence of spontaneous symmetry breaking, we choose instead

$$G = \partial^\mu A_\mu - \xi g\nu b$$

which reduces to $\partial^\mu A_\mu$ when $v = 0$. Multiplying out G^2 , we have

$$\mathcal{L}_{gf} = -\frac{1}{2\xi}\partial^\mu A_\mu \partial^\nu A_\nu - gvA_\mu\partial^\mu b - \frac{1}{2}\xi g^2 v^2 b^2$$

Note that the second term cancels the annoying cross term between the vector field and the derivative of b. Also, the last term on the gives a mass $\sqrt{\xi}M$ to the b field ($M = gv$).

We must still evaluate $\mathcal{L}_{gh}$. To do so, we note under gauge transformation,

$$h \to h + g\theta b \quad b \to b - g\theta(v + h)$$

Then we have

$$\frac{\delta G}{\delta\theta} = -\partial^2 + \xi g^2 v(v + h)$$

The ghost Lagrangian is

$$\mathcal{L}_{gh} = -\partial^\mu\bar{c}\partial_\mu c - \xi g^2 v^2\bar{c}c - \xi g^2 vh\bar{c}c$$

We see from the second term that the ghost has acquired the same mass as the b field.

Now let us examine the vector field. Including $\mathcal{L}_{gh}$, the terms that are quadratic in the vector field can be written as

$$\mathcal{L}_0 = -\frac{1}{2}A_\mu\left[\eta^{\mu\nu}(-\partial^2 + M^2) + (1-\xi)\partial^\mu\partial^\nu\right]A_\nu$$

The propagator for vector field is

$$S_F(k) = \frac{-iP^{\mu\nu}}{k^2 + M^2 - i\epsilon} + \frac{\xi k^\mu k^\nu/k^2}{k^2 + \xi M^2 - i\epsilon}$$

where $P^{\mu\nu} = g^{\mu\nu} - k^\mu k^\nu/k^2$ projects onto the transverse subspace. We see that the transverse components of the vector field propagate with mass M, while the longitudinal component propagates with the same mass as the b and ghost fields, $\sqrt{\xi}M$.

Since their masses depend on ξ, the ghosts, the b field, and the longitudinal component of the vector field must all represent unphysical particles that do not appear in incoming or outgoing states. When $\xi \to \infty$, we can recover the unitary gauge. The details can be found in section 76 of *Quantum Field Theory (Mark Srednichi)*.

26.6.2 Spontaneously broken nonabelian gauge theory

It will be convenient to work with real scalar fields. We therefore decompose any complex scalar fields into pairs of real ones, and organize all the real scalar fields into a big list ϕ_i, $i = 1, \cdots, N$. These real scalar fields form a (possibly reducible) representation R of the gauge group. Let T^a be the gauge-group generator matrices; they are linear combinations of the generators of the $SO(N)$ group that rotates all components of ϕ_i into each other. Because these generators are hermitian and antisymmetric, so are the T^As. The Lagrangian for our theory can now be written as

$$\mathcal{L} = -\frac{1}{2}D^\mu\phi D_\mu\phi - V(\phi) - \frac{1}{4}F^{a\mu\nu}F^a_{\mu\nu}$$

where

$$(D_\mu\phi)_i = \partial_\mu\phi_i - ig_a A^a_\mu T^a_{ij}\phi_j$$

is the covariant derivative.
Now we suppose that the potential is minimized when ϕ has a VEV

$$\langle 0|\phi_i|0\rangle = v_i$$

A generator T^a is unbroken if $T^a_{ij}v_j = 0$, and broken if $T^a_{ij}v_j \neq 0$.

Each broken generator results in a massless Goldstone boson. We note that the potential must be invariant under a global gauge transformation, so

$$\frac{\partial V}{\partial\phi_j}T^a_{jk}\phi_k = 0$$

We differentiate it with respect to ϕ_i to get

$$\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}T^a_{jk}\phi_k + \frac{\partial V}{\partial\phi_j}T^a_{ji} = 0$$

Now set $\phi_i = v_i$; then $\frac{\partial V}{\partial\phi_i}$ vanishes. Also, we can identify

$$(m^2)_{ij} = \left.\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\right|_{\phi_i=v_i}$$

as the mass-squared matrix for the scalars (after spontaneous symmetry breaking). Thus we have

$$(m^2)_{ij}(T^a v)_j = 0$$

We see that if $T^a v \neq 0$, then $T^a v$ is an eigenvector of the mass-squared matrix with eigenvalue zero. So there is a zero eigenvalue for every linearly independent broken generator.

Let us write

$$\phi_i(x) = v_i + \chi_i(x)$$

The covariant derivative of becomes

$$(D_\mu\phi)_i = \partial_\mu\chi_i - ig_a A^a_\mu T^a_{ij}(v+\chi)_j$$

It is now convenient to define a set of real antisymmetric matrices

$$\tau^a_{ij} \equiv ig_a T^a_{ij}$$

and the real rectangular matrix

$$F^a_{\ i} \equiv \tau^a_{ij} v_j$$

We can now write

$$(D_\mu \phi)_i = \partial_\mu \chi_i - A^a_\mu (F^a + \tau^a \chi)_i$$

The kinetic term for ϕ becomes

$$\begin{aligned} -\frac{1}{2} D^\mu \phi D_\mu \phi = & - \frac{1}{2} \partial_\mu \chi_i \partial^\mu \chi_i - \frac{1}{2} (F^a_{\ i} F^b_{\ i}) A^{a\mu} A^b_\mu + F^a_{\ i} A^a_\mu \partial^\mu \chi_i \\ & + A^a_\mu \chi_i \tau^a_{ij} \partial^\mu \chi_j - A^{a\mu} A^b_\mu F^a_{\ i} \tau^b_{ij} \chi_j - \frac{1}{2} A^{a\mu} A^b_\mu \chi_i (\tau^a \tau^b)_{ij} \chi_j \end{aligned}$$

We see that the mass-squared matrix for the vector fields is

$$(M^2)^{ab} = F^a_{\ i} F^b_{\ i} = (FF^T)^{ab}$$

A theorem of linear algebra states that every real rectangular matrix can be written as

$$F^a_{\ i} = S^{ac} (M^c \delta^c_{\ j}) R_{ji}$$

where S and R are orthogonal matrices, and the diagonal entries M^c are real and nonnegative. We see that these diagonal entries are the masses of the vector fields. The vector fields of definite mass are then given by $\widetilde{A}^a_\mu = S^{ba} A^b_\mu$.

Now we are ready to fix R_ξ gauge. To do so, we add to $\mathcal{L}$ the gauge-fixing and ghost terms

$$\mathcal{L}_{gf} + \mathcal{L}_{gh} = -\frac{G^a G^a}{2\xi} - \overline{c}^a \frac{\delta G^a}{\delta \theta^b} c^b$$

where we choose

$$G^a = \partial^\mu A^a_\mu - \xi F^a_{\ i} \chi_i$$

Then we have

$$\mathcal{L}_{gf} = -\frac{1}{2\xi} \partial^\mu A^a_\mu \partial^\nu A^a_\nu - F^a_{\ i} A^a_\mu \partial^\mu \chi_i - \frac{1}{2} \xi (F^a_{\ i} F^a_{\ j}) \chi_i \chi_j$$

The last term makes a contribution to the mass-squared matrix for the χ fields,

$$\xi (M^2)_{ij} = \xi F^a_{\ i} F^a_{\ j} = \xi (F^T F)_{ij}$$

The eigenvalues of this matrix are $\sqrt{\xi} M^a$, where M^a are the vector-boson masses. The mass-squared matrix ξM^2 should be added to the mass-squared matrix m^2 that we get from the potential. We can verify that

$$(m^2)_{ij} (\xi M^2)_{jk} = 0$$

Thus these two contributions to the mass-squared matrix of the scalar fields live in orthogonal subspaces. The m^2 subspace consists of the physical, massive scalars, and the ξM^2 subspace

consists of the unphysical Goldstone bosons; these are the fields that would be set to zero in unitary gauge.

Finally, we must evaluate $\mathcal{L}_{gh}$. Under an infinitesimal gauge transformation, we have

$$A^a_\mu \to A^a_\mu - D^\mu_{ab}\theta^b \quad \chi_i \to \chi_i - \theta^a \tau^a_{ij}(v+\chi)_j$$

Thus we have

$$\frac{\delta G^a}{\delta \theta^b} = -D^\mu_{ab} + \xi (M^2)^{ab} + \xi F^a_{\ j} \tau^b_{jk} \chi_k$$

and so the ghost lagrangian is

$$\mathcal{L}_{gh} = -\partial^\mu \overline{c}^a D^{ab}_\mu c^b - \xi (M^2)^{ab} \overline{c}^a c^b - \xi F^a_{\ j} \tau^b_{jk} \chi_k \overline{c}^a c^b$$

The ghost fields of definite mass are $\widetilde{c}^a = S^{ba} c^b$ and $\widetilde{\overline{c}}^a = S^{ba} \overline{c}^b$.

26.7 The Standard Model

26.7.1 Gauge and Higgs sector

We begin with the electroweak part of the gauge group, $SU(2) \times U(1)$, and the complex scalar field ϕ, known as the Higgs field, in the representation $(2, -\frac{1}{2})$. The Higgs field acquires a nonzero VEV that spontaneously breaks $SU(2) \times U(1)$ to $U(1)$; the unbroken $U(1)$ is identified as electromagnetism. We begin with the covariant derivative of the Higgs field ϕ,

$$(D_\mu \phi)_i = \partial_\mu \phi_i - i[g_2 A^a_\mu T^a + g_1 B_\mu Y]_i^{\ j} \phi_j$$

where $T^a = \frac{1}{2}\sigma^a$ and $Y = -\frac{1}{2}I$; Y is the hypercharge generator. It will prove useful to write out $g_2 A^a_\mu T^a + g_1 B_\mu Y$ in matrix form,

$$\frac{1}{2}\begin{pmatrix} g_2 A^3_\mu - g_1 B_\mu & g_2(A^1_\mu - iA^2_\mu) \\ g_2(A^1_\mu + iA^2_\mu) & -g_2 A^3_\mu - g_1 B_\mu \end{pmatrix}$$

Now suppose that ϕ has a potential

$$V(\phi) = \frac{1}{4}\lambda(\phi^\dagger \phi - \frac{1}{2}v^2)^2$$

This potential gives ϕ a nonzero VEV. We can make a global gauge transformation to bring this VEV entirely into the first component, and furthermore make it real, so that

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} v \\ 0 \end{pmatrix}$$

The kinetic term for ϕ is $-(D_\mu\phi)^\dagger D_\mu \phi$. After replacing ϕ by its VEV, we find a mass term for the gauge fields,

$$\mathcal{L}_{mass} = -\frac{1}{8}v^2 \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} g_2 A^3_\mu - g_1 B_\mu & g_2(A^1_\mu - iA^2_\mu) \\ g_2(A^1_\mu + iA^2_\mu) & -g_2 A^3_\mu - g_1 B_\mu \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Define the weak mixing angle

$$\theta_W \equiv \tan^{-1}(g_1/g_2)$$

and the fields

$$\begin{aligned} W_\mu^\pm &\equiv \frac{1}{\sqrt{2}}(A_\mu^1 \mp iA_\mu^2) \\ Z_\mu &\equiv c_w A_\mu^3 - s_w B_\mu \\ A_\mu &\equiv s_w A_\mu^3 + c_w B_\mu \end{aligned}$$

In terms of these fields, we have

$$\mathcal{L}_{mass} = -M_W^2 W^{+\mu} W_\mu^- - \frac{1}{2} M_Z^2 Z^\mu Z_\mu$$

where we have identified

$$M_W = \frac{g_2 v}{2} \quad M_Z = \frac{M_W}{\cos\theta_W}$$

Note that the A_μ field remains massless; this signifies that there is an unbroken $U(1)$ subgroup. We will identify this unbroken $U(1)$ with the gauge group of electromagnetism.

The two complex components of the ϕ field yield four real scalar fields; three of these become the longitudinal components of the $W^\pm$ and Z^0. The remaining scalar field must be able to account for shifts in the overall scale of ϕ. Thus we can write, in unitary gauge,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix}$$

where H is a real scalar field; the corresponding particle is the Higgs boson. The potential now reads

$$V = \frac{1}{4}\lambda v^2 H^2 + \frac{1}{4}\lambda v H^3 + \frac{1}{16}\lambda H^4$$

We see that the mass of the Higgs boson is given by $m_H^2 = \frac{1}{2}\lambda v^2$. The kinetic term for H comes from the kinetic term for ϕ, and is the usual one for a real scalar field, $-\frac{1}{2}\partial_\mu H \partial^\mu H$. Finally, recall that the mass term for the gauge fields is proportional to v^2. Hence it should be multiplied by a factor of $(1 + \frac{H}{v})^2$.

Now we have to work out the kinetic terms for the gauge fields. We have

$$\mathcal{L} = -\frac{1}{4} F^{a\mu\nu} F_{\mu\nu}^a - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}$$

We find

$$\frac{1}{\sqrt{2}}(F_{\mu\nu}^1 - iF_{\mu\nu}^2) = D_\mu W_\nu^+ - D_\nu W_\mu^+$$

$$\frac{1}{\sqrt{2}}(F_{\mu\nu}^1 + iF_{\mu\nu}^2) = D_\mu W_\nu^- - D_\nu W_\mu^-$$

where we have defined a covariant derivative that acts on W^+,

$$D_\mu \equiv \partial_\mu - ig_2 A_\mu^3 = \partial_\mu - ig_2(s_W A_\mu + c_W Z_\mu)$$

If we identify A_μ as the electromagnetic vector potential, and assign electric charge $Q = +1$ (in units of the proton charge) to the W^+, then we must identify the electromagnetic coupling constant e as

$$e = g_2 \sin\theta_W$$

Here we are adopting the convention that e is positive. (In our treatment of quantum electrodynamics, we used the convention that e is negative, but that is less convenient in the present context.) We also have

$$F^3_{\mu\nu} = s_W F_{\mu\nu} + c_W Z_{\mu\nu} - ig_2(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu)$$

$$B_{\mu\nu} = c_W F_{\mu\nu} - s_W Z_{\mu\nu}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual electromagnetic field strength, and

$$Z_{\mu\nu} \equiv \partial_\mu Z_\nu - \partial_\nu Z_\mu$$

is the abelian field strength associated with the Z_μ field.

Now we can assemble all of this into the complete Lagrangian for the electroweak gauge fields and the Higgs boson in unitary gauge. We will express g_2 in terms of e and θ_W, and λ in terms of m_H and v. We ultimately get

$$\begin{aligned}
\mathcal{L} = \; & - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - D^{\dagger\mu}W^{-\nu}D_\mu W^+_\nu + D^{\dagger\mu}W^{-\nu}D_\nu W^+_\mu \\
& + ie(F^{\mu\nu} + \cot\theta_W Z^{\mu\nu})W^+_\mu W^-_\nu \\
& - \frac{1}{2}\frac{e^2}{\sin^2\theta_W}(W^{+\mu}W^-_\mu W^{+\nu}W^-_\nu - W^{+\mu}W^+_\mu W^{-\nu}W^-_\nu) \\
& - (M^2_W W^{+\mu}W^-_\mu + \frac{1}{2}M^2_Z Z^\mu Z_\mu)(1 + \frac{H}{v})^2 \\
& - \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m^2_H H^2 - \frac{1}{2}m^2_H v^{-1}H^3 - \frac{1}{8}m^2_H v^{-2}H^4
\end{aligned}$$

where

$$D_\mu = \partial_\mu - ie(A_\mu + \cot\theta_W Z_\mu)$$

26.7.2 Lepton sector

Leptons are spin-one-half particles that are singlets of the color group. There are six different flavours of lepton. The six flavours are naturally grouped into three families or generations: e and ν_e, μ and ν_μ, τ and ν_τ.

Let us begin by describing a single lepton family, the electron and its neutrino. We introduce left-handed Weyl fields l and $\bar{e}$ in the representations $(2, -\frac{1}{2})$ and $(1, +1)$ of $SU(2) \times U(1)$. Here the bar over the e in the field $\bar{e}$ is part of the name of the field, and does not denote any sort of conjugation. The covariant derivatives of these fields are

$$(D_\mu l)_i = \partial_\mu l_i - ig_2 A^a_\mu (T^a)_i{}^j l_j - ig_1(-\frac{1}{2})B_\mu l_i$$

$$D_\mu \bar{e} = \partial_\mu \bar{e} - ig_1(+1)B_\mu \bar{e}$$

and their kinetic terms are

$$\mathcal{L}_{kin} = il^{\dagger i}\overline{\sigma}^\mu (D_\mu l)_i + i\bar{e}^\dagger \overline{\sigma}^\mu D_\mu \bar{e}$$

We cannot write down a mass term involving l and/or $\bar{e}$ because there is no gauge-group singlet contained in any of the products

$$(2, -\frac{1}{2}) \otimes (2, -\frac{1}{2}) \quad (2, -\frac{1}{2}) \otimes (1, +1) \quad (1, +1) \otimes (1, +1)$$

However, we are able to write down a Yukawa coupling of the form

$$\mathcal{L}_{Yuk} = -y\epsilon^{ij}\phi_i l_j \bar{e} + \text{h.c.}$$

where ϕ is the Higgs field in the $(2, -\frac{1}{2})$ representation, and y is the Yukawa coupling constant. A gauge-invariant Yukawa coupling is possible because there is a singlet on the right-hand side of

$$(2, -\frac{1}{2}) \otimes (2, -\frac{1}{2}) \otimes (1, +1) = (1, 0) \oplus (3, 0)$$

In unitary gauge, we replace ϕ_1 with $\frac{1}{\sqrt{2}}(v + H)$, where H is the real scalar field representing the physical Higgs boson, and ϕ_2 with zero. The Yukawa term becomes

$$\mathcal{L}_{Yuk} = -\frac{1}{\sqrt{2}} y(v + H)(l_2 \bar{e} + \text{h.c.})$$

It is now convenient to assign new names to the $SU(2)$ components of l,

$$l = \begin{pmatrix} \nu \\ e \end{pmatrix}$$

So, we have

$$\mathcal{L}_{Yuk} = -\frac{1}{\sqrt{2}} y(v + H)\overline{\mathcal{E}}\mathcal{E}$$

where we have defined a Dirac field for the electron

$$\mathcal{E} \equiv \begin{pmatrix} e \\ \bar{e}^\dagger \end{pmatrix}$$

We see that the electron has acquired a mass $m_e = \frac{yv}{\sqrt{2}}$, while neutrino has remained massless. We can describe the neutrino with a Majorana field

$$\mathcal{N} \equiv \begin{pmatrix} \nu \\ \nu^\dagger \end{pmatrix}$$

However, it is often more convenient to work with

$$\mathcal{N}_L \equiv P_L \mathcal{N} = \begin{pmatrix} \nu \\ 0 \end{pmatrix}$$

We can think of $\mathcal{N}_L$ as a Dirac field; for example, the neutrino kinetic term can be written as $i\overline{\mathcal{N}_L}\partial\!\!\!/\mathcal{N}_L$.

Now we express the covariant derivatives in terms of the $W_\mu^\pm$, Z_μ , and A_μ fields. We have

$$g_2 A_\mu^1 T^1 + g_2 A_\mu^2 T^2 = \frac{g_2}{\sqrt{2}} \begin{pmatrix} 0 & W_\mu^+ \\ W_\mu^- & 0 \end{pmatrix}$$

and

$$g_2 A_\mu^3 T^3 + g_1 B_\mu Y = e(T^3 + Y)A_\mu + e(\cot\theta_W T^3 - \tan\theta_W Y)Z_\mu$$

Since we identify A_μ as the electromagnetic field and e as the electromagnetic coupling constant (with the convention that e is positive), we identify

$$Q = T^3 + Y$$

as the generator of electric charge. Then we have

$$Q\nu = 0 \quad Qe = -e \quad Q\bar{e} = +\bar{e}$$

It is convenient to replace Y with $Q - T^3$. We find

$$g_2 A_\mu^3 T^3 + g_1 B_\mu Y = eQA_\mu + \frac{e}{s_W c_W}(T^3 - s_W^2 Q)Z_\mu$$

In terms of the four-component fields, we have

$$(g_2 A_\mu^3 T^3 + g_1 B_\mu Y)\mathcal{E} = \left[-eA_\mu + \frac{e}{s_W c_W}(-\frac{1}{2}P_L + s_W^2)Z_\mu\right]\mathcal{E}$$

$$(g_2 A_\mu^3 T^3 + g_1 B_\mu Y)\mathcal{N}_L = \frac{e}{s_W c_W}(+\frac{1}{2})Z_\mu \mathcal{N}_L$$

The couplings of the gauge fields to the leptons can be written as

$$\mathcal{L}_{int} = \frac{1}{\sqrt{2}} g_2 W_\mu^+ J^{-\mu} + \frac{1}{\sqrt{2}} g_2 W_\mu^- J^{+\mu} + \frac{e}{s_W c_W} Z_\mu J_Z^\mu + eA_\mu J_{EM}^\mu$$

where

$$\begin{aligned} J^{+\mu} &\equiv \overline{\mathcal{E}}_L \gamma^\mu \mathcal{N}_L \\ J^{-\mu} &\equiv \overline{\mathcal{N}}_L \gamma^\mu \mathcal{E}_L \\ J_Z^\mu &\equiv J_3^\mu - s_W^2 J_{EM}^\mu \\ J_3^\mu &\equiv \frac{1}{2}\overline{\mathcal{N}}_L \gamma^\mu \mathcal{N}_L - \frac{1}{2}\overline{\mathcal{E}}_L \gamma^\mu \mathcal{E}_L \\ J_{EM}^\mu &\equiv -\overline{\mathcal{E}}\gamma^\mu \mathcal{E}_L \end{aligned}$$

Having worked out the interactions of a single lepton generation, we now examine what happens when there is more than one of them. Let us consider the fields l_{iI} and $\bar{e}_I$, where $I = 1, 2, 3$ is a generation index. The kinetic term for all these fields is

$$\mathcal{L}_{kin} = i l_I^{\dagger i} \overline{\sigma}^\mu (D_\mu)_i^{\ j} l_{jI} + i \bar{e}_I^\dagger \overline{\sigma}^\mu D_\mu \bar{e}_I$$

where the repeated generation index is summed. The most general Yukawa term we can write down now reads

$$\mathcal{L}_{Yuk} = -\epsilon^{ij}\phi_i l_{jI} y_{IJ}\bar{e}_J + \text{h.c.}$$

where y_{IJ} is a complex 3×3 matrix, and the generation indices are summed. We can make unitary transformations in generation space on the fields: $l_I \to L_{IJ}l_J$ and $\bar{e}_I \to \overline{E}_{IJ}\bar{e}_J$, where L and $\overline{E}$ are independent unitary matrices. The kinetic terms are unchanged, and the Yukawa matrix y is replaced with $L^T y\overline{E}$. We can choose L and $\overline{E}$ so that $L^T y\overline{E}$ is diagonal with positive real entries y_I. The charged leptons then have masses $m_{eI} = \frac{y_I v}{\sqrt{2}}$, and the neutrinos remain massless. In the currents, we simply add a generation index I to each field, and sum over it.

26.7.3 Quark sector

Quarks are spin-one-half particles that are triplets of the color group. There are six different flavours of quark. The six flavours are naturally grouped into three families or generations: u and d, c and s, t and b.

Let us begin by describing a single quark family, the up and down quarks. We introduce left-handed Weyl fields q, $\bar{u}$, and $\bar{d}$ in the representations $(3, 2, +\frac{1}{6})$, $(\bar{3}, 1, -\frac{2}{3})$, and $(\bar{3}, 1, +\frac{1}{3})$ of $SU(3) \times SU(2) \times U(1)$. Here the bar over the letter in the fields is part of the name of the field, and does not denote any sort of conjugation The covariant derivatives of these fields are

$$\begin{aligned}
(D_\mu q)_{\alpha i} &= \partial_\mu q_{\alpha i} - ig_3 A^a_\mu (T^a_3)_\alpha{}^\beta q_{\beta i} - ig_2 A^a_\mu (T^a_2)_i{}^j q_{\alpha j} - ig_1(+\frac{1}{6})B_\mu q_{\alpha i} \\
(D_\mu \bar{u})^\alpha &= \partial_\mu \bar{u}^\alpha - ig_3 A^a_\mu (T^a_{\bar{3}})^\alpha{}_\beta \bar{u}^\beta - ig_1(-\frac{2}{3})B_\mu \bar{u}^\alpha \\
(D_\mu \bar{d})^\alpha &= \partial_\mu \bar{d}^\alpha - ig_3 A^a_\mu (T^a_{\bar{3}})^\alpha{}_\beta \bar{d}^\beta - ig_1(+\frac{1}{3})B_\mu \bar{d}^\alpha
\end{aligned}$$

We rely on context to distinguish the $SU(3)$ gauge fields from the $SU(2)$ gauge fields. The kinetic terms for q, $\bar{u}$, and $\bar{d}$ are

$$\mathcal{L}_{kin} = iq^{\dagger\alpha i}\overline{\sigma}^\mu (D_\mu q)_{\alpha i} + i\bar{u}^\dagger_\alpha \overline{\sigma}^\mu (D_\mu \bar{u})^\alpha + i\bar{d}^\dagger_\alpha \overline{\sigma}^\mu (D_\mu \bar{d})^\alpha$$

We cannot write down a mass term involving q, $\bar{u}$, and/or $\bar{d}$ because there is no gauge-group singlet contained in any of the products of their representations. But we are able to write down Yukawa couplings of the form

$$\mathcal{L}_{Yuk} = -y'\epsilon^{ij}\phi_i q_{\alpha j}\bar{d}^\alpha - y"\phi^{\dagger i} q_{\alpha i}\bar{u}^\alpha + \text{h.c.}$$

where ϕ is the Higgs field in the $(1, 2, -\frac{1}{2})$ representation, and y' and $y"$ are the Yukawa coupling constants. These gauge invariant Yukawa couplings are possible because there are singlets on the right-hand sides of

$$(1, 2, -\frac{1}{2}) \otimes (3, 2, +\frac{1}{6}) \otimes (\bar{3}, 1, +\frac{1}{3}) = (1, 1, 0)\oplus$$

$$(1, 2, \frac{1}{2}) \otimes (3, 2, +\frac{1}{6}) \otimes (\bar{3}, 1, -\frac{2}{3}) = (1, 1, 0)\oplus$$

In unitary gauge, we replace ϕ_1 with $\frac{1}{\sqrt{2}}(v+H)$, where H is the real scalar field representing the physical Higgs boson, and ϕ_2 with zero. The Yukawa term becomes

$$\mathcal{L}_{Yuk} = -\frac{1}{\sqrt{2}}y'(v+H)q_{\alpha 2}\bar{d}^{\alpha} - \frac{1}{\sqrt{2}}y''(v+H)q_{\alpha 1}\bar{u}^{\alpha} + \text{h.c.}$$

It is now convenient to assign new names to the $SU(2)$ components of q,

$$q = \begin{pmatrix} u \\ q \end{pmatrix}$$

Then we have

$$\mathcal{L}_{Yuk} = -\frac{1}{\sqrt{2}}y'(v+H)\overline{\mathcal{D}}_{\alpha}\mathcal{D}_{\alpha} - \frac{1}{\sqrt{2}}y''(v+H)\overline{\mathcal{U}}_{\alpha}\mathcal{U}_{\alpha}$$

where we have defined Dirac fields for the down and up quarks

$$\mathcal{D}_{\alpha} \equiv \begin{pmatrix} d_{\alpha} \\ \bar{d}^{\dagger}_{\alpha} \end{pmatrix} \quad \mathcal{U}_{\alpha} \equiv \begin{pmatrix} u_{\alpha} \\ \bar{u}^{\dagger}_{\alpha} \end{pmatrix}$$

We see that the up and down quarks have acquired masses

$$m_d = \frac{y'v}{\sqrt{2}} \quad m_u = \frac{y''v}{\sqrt{2}}$$

Now we express the covariant derivatives in terms of the $W^{\pm}_{\mu}$, Z_{μ} , and A_{μ} fields. We have

$$g_2 A^1_{\mu} T^1 + g_2 A^2_{\mu} T^2 = \frac{g_2}{\sqrt{2}} \begin{pmatrix} 0 & W^+_{\mu} \\ W^-_{\mu} & 0 \end{pmatrix}$$

$$g_2 A^3_{\mu} T^3 + g_1 B_{\mu} Y = e(T^3 + Y)A_{\mu} + e(\cot\theta_W T^3 - \tan\theta_W Y)Z_{\mu}$$

and

$$Q = T^3 + Y$$

We see that

$$Qu = +\frac{2}{3}u \quad Qd = -\frac{1}{3}d \quad Q\bar{u} = -\frac{2}{3}\bar{u} \quad Q\bar{d} = +\frac{1}{3}\bar{d}$$

This is just the set of electric charge assignments that we expect for the up and down quarks. In terms of the four-component fields, we have

$$(g_2 A^3_{\mu} T^3 + g_1 B_{\mu} Y)\mathcal{U} = \left[+\frac{2}{3}eA_{\mu} + \frac{e}{s_W c_W}(\frac{1}{2}P_L - \frac{2}{3}s^2_W)Z_{\mu}\right]\mathcal{U}$$

$$(g_2 A^3_{\mu} T^3 + g_1 B_{\mu} Y)\mathcal{U} = \left[-\frac{1}{3}eA_{\mu} + \frac{e}{s_W c_W}(-\frac{1}{2}P_L + \frac{1}{3}s^2_W)Z_{\mu}\right]\mathcal{U}$$

The coupings of the electroweak gauge fields to the quarks is

$$\mathcal{L}_{int} = \frac{1}{\sqrt{2}}g_2 W^+_{\mu} J^{-\mu} + \frac{1}{\sqrt{2}}g_2 W^-_{\mu} J^{+\mu} + \frac{e}{s_W c_W} Z_{\mu} J^{\mu}_Z + eA_{\mu} J^{\mu}_{EM}$$

where we have defined the currents

$$\begin{aligned}
J^{+\mu} &\equiv \overline{\mathcal{D}}_L\gamma^\mu\mathcal{U}_L \\
J^{-\mu} &\equiv \overline{\mathcal{U}}_L\gamma^\mu\mathcal{U}_L \\
J_Z^\mu &\equiv J_3^\mu - s_W^2 J_{EM}^\mu \\
J_3^\mu &\equiv \frac{1}{2}\overline{\mathcal{U}}_L\gamma^\mu\mathcal{U}_L - \frac{1}{2}\overline{\mathcal{D}}_L\gamma^\mu\mathcal{D}_L \\
J_{EM}^\mu &\equiv +\frac{2}{3}\overline{\mathcal{U}}\gamma^\mu\mathcal{U}_L - \frac{1}{3}\overline{\mathcal{D}}\gamma^\mu\mathcal{D}_L
\end{aligned}$$

Having worked out the interactions of a single quark generation, we now examine what happens when there is more than one of them. Let us consider the fields $q_{\alpha iI}$, $\bar{u}_I$ and $\bar{d}_I$, where $I = 1, 2, 3$ is a generation index. The kinetic term for all these fields is

$$\mathcal{L}_{kin} = iq^{\dagger\alpha iI}\overline{\sigma}^\mu(D_\mu)_{\alpha i}{}^{\beta j}q_{\beta jI} + i\bar{u}^\dagger_{\alpha I}\overline{\sigma}^\mu(D_\mu)^\alpha{}_\beta\bar{u}^\beta_I + i\bar{d}^\dagger_{\alpha I}\overline{\sigma}^\mu(D_\mu)^\alpha{}_\beta\bar{d}^\beta_I$$

where the repeated generation index is summed. The most general Yukawa term we can write down now reads

$$\mathcal{L}_{Yuk} = -\epsilon^{ij}\phi_i q_{\alpha jI}y'_{IJ}\bar{d}^\alpha_J - \phi^{\dagger i}q_{\alpha jI}y''_{IJ}\bar{u}^\alpha_J + \text{h.c.}$$

y' and y'' are complex 3×3 matrices. In unitary gauge, this becomes

$$\mathcal{L}_{Yuk} = -\frac{1}{\sqrt{2}}(v+H)d_{\alpha I}y'_{IJ}\bar{d}^\alpha_J - \frac{1}{\sqrt{2}}(v+H)u_{\alpha I}y''_{IJ}\bar{u}^\alpha_J + \text{h.c.}$$

We can make unitary transformations in generation space on the fields: $d_I \to D_{IJ}d_J$, $\bar{d}_I \to \overline{D}_{IJ}\bar{d}_J$, $u_I \to U_{IJ}u_J$ and $\bar{u}_I \to \overline{U}_{IJ}\bar{u}_J$, where U, $\overline{U}$, D, $\overline{D}$ are independent unitary matrices. The kinetic terms are unchanged (except for the couplings to the $W^\pm$, as we will discuss momentarily), and the Yukawa matrices y' and y'' are replaced with $D^T y'\overline{D}$ and $U^T y''\overline{U}$. We can choose U, $\overline{U}$, D, $\overline{D}$ so that $D^T y'\overline{D}$ and $U^T y''\overline{U}$ are diagonal with positive real entries y'_I and $y"_I$. The down quarks D_I then have masses $m_{dI} = \frac{y'_I v}{\sqrt{2}}$, and the up quarks u_I have masses $m_{uI} = \frac{y''_I v}{\sqrt{2}}$. In the neutral currents, we simply add a generation index I to each field, and sum over it. The charged currents are more complicated, however; they become

$$J^{+\mu} = \overline{\mathcal{D}}_{LI}(V^\dagger)_{IJ}\gamma^\mu\mathcal{U}_{LJ}$$

$$J^{-\mu} = \overline{\mathcal{U}}_{LI}V_{IJ}\gamma^\mu\mathcal{D}_{LJ}$$

where $V \equiv U^\dagger D$ is the Cabibbo–Kobayasi–Maskawa matrix (or CKM matrix for short).
A 3×3 unitary matrix has nine real parameters. However, we are still free to make the independent phase rotations $\mathcal{D}_I = e^{i\alpha_I}\mathcal{D}_I$ and $\mathcal{U}_I = e^{i\beta_I}\mathcal{U}_I$, as these leave the kinetic and mass terms invariant. These phase changes allow us to make the first row and column of V_{IJ} real, eliminating five of the nine parameters. The remaining four can be chosen as θ_1 (the Cabibbo angle),θ_2 ,θ_3 , and δ, where

$$V = \begin{pmatrix} c_1 & +s_1c_3 & +s_1s_3 \\ -s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ -s_1s_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix}$$

and $c_i = \cos\theta_i$ and $s_i = \sin\theta_i$. The measured values of these angles are $s_1 = 0.224$, $s_2 = 0.041$, $s_3 = 0.016$, and $\delta = 40°$. Note that the charged currents now have some terms with a phase factor $e^{i\delta}$, and some without. Since the time-reversal operator T is antiunitary, the charged currents do not transform in a simple way under time reversal. This implies that the charged current terms in $\mathcal{L}_{int}$ are not time-reversal invariant; hence the electroweak interactions violate time-reversal symmetry. Since CPT is always a good symmetry, time-reversal violation is equivalent to CP violation; δ is therefore sometimes called the CP violating phase.

Part VI

Statistical mechanics

Chapter 27
Thermodynamics

27.1 Introduction to thermodynamics

A thermodynamic system is a macroscopic system whose behaviour is identified thanks to a small and finite number of quantities - the thermodynamic properties.
There is a certain degree of circularity in the definition of thermodynamic parameters, which is resolved by experiment. One considers only a restricted set of manipulations on thermodynamic systems. In practice, one allows them to be put in contact with one another, or one acts upon them by changing a few macroscopic properties such as their volume or the electric or magnetic field in which they are immersed. One then identifies a number of properties such that, if they are known before the manipulation, their values after the manipulation can be predicted. The smallest set of properties that allows one to successfully perform such a prediction can be selected as the basis for a thermodynamic description of the system.

If the state of a thermodynamic system can be fully characterized by the values of the thermodynamic variables, and if these values are invariant over time, one says that it is in a state of thermodynamic equilibrium. Thermodynamic equilibrium occurs when all fast processes have already occurred, while the slow ones have yet to take place. Clearly the distinction between fast and slow processes is dependent on the observation time τ that is being considered.

A system can be shown to be in equilibrium if the observation time is fairly short, while it is no longer possible to consider it in equilibrium for longer observation times. A more curious situation is that the same system can be considered in equilibrium, but with different properties, for different observation times.

Let us consider two thermodynamic systems, 1 and 2, that can be made to interact with one another. Variables like the volume V, the number of particles N, and the internal energy U, whose value (relative to the total system) is equal to the sum of the values they assume in the single systems, are called additive or extensive. Strictly speaking, internal energy is not extensive, unless the interaction between 1 and 2 can be neglected.

The fundamental hypothesis of thermodynamics is that it should be possible to characterize the state of a thermodynamic system by specifying the values of a certain set $(X_0, X_1, \cdots, X_r)$ of extensive variables. For example, X_0 could be the internal energy U, X_1 the number of particles N, X_2 the volume V of the system, and so on.
The central problem of thermodynamics is that **given the initial state of equilibrium of sev-**

eral thermodynamic systems that are allowed to interact, determine the final thermodynamic state of equilibrium.

The interaction between thermodynamic systems is usually represented by idealized walls that allow the passage of one (or more) extensive quantities from one system to the other. Among the various possibilities, the following are usually considered:

- **Thermally conductive walls** These allow the passage of energy, but not of volume or particles.
- **Semipermeable walls** These allow the passage of particles belonging to a given chemical species (and consequently also of energy)

The space of possible states of equilibrium (compatible with constraints and initial conditions) is called the space of virtual states. The initial state is obviously a (specific) virtual state. The central problem of thermodynamics can obviously be restated as follows:
Characterize the actual state of equilibrium among all virtual states.

27.2 Entropy formulation of thermodynamics

27.2.1 Property of entropy function

There exists a function S of the extensive variables $(X_0, X_1, \cdots, X_r)$, called the entropy, that assumes the maximum value for a state of equilibrium among all virtual states and that possesses the following properties:

1. **Extensivity**: If 1 and 2 are thermodynamic systems, then

$$S^{(1\cup 2)} = S^{(1)} + S^{(2)}$$

2. **Convexity**: If $X^1 = (X_0^1, X_1^1, \cdots, X_r^1)$ and $X^2 = (X_0^2, X_1^2, \cdots, X_r^2)$ are two thermodynamic states of the same system, then for any a between 0 and 1, one obtains

$$S[(1-\alpha)X^1 + \alpha X^2] \geq (1-\alpha)S(X^1) + \alpha S(X^2)$$

From this expression, if we take the derivative with respect to α at $\alpha = 0$, we obtain

$$\sum_{i=0}^{r} \left.\frac{\partial S}{\partial X_i}\right|_{X^1} (X_i^2 - X_i^1) \geq S(X^2) - S(X^1)$$

which expresses the fact that the surface $S(X_0, X_1, \cdots, X_r)$ is always below the plane that is tangent to each of its points. (We adpot the convention that convex means upper convex).

3. **Monotonicity**: $S(U, X_1, \cdots, X_r)$ is a monotonically increasing function of the internal energy U:

$$\left.\frac{\partial S}{\partial U}\right|_{X_1,\cdots,X_r} = \frac{1}{T} > 0$$

The entropy postulate allows one to solve the central problem of thermodynamics, by referring it back to the solution of a constrained extremum problem: **The equilibrium state corresponds to the maximum entropy compatible with the constraints.**

27.2.2 Simple problems

Thermal Contact

Let us consider two systems, 1 and 2, that are in contact by means of a thermally conductive wall. The virtual state space is therefore defined by the relations:

$$U^{(1)} + U^{(2)} = U = \text{const.}$$

$$X_i^{(1)} = \text{const.} \quad X_i^{(2)} = \text{const.}. \quad r = 1, \cdots, r$$

Let us look for the maximum of S as a function of $U^{(1)}$:

$$\frac{\partial S}{\partial U^{(1)}} = \left.\frac{\partial S^{(1)}}{\partial U^{(1)}}\right|_{U^{(1)}} - \left.\frac{\partial S^{(2)}}{\partial U^{(2)}}\right|_{U-U^{(1)}}$$

If we denote the value of $U^{(1)}$ at equilibrium by $U_{\text{eq}}^{(1)}$, then we have

$$\left.\frac{\partial S^{(1)}}{\partial U^{(1)}}\right|_{U_{\text{eq}}^{(1)}} = \left.\frac{\partial S^{(2)}}{\partial U^{(2)}}\right|_{U_{\text{eq}}^{(2)}}$$

Due to entropy's convexity, we can futher derive that

$$\left[\left.\frac{\partial S^{(1)}}{\partial U^{(1)}}\right|_{U_{\text{in}}^{(1)}} - \left.\frac{\partial S^{(2)}}{\partial U^{(2)}}\right|_{U_{\text{in}}^{(2)}}\right](U_{\text{eq}}^{(1)} - U_{\text{in}}^{(1)}) \geq 0$$

Let us introduce the quantity

$$T = \left(\frac{\partial S}{\partial U}\right)^{-1}$$

According to our hypotheses, this quantity is positive. We obtained the following results:

- At equilibrium, T is the same in all subsystems that are in reciprocal contact by means of thermally conductive walls.
- In order to reach equilibrium, energy shifts from systems with higher values of T toward systems with lower values of T.

Later, we will show that T is the temperature of the system.

A Thermally Conductive and Mobile Wall

In this case, the two systems can also exchange volume V, in addition to internal energy U. If we introduce the quantity p by

$$\frac{p}{T} = \frac{\partial S}{\partial V}$$

The two equilibrium conditions are

$$T^{(1)} = T^{(2)} \quad p^{(1)} = p^{(2)}$$

One can easily prove that between two systems, both initially at the same temperature, volume is initially released by the system in which p is lower to the system in which p is higher. Later, we will show that p is the pressure of the system.

A Semipermeable Wall

Let us consider a system composed of several chemical species, and let us introduce the number of molecules $N_1, \cdots, N_r$ belonging to the chemical species that constitute it as part of the thermodynamic variables. Let us suppose that two systems of this type are separated by a wall that only allows the k-th chemical species to pass. Clearly, it is impossible for the exchange of molecules to occur without an exchange of energy. If we introduce the quantity μ_i by

$$\frac{\mu_i}{T} = \frac{\partial S}{\partial N_i}$$

The equilibrium conditions will therefore be

$$T^{(1)} = T^{(2)} \quad \mu^{(1)} = \mu^{(2)}$$

We will define μ_i as the chemical potential of the specie i.

27.2.3 Heat and Work

From mechanics (and from electromagnetism), we can derive an expression for the infinitesimal mechanical work performed on the system by varying the extensive quantities. One usually adopts a sign convention according to which work is considered positive if the system performs work on the outside. Following this convention, the expression of infinitesimal work is given by

$$\delta W = -\sum_{i=1}^{r} f_i dX_i$$

On the one hand, we have

$$dS = \frac{dU}{T} + \sum_{i=1}^{r} \frac{\partial S}{\partial X_i}\bigg|_{U,\cdots,X_r} dX_i$$

It can be written as

$$dU = TdS - \sum_{i=1}^{r} T\frac{\partial S}{\partial X_i}\bigg|_{U,\cdots,X_r} dX_i$$

On the other hand, we have

$$dU = \delta Q - \delta W$$

So we can get

$$\delta Q = TdS \qquad \frac{\partial S}{\partial X_i}\bigg|_{U,\cdots,X_r} = -\frac{f_i}{T}$$

Temperature

Let us consider a system made up of a thermal engine and two heat reservoirs with $T_1 > T_2$. A heat reservoir is a system for which T is independent of U. The whole compound system is enclosed in a container that allows it to exchange energy with the environment only in a purely mechanical way.
Let the system evolve from an initial equilibrium condition, in which the first heat reservoir has internal energy U_1, the second has internal energy U_2 , and the thermal engine is in some equilibrium state, to a final equilibrium state in which the first heat reservoir has internal energy U_1', the second has U_2'. So $W = (U_1 + U_2) - (U_1' + U_2')$, and the thermal engine is back to its initial state. By definition, the efficiency of the engine is given by $\eta = \frac{W}{U_1 - U_1'}$.
In a transformation of this kind, the total entropy of the compound system cannot become smaller. Thus, we have

$$S^{(1)}(U_1) + S^{(2)}(U_2) \leq S^{(1)}(U_1') + S^{(2)}(U_2')$$

On the other hand, since we are dealing with heat reservoirs, we have

$$S^{(i)}(U_i') = S^{(i)}(U_i) + \frac{U_i' - U_i}{T_i} \quad i = 1, 2$$

Thus, we have

$$\frac{U_1 - U_1'}{T_1} \leq \frac{U_2' - U_2}{T_2}$$

Therefore,

$$\eta \leq 1 - \frac{T_2}{T_1}$$

Compared with the maximum efficiency evaluated in elementary thermodynamics, we can conclude that T is the absolute temperature, up to an overall factor, which can be fixed to 1 by rescaling the S.

Pressure

Let us consider an infinitesimal variation of V. In this case, mechanics tells us that the work performed by the system is given by $\delta W = PdV$. So we have

$$\left.\frac{\partial S}{\partial V}\right|_{U,\cdots,X_r} = \frac{P}{T}$$

This allows us to identify the pressure P with the quantity p we defined previousl.

The Fundamental Equation

The equation

$$S = S(X_0 = U, X_1, \cdots, X_r)$$

is called the fundamental equation, and it represents a complete description of the thermodynamics of the system being considered.

27.3 Thermodynamic potential

27.3.1 Energy Scheme

We can also use a different (but equivalent) formulation of the fundamental principle of thermodynamics, in which entropy assumes the role of an independent variable, while energy becomes a dependent variable that satisfies a variational principle. This formalism is known as the energy scheme. In this formalism, the maximum entropy principle is replaced by the principle of minimum internal energy:
Among all states with a specific entropy value, the state of equilibrium is that in which internal energy is minimal.

Let ΔX be a virtual variation of the extensive variables (excluding internal energy U) with respect to the equilibrium value X_{eq}. Then

$$\Delta S = S(U, X_{\text{eq}} + \Delta X) - S(U, X_{\text{eq}}) \leq 0$$

Since S is a monotonically increasing function of U, there exists a value $U' > U$ such that $S(U', X_{\text{eq}} + \Delta X) = S(U, X_{\text{eq}})$. Therefore, if S is kept constant, as the system moves out of equilibrium, U cannot but increase. This is what we wanted to prove.
Therefore, the fundamental equation in the energy scheme is

$$U = U(S, X_1, \cdots, X_r)$$

U's differential assumes the form

$$dU = TdS + \sum_{i=1}^{r} f_i dX_i$$

Further more, we can derive that

$$U[(1-\alpha)Y^1 + \alpha Y^2] \leq (1-\alpha)U(Y^1) + \alpha U(Y^2)$$

which implies that the internal energy function is concave.

27.3.2 Intensive Variables and Thermodynamic Potentials

The derivatives $f_i = \frac{\partial U}{\partial X_i}$ of the internal energy U with respect to extensive quantities S, $\{X_i\}$ are called intensive quantities. For uniformity's sake, we define $f_0 \equiv \frac{\partial U}{\partial X_i} = T$. A given quantity f_i is called the conjugate of the corresponding variable X_i, and vice versa. The temperature T and entropy S are therefore conjugates, as are the pressure (with the opposite sign) $-P$ and the volume V.

Since both U and X_i are extensive, in a homogeneous system, intensive variables are not dependent on system size. Moreover, if a system is composed of several subsystems that can exchange the extensive quantity X_i, the corresponding intensive variable f_i assumes the same value in those subsystems that are in contact at equilibrium.

We now want to identify the state of equilibrium among all states that exhibit a given value of an intensive variable f_i. Specifically, for $i = 0$, we are confronted with the case of system with a fixed temperature, i.e. heat reservoir.

Let us now define the Helmholtz free energy $F(T, X)$ by the relation

$$F(T, X) = U(S(T, X), X) - TS(T, X)$$

where $X = \{X_1, \cdots, X_r\}$. We can then show that the thermodynamical equilibrium in these conditions is characterized by the following variational principle:
The value of the Helmholtz free energy is minimal for the equilibrium state among all virtual states at the given temperature T.

Let us now consider more generally the Legendre transform of the internal energy U with respect to the intensive variable f_i:

$$\Phi(S, f_1, X_2, \cdots, X_r) = U(S, X_1(S, f_1, X_2, \cdots, X_r), X_2, \cdots, X_r) - f_1 X_1(S, f_1, X_2, \cdots, X_r)$$

where $X_1(S, f_1, X_2, \cdots, X_r), X_2, \cdots, X_r)$ is determined by the condition

$$f_1 = \left.\frac{\partial U}{\partial X_1}\right|_{S, X_2, \cdots, X_r}$$

Then, the state of equilibrium is specified by the following criterion:
Among all the states that have the same value as f_1, the state of equilibrium is that which corresponds to the minimum value of Φ.

Let us observe that the partial derivative of Φ, performed with respect to f_1, with the other extensive variables kept fixed, yields the value of the extensive variable X_1:

$$\left.\frac{\partial \Phi}{\partial f_1}\right|_{S, X_2, \cdots, X_r} = -X_1(S, f_1, X_2, \cdots, X_r)$$

Nothing prevents us from considering two or more intensive variables as fixed—for example, $f_0 = T$ and f_1. Similar considerations will then lead us to introduce the thermodynamic potential $\Phi(T, f_1, X_2, \cdots, X_r)$, obtained as a Legendre transform of U with respect to S and X_1:

$$\Phi(T, f_1, X_2, \cdots, X_r) = U - TS - f_1 X_1$$

This thermodynamic potential assumes at equilibrium the minimum value among all the states with the same values of T and f_1. We can therefore obtain a whole series of thermodynamic potentials, by using a Legendre transform with respect to the extensive variables X_i. We cannot however eliminate all extensive variables in this manner. We will see later that if we did this, the resulting thermodynamic potential would identically vanish. For the time being, it is sufficient to observe that the Φ potentials are extensive quantities, and one cannot see how they could

be a function only of intensive quantities like the f_i.
A general thermodynamic potential

$$\Phi(S, f_1, \cdots, f_k, X_{k+1}, \cdots, X_r) = U - \sum_{i=1}^{k} f_i X_i$$

is concave as a function of the remaining extensive variables, for fixed values of the intensive variables $f_1, \cdots, f_k$. Φ on the other hand is convex as a function of the intensive variables f_is, when the extensive variables are fixed.

27.3.3 Free Energy and Maxwell Relations

The natural variables of F are the temperature T and the extensive variables $X_1, \cdots, X_r$, entropy excluded. Entropy can be obtained by taking the derivative of F with respect to T. The expression for the differential of F is

$$dF = -SdT + \sum_{i=1}^{r} f_i dX_i$$

More specifically, by setting $X_1 = V$, one has

$$-P = \left.\frac{\partial F}{\partial V}\right|_{T, X_2, \cdots, X_r}$$

If we now take this equation's derivative with respect to T and we use the theorem of the equality of mixed derivatives, we obtain

$$-\left.\frac{\partial P}{\partial T}\right|_{V, X_2, \cdots, X_r} = -\frac{\partial^2 F}{\partial T \partial V} = -\left.\frac{\partial S}{\partial V}\right|_{T, X_2, \cdots, X_r}$$

These relations between thermodynamic derivatives that derive from the equality of mixed derivatives of thermodynamic potentials are called Maxwell relations.
The free energy designation is derived from the following property. If a system is put in contact with a reservoir at temperature T, we can prove the maximum quantity of work $W_{\max}$ that it can perform on its environment is equal to the variation in free energy between the initial and final states. In other words, one has

$$W \leq F_{\text{in}} - F_{\text{fin}}$$

27.3.4 Gibbs Free Energy and Enthalpy

Transforming F according to Legendre with respect to V, we obtain a new thermodynamic potential, called the Gibbs free energy:

$$G(T, P, X_2, \cdots, X_r) = F + PV = U - TS - PV$$

The variational principle satisfied by the Gibbs free energy is the following:
Among all states that have the same temperature and pressure values, the state of equilibrium is that in which the Gibbs free energy assumes the minimum value.
G's differential is expressed as follows:

$$dG = -SdT + VdP + \sum_{i=2}^{r} f_i dX_i$$

We can prove that if a system is brought toward equilibrium while temperature and pressure are kept constant, the maximum work that can be performed on its environment is given precisely by the difference between the initial and final values of G.

If on the other hand, we Legendre transform the internal energy U with respect to V, we obtain a new thermodynamic potential, usually denoted by H and called enthalpy:

$$H(S, P, X_2, \cdots, X_r) = U + PV$$

Enthalpy governs the equilibrium of adiabatic processes that occur while pressure is constant:
Among all states that have the same entropy and pressure values, the state of equilibrium is the one that corresponds to the minimum value of enthalpy.

If a system relaxes toward equilibrium while the pressure is kept constant, the maximum heat that can be produced by the system is equal to its variation in enthalpy. For this reason, enthalpy it is also called free heat. The differential of H is given by $dH = TdS + VdP + \sum_{i=2}^{r} f_i dX_i$

The equality of the mixed derivatives of G and H yield two more Maxwell relations:

$$-\left.\frac{\partial S}{\partial P}\right|_{T,X_2,\cdots,X_r} = \left.\frac{\partial V}{\partial T}\right|_{P,X_2,\cdots,X_r}$$

$$\left.\frac{\partial T}{\partial P}\right|_{S,X_2,\cdots,X_r} = \left.\frac{\partial V}{\partial S}\right|_{P,X_2,\cdots,X_r}$$

27.3.5 Other Thermodynamic Potentials

The Legendre transform of F with respect to N produces a thermodynamic potential (often written as Ω) that depends on T, on volume V, on chemical potential μ, and on the other extensive variables:

$$\Omega(T, V, \mu) = F - \mu N$$

Its differential is expressed as follows:

$$d\Omega = -SdT - PdV - Nd\mu$$

If one transforms U instead, one obtains a rarely used potential that depends on S, V, and μ, which we will designate as Φ:

$$\Phi(S, V, \mu) = U - \mu N$$

Its differential is given by

$$d\Phi = TdS - PdV - Nd\mu$$

We have the following Maxwell relations

$$\left.\frac{\partial S}{\partial \mu}\right|_{T,V} = \left.\frac{\partial N}{\partial T}\right|_{\mu,V}$$

$$\left.\frac{\partial S}{\partial N}\right|_{T,V} = -\left.\frac{\partial \mu}{\partial T}\right|_{N,V}$$

27.4 The Euler and Gibbs-Duhem Equations

$$U(\lambda S, \lambda X_1, \cdots, \lambda X_r) = \lambda U(S, X_1, \cdots, X_r)$$

By taking the derivative of this equation with respect to λ and setting $\lambda = 1$, we obtain the Euler equation

$$TS + \sum_{i=1}^{r} f_i X_i = U$$

More particularly, for simple fluids, one obtains

$$U = TS - PV + \mu N$$

which among other things implies that

$$\mu = (U - TS + PV)/N = G/N$$

$$\Omega = U - TS - \mu N = -PV$$

More particularly, from the Euler equation, it follows that the Legendre transform of U with respect to all extensive variables vanishes identically. Let us note that the interpretation of the chemical potential as a per particle density of Gibbs free energy is valid only in the case of simple fluids—in the case of a mixture of several chemical species, it is no longer valid. If we take the Euler equation's differential, we obtain

$$dU = TdS + SdT + \sum_{i=1}^{r} f_i dX_i + X_i df_i$$

By subtracting both sides of this equation from the usual expression of dU, we obtain the Gibbs-Duhem equation:

$$SdT + \sum_{i=1}^{r} X_i df_i = 0$$

In the case of simple fluids, for example, one arrives at

$$SdT - VdP + Nd\mu = 0$$

By dividing with respect to the number of particles N, one obtains the Gibbs-Duhem equation in the form

$$d\mu = vdP - sdT$$

where v represents volume per particle and s entropy per particle.

Relations between the densities and the intensive variables obtained by deriving the fundamental equation are called equations of state. If, for example, we consider the Gibbs free energy for a simple fluid, we arrive at

$$V = \left.\frac{\partial G}{\partial P}\right|_{T,N}$$

from which we obtain

$$v = \frac{V}{N} = v(P,T)$$

where we have made use of the fact that G is extensive, and therefore proportional to N. In the case of the simple fluid, we have another independent equation of state:

$$s = \frac{S}{N} = -\frac{1}{N}\left.\frac{\partial G}{\partial T}\right|_{P}$$

which expresses the entropy per particle s as a function of P and T. In reality, the two equations of state are not completely independent, because of the Maxwell relations:

$$-\left.\frac{\partial s}{\partial P}\right|_{T} = \left.\frac{\partial v}{\partial T}\right|_{P}$$

The ideal gas is a simple fluid that satisfies the equation of state

$$P = \frac{NT}{V}$$

Maxwell relations allow us to prove that in an ideal gas, the internal energy only depends on N and T (and not on the volume V). Moreover, entropy is the sum of a term that depends only on T with one that depends only on V.

27.5 Thermodynamic systems with multi-components

27.5.1 Chemical Reactions

Let us now consider a mixture of r chemical species, $A_1, \cdots, A_r$, which can be transformed into one other by a reaction of the following type:

$$\nu_1 A_1 + \cdots + \nu_k A_k \leftrightarrows \nu_{k+1} A_{k+1} + \cdots + \nu_r A_r$$

We can conventionally assign negative stoichiometric coefficients to the products, so as to write this formula as a formal equation:

$$\sum_{i=1}^{r} \nu_i A_i = 0$$

If temperature and pressure are kept fixed, we can calculate the variation of Gibbs free energy for a certain variation in the number of particles due to the reaction

$$\delta G = \sum_i \left.\frac{\partial G}{\partial N_i}\right|_{P,T} \delta N_i \propto \sum_i \left.\frac{\partial G}{\partial N_i}\right|_{P,T} \nu_i = \sum_i \mu_i \nu_i$$

Since at equilibrium one must have $\delta G = 0$ for any virtual variation of the N_i, one will have

$$\sum_i \mu_i \nu_i = 0$$

where the stoichiometric coefficients ν_i are taken with their corresponding signs.

27.5.2 Phase Coexistence

It frequently happens that two systems characterized by different thermodynamic density values can maintain thermodynamic equilibrium even in the absence of constraints on the mutual exchange of extensive quantities. This situation is called phase coexistence.
In the case of a simple fluid, it is realized, for example, when a liquid coexists with its vapor inside a container. In this case, the intensive variables assume the same value in both systems, while densities assume different values. In these cases, we refer to each of the coexisting homogeneous systems as a phase.
One can describe phase coexistence by saying that the equation of state

$$v = v(P, T)$$

does not admit of a unique solution, but instead allows for at least the two solutions $v = v_{\text{liq}}$ and $v = v_{\text{vap}}$ which correspond to the liquid and vapor, respectively.
Since the liquid and vapor coexist and can exchange particles, the chemical potential of the liquid has to be equal to that of the vapor:

$$\mu_{\text{liq}}(P, T) = \mu_{\text{vap}}(P, T)$$

On the other hand, we know that for a simple fluid, the chemical potential is equal to the Gibbs free energy per particle. We can verify that the Gibbs free energy in the total system does not depend on the number of particles that make up the liquid and the vapor system:

$$G = G_{\text{liq}} + G_{\text{vap}} = N_{\text{liq}}\mu_{\text{liq}} + N_{\text{vap}}\mu_{\text{vap}} = (N_{\text{liq}} + N_{\text{vap}})\mu = N\mu$$

The volume per particle of the system is given by

$$v = \frac{V_{\text{liq}} + V_{\text{vap}}}{N} = \frac{N_{\text{liq}} v_{\text{liq}} + N_{\text{vap}} v_{\text{vap}}}{N} = x_{\text{liq}} v_{\text{liq}} + x_{\text{vap}} v_{\text{vap}}$$

where x_{liq} is the fraction of particles in the liquid and $x_{\text{vap}} = 1 - x_{\text{liq}}$ that of the particles in the vapor. As a consequence, the value of v lies somewhere between v_{liq} and v_{vap}.

27.5.3 The Clausius-Clapeyron Equation

If we consider a thermodynamic system as a function of its intensive variables, we can identify some regions in which the thermodynamic properties vary regularly with variations of their values. These regions represent thermodynamically stable phases and are limited by curves that represent phase transitions. The phase transitions can be discontinuous, like the phase coexistence we just discussed, or continuous. In the first case, the densities present a discontinuity at the transition, while in the second, they vary with continuity, even though their derivatives can exhibit some singularities. One often also employs the following terminology: discontinuous transitions are called first order transitions, while continuous ones are called second order transitions.
In the case of a simple fluid, it is possible to identify the transition curve within the plane of the intensive variables (P, T) - in other words, the curve $P = P_t(T)$ - from the condition of equality of the chemical potential μ between the two coexisting phases:

$$\mu_{\text{liq}}(P_t(T), T) = \mu_{\text{vap}}(P_t(T), T)$$

The pressure $P_t(T)$ at which the transition occurs is also called vapor pressure. It is possible to relate this curve locally with the discontinuity of densities at transition. To obtain this relation, let us take the total derivative of this equation with respect to T, along the transition line $P_t(T)$. We obtain

$$\left.\frac{\partial \mu_{\text{liq}}}{\partial P}\right|_T \frac{dP_t}{dT} + \left.\frac{\partial \mu_{\text{liq}}}{\partial T}\right|_P = \left.\frac{\partial \mu_{\text{vap}}}{\partial P}\right|_T \frac{dP_t}{dT} + \left.\frac{\partial \mu_{\text{vap}}}{\partial T}\right|_P$$

Therefore,

$$\frac{dP_t}{dT} = \frac{s_{\text{vap}} - s_{\text{liq}}}{v_{\text{vap}} - v_{\text{liq}}}$$

This equation, which can be applied to each case of phase coexistence, is called the Clausius - Clapeyron equation.

27.5.4 The Coexistence Curve

We can represent the phase diagram in the plane (v, T), in which the intensive variable T is accompanied by the density v, the volume per particle. In this manner, phase coexistence is represented by the existence of a forbidden region $v_{\text{liq}}(T) < v < v_{\text{vap}}(T)$ in the plane. Outside this region, it is possible to obtain any given value of v in a homogeneous system. Within this region, instead, the system separates into a liquid and a vapor phase. The x_{liq} fraction of particles in the liquid phase (and the analogous fraction in the vapor phase) are determined by the condition that the entire system's volume per particle be equal to v. One thus obtains

$$x_{\text{liq}} = \frac{v_{\text{vap}} - v}{v_{\text{vap}} - v_{\text{liq}}}$$

This result is known as the lever rule.

27.5.5 Coexistence of Several Phases

Let us now consider a mixture of particles belonging to r different chemical species. Let us suppose that we are looking for the coexistence of q phases. At equilibrium, all the intensive

variables must assume the same value in the coexisting phases. We will therefore have a specific value for the pressure and the temperature, and in addition the chemical potential of each species will have to assume the same value in all the different phases. If we denote the chemical potential of species i in phase α as μ_i^α , we will then have

$$\mu_i^\alpha = \mu_i \quad i = 1, \cdots, r \quad \alpha = 1, \cdots,$$

In this equation, μ_i is the shared value taken by the chemical potential of species i. We thus obtain $r(q-1)$ equations for $q(r-1)+2$ unknown values. These unknown values are P, T, and the $q(r-1)$ independent densities x_i^α of species i in phase α. Generically speaking, therefore, $f = 2 - q + r$ free parameters remain. For $f = 0$, coexistence will occur in isolated points of the phase diagram, for $f = 1$, along a line, and so on. The quantity f is called variance. It is called the Gibbs phase rule.

27.5.6 The Critical Point

Let us once again consider the simple fluid and observe that the coexistence of liquid and vapour cannot be obtained for temperatures higher than a certain temperature T_c, called the critical temperature. To be more precise, the transition curve ends at a point (P_c, T_c), where P_c is the critical pressure. For $T < T_c$, the difference $v_{\text{vap}} - v_{\text{liq}}$ tends continuously toward zero when T gets close to T_c - the discontinuity of thermodynamic densities tends to vanish, or in other words, the transition goes from being discontinuous to being continuous (and finally to disappear at higher temperatures).
The critical point is a thermodynamic state with exceptional properties. For example, since for $T < T_c$ and $P = P_t(T)$ one gets $\partial P/\partial V|_T = 0$ (within the coexistence curve), this relation must ultimately be valid also at the critical point - in other words, the system's compressibility diverges:

$$\chi = -\frac{1}{V}\frac{\partial V}{\partial P}\bigg|_T \to \infty \text{ for } T \to T_c$$

Various thermodynamic properties exhibit analogous singularities.

Chapter 28
Principles of Statistical Mechanics and Ensembles

28.1 Density matrix

In quantum mechanics, the state of a system is a vector in Hilbert space, denoted as $|\psi\rangle$. A physical observable is an operator on this Hilbert space, denoted as O. The expectation value of the measurement of the observable is $\langle\psi|O|\psi\rangle$. It is easy to verify that

$$\langle\psi|O|\psi\rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|O)$$

Now, consider a system whose space of states is the direct product of two subspace, i.e.

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

An arbitrary state can be decomposed as

$$|\psi\rangle = C_{iI}|i\rangle_A \otimes |I\rangle_B$$

So, we have

$$|\psi\rangle\langle\psi| = C_{iI}C^*_{jJ}|i, I\rangle\langle j, J|$$

We define the partial trace of $|\psi\rangle\langle\psi|$ on B as

$$\mathrm{Tr}_B(|\psi\rangle\langle\psi|) \equiv \sum_I C_{iI}C^*_{jI}|i\rangle\langle j|$$

It is an operator on $\mathcal{H}_A$. Now, suppose there is an observable which measures only on A, i.e.

$$\langle i, I|O|j, I\rangle = \delta_{IJ}\langle i|O_A|j\rangle$$

where O_A is an operator on $\mathcal{H}_A$. We can verify that

$$\langle\psi|O|\psi\rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|O) = \mathrm{Tr}_A\left[\mathrm{Tr}_B(|\psi\rangle\langle\psi|)O_A\right]$$

Now, if we take A as the system and B the environment, a piratical observable measures only on system. For any system which is coupled to environment, its state can be described by an operator

$$\rho = \mathrm{Tr}_{\mathrm{env}}(|\psi\rangle\langle\psi|)$$

So the expectation value of the measurement on the system is

$$\mathrm{Tr}[\rho O_{\mathrm{sys}}]$$

We can verify that

$$\mathrm{Tr}\rho = 1, \quad \rho^\dagger = \rho$$

and any eigenvalue of ρ must lie between 0 and 1. Suppose rho can be diagonalized as

$$\rho = \sum_i p_i |i\rangle\langle i|$$

We have

$$\mathrm{Tr}[\rho O] = p_i \langle i|O|i\rangle$$

So it is reasonable to assume p_i as the (classical) probability of the system in (pure) state $|i\rangle$. One fundamental postulate of statistical mechanics is that the entropy operator of the system is

$$\hat{S} = -\ln \rho$$

So, the expectation value of entropy is

$$S = \mathrm{Tr}[-\rho \ln \rho]$$

28.2 Statistical ensemble

28.2.1 Micro-canonical ensemble

Micro-canonical ensemble describes a system which is weakly coupled to the environment. The volume V and the number of the particles N are fixed. The energy of the system lies in a narrow range between $E - \Delta E$ and $E + \Delta E$. The total number of distinct microstates accessible to a system is then denoted by the symbol $\Gamma(V, N, E; \Delta)$ and, by assumption, any one of these microstates is as likely to occur as any other.
Accordingly, the density matrix in the energy representation will be of the form

$$\rho_{mn} = \rho_m \delta_{mn}$$

with

$$\rho_n = \begin{cases} 1/\Gamma \text{ for each of the accessible states} \\ 0 \text{ for all other states} \end{cases}$$

The entropy of the system is

$$S = \ln \Gamma$$

28.2.2 Canonical ensemble

Canonical ensemble describes a system which can exchange energy with the environment. The density matrix of the system is

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}[e^{-\beta H}]}$$

Now, we define

$$Z(\beta, V, N) \equiv \mathrm{Tr}[e^{-\beta H}] \quad F(\beta, V, N) \equiv -\ln Z/\beta$$

The energy of the system is

$$U = \mathrm{Tr}[\rho H] = -\left.\frac{\partial \ln Z}{\partial \beta}\right|_{V,N}$$

If we further define $T \equiv 1/\beta$, we have

$$U = F - T\left.\frac{\partial F}{\partial T}\right|_{V,N}$$

The entropy of the system is

$$S = \mathrm{Tr}[-\rho \ln \rho] = \frac{U - F}{T}$$

We can further derive that

$$\left.\frac{\partial U}{\partial S}\right|_{V,N} = T$$

Now, we can identify T as the absolute temperature and F as free energy in thermodynamics.

28.2.3 Grand-canonical ensemble

Grand canonical ensemble describes a system which can exchange energy and particles with the environment. The density matrix of the system is

$$\rho = \frac{e^{-\beta(H-\mu N)}}{\mathrm{Tr}[e^{-\beta(H-\mu N)}]}$$

Now, we define

$$Z_\Omega(\beta, V, \mu) \equiv \mathrm{Tr}[e^{-\beta(H-\mu N)}] \quad \Omega(\beta, V, N) \equiv -\ln Z_\Omega/\beta$$

The particle number of the system is

$$N = \mathrm{Tr}[\rho N] = \frac{1}{\beta}\left.\frac{\partial \ln Z_\Omega}{\partial \mu}\right|_{V,\beta} = -\left.\frac{\partial \Omega}{\partial \mu}\right|_{V,T}$$

We also have

$$U - \mu N = -\left.\frac{\partial \ln Z_\Omega}{\partial \beta}\right|_{V,\mu} = \Omega - T\left.\frac{\partial \Omega}{\partial T}\right|_{V,\mu}$$

The entropy of the system is

$$S = \mathrm{Tr}[-\rho \ln \rho] = \frac{U - \mu N - \Omega}{T}$$

We can further derive that

$$\left.\frac{\partial U}{\partial N}\right|_{V,S} = \mu$$

Now, we can identify μ as the chemical potential and Ω as grand canonical potential in thermodynamics.

28.3 Fluctuations

28.3.1 Canonical Ensemble

The density matrix for canonical ensemble is

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}[e^{-\beta H}]}$$

We have

$$\left.\frac{\partial \rho}{\partial \beta}\right|_{N,V} = -\rho H + \rho \mathrm{Tr}[\rho H]$$

Since $U = \mathrm{Tr}[\rho H]$, we can derive that

$$\left.\frac{\partial U}{\partial \beta}\right|_{N,V} = -\mathrm{Tr}[\rho H^2] + (\mathrm{Tr}[\rho H])^2 = -\langle E^2\rangle + \langle E\rangle^2 = -\langle(\Delta E)^2\rangle$$

So, the relative fluctuation of energy in canonical ensemble is

$$\frac{\sqrt{\langle(\Delta E)^2\rangle}}{\langle E\rangle} = \frac{T}{U}\sqrt{\left.\frac{\partial U}{\partial T}\right|_{N,V}} = T\frac{\sqrt{C_V}}{T} \sim O(N^{-1/2})$$

For large N (which is true for every statistical system) the relative r.m.s. fluctuation in the values of E is quite negligible.

28.3.2 Grand canonical ensemble

Density and energy fluctuations in the grand canonical ensemble is much more complicated. The detailed discussion can be found in section 4.5 of *Statistical Mechanics (R.K.Pathria & Paul D.Beale)*.
The density fluctuation is

$$\frac{\langle(\Delta n)^2\rangle}{\langle n\rangle^2} = \frac{T}{V}\kappa_T$$

where $n = \frac{N}{V}$ is the number density and $\kappa_T = -\frac{1}{v}\left.\frac{\partial v}{\partial P}\right|_T$ is the isothermal compressibility of the system.

Thus, the relative root-mean-square fluctuation in the particle density of the given system is ordinarily $O(N^{-1/2})$ and, hence, negligible. However, there are exceptions, like the ones met with in situations accompanying phase transitions. In those situations, the compressibility of a given system can become excessively large, as is evidenced by an almost "flattening" of the isotherms. In the region of phase transitions, especially at the critical points, we encounter unusually large fluctuations in the particle density of the system. Such fluctuations indeed exist and account for phenomena like critical opalescence. It is clear that under these circumstances the formalism of the grand canonical ensemble could, in principle, lead to results that are not necessarily identical to the ones following from the corresponding canonical ensemble. In such cases, it is the formalism of the grand canonical ensemble that will have to

be preferred because only this one will provide a correct picture of the actual physical situation.

The energy fluctuation in grand canonical ensemble is

$$\langle(\Delta E)^2\rangle = T^2 C_V + \left(\left.\frac{\partial U}{\partial N}\right|_{T,V}\right)^2 \langle(\Delta N)^2\rangle$$

The mean-square fluctuation in the energy of a system in the grand canonical ensemble is equal to the value it would have in the canonical ensemble plus a contribution arising from the fact that now the particle number N is also fluctuating. Again, under ordinary circumstances, the relative root-mean-square fluctuation in the energy density of the system would be practically negligible. However, in the region of phase transitions, unusually large fluctuations in the value of this variable can arise by virtue of the second term in the formula.

Chapter 29
Interaction-free Systems

29.1 General discussion

29.1.1 Bose-Einstein Statistics

If the system is composed of interaction-free Bosons, the state of the system can be denoted as

$$|n_1, n_2, \cdots, n_i, \cdots\rangle$$

where n_i is the number of particles in state $|i\rangle$. Here, we choose $|i\rangle$ to the energy eigenstate with energy ϵ_i.
If we adopt grand-canonical ensemble, we have

$$Z_\Omega = \mathrm{Tr}[e^{-\beta(H-\mu N)}] = \sum_{n_1,\cdots,n_i,\cdots} e^{-\beta\sum_i n_i(\epsilon_i-\mu)} = \prod_i \sum_{n_i=0}^{\infty} [e^{-\beta(\epsilon_i-\mu)}]^{n_i} = \prod_i \frac{1}{1-e^{-\beta(\epsilon_i-\mu)}}$$

So, the grand canonical potential of the system is

$$\Omega = -\beta^{-1}\ln Z_\Omega = T\sum_i \ln(1-e^{-\beta(\epsilon_i-\mu)})$$

Note the chemical potential of the system must satisfy that

$$\mu < \epsilon_0$$

where ϵ_0 is the energy of the ground state.
To derive further results, we prefer to introduce a parameter z, defined by the relation

$$z \equiv e^{\beta\mu}$$

referred to as the fugacity of the system.
The expectation value of particle number is

$$N = -\left.\frac{\partial\Omega}{\partial\mu}\right|_{T,V} = \sum_i \frac{1}{e^{\beta\epsilon_i}z^{-1}-1}$$

The expectation value of energy is

$$U = \sum_i \frac{\epsilon_i}{e^{\beta\epsilon_i}z^{-1}-1}$$

The expectation value of particle number on level i is

$$n_i = \frac{1}{e^{\beta\epsilon_i}z^{-1}-1}$$

29.1.2 Fermi-Dirac Statistics

If the system is composed of interaction-free Fermions, the state of the system can be denoted as

$$|n_1, n_2, \cdots, n_i, \cdots\rangle$$

where n_i is the number of particles in state $|i\rangle$ and $n_i = 0$ or 1. Here, we choose $|i\rangle$ to the energy eigenstate with energy ϵ_i.
If we adopt grand-canonical ensemble, we have

$$Z_\Omega = \mathrm{Tr}[e^{-\beta(H-\mu N)}] = \sum_{n_1,\cdots,n_i,\cdots} e^{-\beta\sum_i n_i(\epsilon_i-\mu)} = \prod_i \sum_{n_i=0}^{1} [e^{-\beta(\epsilon_i-\mu)}]^{n_i} = \prod_i 1 + e^{-\beta(\epsilon_i-\mu)}$$

So, the grand canonical potential of the system is

$$\Omega = -\beta^{-1}\ln Z_\Omega = -T\sum_i \ln(1 + e^{-\beta(\epsilon_i-\mu)})$$

The expectation value of particle number is

$$N = \sum_i \frac{1}{e^{\beta\epsilon_i}z^{-1} + 1}$$

The expectation value of energy is

$$U = \sum_i \frac{\epsilon_i}{e^{\beta\epsilon_i}z^{-1} + 1}$$

The expectation value of particle number on level i is

$$n_i = \frac{1}{e^{\beta\epsilon_i}z^{-1} + 1}$$

29.1.3 Boltzmann Statistics

If the system is composed of distinguishable interaction-free particles, the state of the system can be denoted as

$$|i_1, i_2, \cdots, i_N\rangle$$

where i_k is the state of k-th particle and N is the number of particle of the system.
If we adopt canonical ensemble, we have

$$Z = \mathrm{Tr}e^{-\beta H} = \sum_{i_k} e^{-\beta\sum_{k=1}^{N}\epsilon_{i_k}} = \prod_{k=1}^{N}\sum_i e^{-\beta\epsilon_i}$$

Now, we define

$$Z_1 \equiv \sum_i e^{-\beta\epsilon_i}$$

we have

$$Z = (Z_1)^N$$

For a particle in a box, we can derive that $Z_1 \propto V \propto N$, so the free energy of the system is

$$F = -NT \ln Z_1 = -NT \ln N + \cdots$$

However, the term $-NT \ln N$ is not extensible. This is called Gibbs paradox. It is raised because our assumption that particles are distinguishable is wrong. This can be amended if we demand that

$$Z = \frac{Z_1^N}{N!}$$

then the non-extensible term is eliminated.
If we adopt the grand-canonical ensemble, we have

$$Z_\Omega = \sum_{N=0}^{\infty} \frac{z^N (Z_1)^N}{N!} = \exp[Z_1 z]$$

So, the grand canonical potential of the system is

$$\Omega = -\beta^{-1} \ln Z_\Omega = -T \sum_i e^{-\beta(\epsilon_i - \mu)}$$

The expectation value of particle number is

$$N = \sum_i \frac{1}{e^{\beta\epsilon_i} z^{-1}}$$

The expectation value of energy is

$$U = \sum_i \frac{\epsilon_i}{e^{\beta\epsilon_i} z^{-1}}$$

The expectation value of particle number on level i is

$$n_i = \frac{1}{e^{\beta\epsilon_i} z^{-1}}$$

As we can see, in the limit of

$$e^{\beta\epsilon_i} \geq e^{\beta\epsilon_0} \gg z$$

Bose-Einstein, Fermi-Dirac and Boltzmann statistics are identical.

29.2 Ideal Boltzmann Gas

29.2.1 Molecules without internal motion

Let us study thermodynamic properties of ideal gas. Our basic assumptions are:

- The temperature is high enough so that interaction between molecules can be neglected, i.e. $e^{-\frac{V}{T}} \ll 1$.
- Bose-Einstein or Fermi-Dirac statistics of the molecules can be well approximated by Boltzmann statistics, i.e. $z \ll 1$.

Suppose the side length of the box is L, the momentum of the particle must be the form of

$$\frac{2\pi}{L}(n_x, n_y, n_z)$$

where n_x,n_y and n_z are integers.
So we have

$$Z_1 = \sum_{n_x,n_y,n_z} e^{-\frac{\beta}{2m}(\frac{2\pi}{L})^2(n_x^2+n_y^2+n_z^2)}$$

If the difference of adjacent energy level is much smaller than β^{-1}, i.e.

$$\frac{\beta}{2m}(\frac{2\pi}{L})^2 \ll 1$$

the summation can be approximately as an integral. In SI units, this condition can be written explicitly as

$$T \gg \frac{h^2}{2mk_BL^2} \sim 10^{-17}K$$

which is always satisfied in practice.
So we have

$$Z_1 = V\int \frac{d^3p}{(2\pi)^3}e^{-\frac{\beta p^2}{2m}} = \frac{V}{\lambda^3}$$

where $\lambda \equiv \frac{2\pi}{\sqrt{2\pi mT}}$

Note: *When calculating the integral above, the following formula may by useful:*

$$\int_0^\infty x^n e^{-a\,x^2}\,dx = \frac{\Gamma(\frac{n+1}{2})}{2\,a^{\frac{n+1}{2}}}$$

The free energy of the system is

$$F = -NT\ln(\frac{V}{\lambda^3}) + T\ln N! = -NT\ln(\frac{V}{N\lambda^3}) - NT$$

Calculation of other thermodynamic quantities is straightforward.

$$S = -\left.\frac{\partial F}{\partial T}\right|_{V,N} = N\ln\frac{V}{\lambda^3} + \frac{5}{2}N$$

$$U = F + TS = \frac{3}{2}NT$$

$$P = -\left.\frac{\partial F}{\partial V}\right|_{T,N} = \frac{NT}{V}$$

$$\mu = \left.\frac{\partial F}{\partial N}\right|_{T,V} = -T\ln\frac{V}{N\lambda^3}$$

$$C_V = \left.\frac{\partial U}{\partial T}\right|_{V,N} = \frac{3}{2}N$$

Note that

$$z = e^{\beta\mu} = \frac{N\lambda^3}{V}$$

We must have

$$\frac{\lambda^3}{v} \ll 1$$

to ensure that Boltzmann statistics is valid. In SI units, this condition can be written explicitly as

$$\frac{T^{3/2}}{\rho} \gg \frac{h^3}{(2\pi k_B)^{3/2} m^{5/2}} \sim O(1) \frac{\mathrm{K}^{3/2}}{\mathrm{kg/m^3}}$$

Usually, before this condition is violated, the interaction between molecules becomes important and the gas may transform to liquid already.

29.2.2 Molecules without internal motion

If the internal motion of molecules is taken into account, we have

$$Z_1 = \frac{V}{\lambda^3} j(T)$$

where

$$j(T) \equiv \sum_i g_i e^{-\beta \epsilon_i}$$

where ϵ_i is the energy associated with a state of internal motion, while g_i is the multiplicity of that state.
The contributions made by the internal motions of the molecules, over and above the translational degrees of freedom, follow straightforwardly from the function $j(T)$.
We obtain

$$F_{\text{int}} = -NT \ln j$$

$$\mu_{\text{int}} = -kT \ln j$$

$$S_{\text{int}} = N \left(\ln j + T \frac{\partial}{\partial T} \ln j \right)$$

$$U_{\text{int}} = NT^2 \frac{\partial}{\partial T} \ln j$$

$$(C_V)_{\text{int}} = N \frac{\partial}{\partial T} \left\{ T^2 \frac{\partial}{\partial T} \ln j \right\}$$

Thus, the central problem in this study is to derive an explicit expression for the function $j(T)$ from a knowledge of the internal states of the molecules. For this, we note that the internal state of a molecule is determined by

- the electronic state
- the state of the nuclei
- the vibrational state
- the rotational state

Rigorously speaking, these four modes of excitation mutually interact; in many cases, however, they can be treated independently of one another. We can then write

$$j(T) = j_{\text{elec}}(T) j_{\text{nuc}}(T) j_{\text{vib}}(T) j_{\text{rot}}(T)$$

with the result that the net contribution made by the internal motions to the various thermodynamic properties of the system is given by a simple sum of the four respective contributions. A detailed discussion on gaseous systems composed of molecules with internal motion can be found in section 6.5 from *Statistical Mechanics (R.K.Pathria & Paul D.Beale)*.

29.3 Ideal Bose Systems

Bose-Einstein condensation

For an ideal Bose gas, we have

$$\ln Z_\Omega = \frac{PV}{T} = -\sum_i \ln(1 - ze^{-\beta\epsilon_i})$$

and

$$N = \sum_i \frac{1}{z^{-1}e^{\beta\epsilon_i} - 1}$$

In view of the fact that, for large V, the spectrum of the single-particle states is almost a continuous one, the summations may be replaced by integrations.
We note that by replacing summation by integration we are inadvertently giving a weight zero to the energy level $\epsilon = 0$. This is wrong because in a quantum mechanical treatment we must give a statistical weight unity to each non-degenerate single-particle state in the system. It is, therefore, advisable to take this particular state out of the sum in question before carrying out the integration; a rigorous justification of this unusual step can be found in Appendix F from *Statistical Mechanics (R.K.Pathria & Paul D.Beale)*
We thus obtain

$$\frac{P}{T} = -\frac{2\pi}{(2\pi)^3}(2m)^{3/2}\int_0^\infty \epsilon^{1/2}\ln(1 - ze^{-\beta\epsilon})d\epsilon - \frac{1}{V}\ln(1 - z)$$

and

$$\frac{N}{V} = \frac{2\pi}{(2\pi)^3}(2m)^{3/2}\int_0^\infty \frac{\epsilon^{1/2}d\epsilon}{z^{-1}e^{\beta\epsilon} - 1} + \frac{1}{V}\frac{z}{1 - z}$$

For $z \ll 1$, which corresponds to situations not far removed from the classical limit, the last term of each equation is of order $1/N$ and, therefore, negligible.
However, as z increases and assumes values close to unity, the term $V^{-1}z/(1 - z)$, which is identically equal to N_0/V (N_0 being the number of particles in the ground state), can well become a significant fraction of the quantity N/V ; this accumulation of a macroscopic fraction of the particles into a single state leads to the phenomenon of Bose-Einstein condensation.
Nevertheless, since $z = N_0/(N_0 + 1)$, the term $-V^{-1}\ln(1 - z)$ is equal to $V^{-1}\ln(N_0 + 1)$, which is at most $O(N^{-1}\ln(N + 1))$; this term is, therefore, negligible for all values of z and

hence may be dropped altogether.

At last, we have

$$\frac{P}{T} = \frac{1}{\lambda^3} g_{5/2}(z)$$

and

$$\frac{N - N_0}{V} = \frac{1}{\lambda^3} g_{3/2}(z)$$

where $g_\nu(z)$ are Bose-Einstein functions defined by

$$g_\nu(z) \equiv \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1} dx}{z^{-1} e^x - 1} = z + \frac{z^2}{2^\nu} + \cdots$$

The internal energy of this system is given by

$$U = -\left. \frac{\partial \ln Z_\Omega}{\partial \beta} \right|_{z,V} = \frac{3T}{2} \frac{V}{\lambda^3} g_{5/2}(z)$$

Thus, quite generally, our system satisfies the relationship

$$U = \frac{3}{2} PV$$

When the fugacity of the system is close to 1, we have

$$1 \ll N_0$$

and

$$\frac{N - N_0}{V} = \frac{1}{\lambda^3} \zeta(3/2)$$

where $\xi(\nu) = g_\nu(1)$. This curious phenomenon of a macroscopically large number of particles accumulating in a single quantum state is generally referred to as the phenomenon of Bose-Einstein condensation.
The condition for the onset of Bose-Einstein condensation is

$$T < T_c \equiv \frac{2\pi}{m} \left\{ \frac{N}{V \zeta(3/2)} \right\}^{2/3}$$

Here, T_c denotes a characteristic temperature that depends on the particle mass m and the particle density N/V in the system. Accordingly, for $T < T_c$, the system may be looked on as a mixture of two "phases":

- a normal phase, consisting of $N_e = N(T/T_c)^{3/2}$ particles distributed over the excited states
- a condensed phase, consisting of $N_0 = N - N_e$ particles accumulated in the ground state

Pressure

Next, we examine the variation of P with T, keeping ν fixed. Now, for $T < T_c$, the pressure is given by

$$P = \frac{T}{\lambda^3}\zeta(5/2)$$

which is proportional to $T^{5/2}$ and is independent of v - implying infinite compressibility. At the transition point the value of the pressure is

$$P(T_c) = \frac{\zeta(5/2)}{\zeta(3/2)}\frac{NT_c}{V} \approx 0.5134\frac{NT_c}{V}$$

Thus, the pressure exerted by the particles of an ideal Bose gas at the transition temperature is about one-half of that exerted by the particles of an equivalent Boltzmannian gas. For $T > T_c$, the pressure is given by

$$P = \frac{N}{V}T\frac{g_{5/2}(z)}{g_{3/2}(z)}$$

As $T \to \infty$, the pressure approaches the classical value.

Specific heat

For $T < T_c$, the specific heat is

$$\frac{C_V}{N} = \frac{3}{2}\frac{V}{N}\zeta(5/2)\frac{d}{dT}(\frac{T}{\lambda^3}) = \frac{15}{4}\zeta(5/2)\frac{v}{\lambda^3}$$

which is proportional to $T^{3/2}$. At $T = T_c$, we have $C_V(T_c) \approx 1.925N$, which is significantly higher than the classical value 1.5. For $T > T_c$, we obtain an implicit formula. First of all,

$$\frac{C_V}{N} = \left[\frac{\partial}{\partial T}\left(\frac{3}{2}T\frac{g_{5/2}(z)}{g_{3/2}(z)}\right)\right]_v$$

Note the recurrence relation of Bose-Einstein function

$$z\frac{\partial g_\nu(z)}{\partial z} = g_{v-1}(z)$$

and recall that $g_{3/2}(z) = \lambda^3/v$, it is easy to get

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_v = -\frac{3}{2T}\frac{g_{3/2}(z)}{g_{1/2}(z)}$$

At last, we have

$$\frac{C_V}{N} = \frac{15}{4}\frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)}$$

In the limit $z \to 1$, the second term vanishes because of the divergence of $g_{1/2}(z)$, while the first term gives exactly the the same result as in the case where $T \to T_c$ from below. The specific heat is, therefore, continuous at the transition point.
Its derivative is, however, discontinuous, the magnitude of the discontinuity being

$$\left.\frac{\partial C_V}{\partial T}\right|_{T=T_c-0} - \left.\frac{\partial C_V}{\partial T}\right|_{T=T_c+0} = \frac{27N}{16\pi T_c}\zeta(3/2)^2 \approx 3.665\frac{N}{T_c}$$

Entropy

Finally, we examine the adiabats of the ideal Bose gas. For this, we need an expression for the entropy of the system. Making use of the thermodynamic formula $U - TS + PV = N\mu$, we get

$$s = \begin{cases} \frac{5}{2}\frac{g_{5/2}(z)}{g_{3/2}(z)} - \ln z & T > T_c \\ \frac{5}{2}\frac{v}{\lambda^3}\zeta(5/2) & T < T_c \end{cases}$$

Now, a reversible adiabatic process implies the constancy of s. For $T > T_c$, this implies the constancy of z as well and in turn the constancy of $v/lambda^3$. For $T \leq T_c$, it again implies the same. We thus obtain, quite generally, the following relationship between the volume and the temperature of the system when it undergoes a reversible adiabatic process:

$$vT^{3/2} = \text{ const.}$$

The corresponding relationship between the pressure and the temperature is

$$P/T^{5/2} = \text{ const.}$$

Eliminating T, we obtain

$$Pv^{5/3} = \text{ const.}$$

as the equation for an adiabat of the ideal Bose gas.

Note: *Incidentally, the foregoing results are exactly the same as for an ideal classical gas. There is, however, a significant difference between the two cases; that is, while the exponent $5/3$ is identically equal to the ratio of the specific heats C_P and C_V in the case of the ideal classical gas, it is not so in the case of the ideal Bose gas.*

In the mixed-phase region, the entropy of the gas may be written as

$$S = N_e \frac{5}{2}\frac{\zeta(5/2)}{\zeta(3/2)}$$

As expected, the N_0 particles that constitute the condensate do not contribute toward the entropy of the system, while the N_e particles that constitute the normal part contribute an amount of $\frac{5}{2}\frac{\zeta(5/2)}{\zeta(3/2)}$ per particle.

Chapter 30
Quantum field theory approach

Chapter 31
Phase Transitions and the Renormalization Group