Advanced Concepts of Theoretical Physics

Uwe-Jens Wiese

January 7, 2013

Contents

Chapter 1

Introduction

Theoretical physics is the attempt to describe Nature quantitatively using the language of mathematics. Indeed there is an interplay between mathematics and theoretical physics, with both disciplines enriching each other. Isaac Newton invented calculus in order to apply it to classical mechanics. Michael Faraday developed the field concept in order to describe electromagnetism. David Hilbert's abstract space of square integrable functions turned out to be the home of quantum mechanical wave functions. Paul Dirac "invented" the δ-function which motivated mathematicians to develop distribution theory. Abelian gauge symmetry was discovered as the basic principle underlying Maxwell's equations. Sophus Lie's abstract concepts of non-Abelian symmetry algebras and symmetry groups underlie Wolfgang Pauli's ideas about non-Abelian gauge symmetry, which were further developed by Robert Mills and Chen-Ning Yang and were later discovered to form the basis of the strong and electroweak interactions. Michael Atiyah and Isidore Singer investigated the topological properties of Dirac operators, which turned out to be essential in the context of anomalous symmetries in quantum field theory, which are present at the classical level but cannot be maintained in the quantization process. Michael Berry realized that topological features of wave functions play a central role in adiabatic processes in quantum mechanics and can be characterized by abstract Abelian and non-Abelian gauge fields in the space of slowly varying external parameters. This list could be extended much further, and the creative process of discovering new mathematics when thinking about Nature or using existing mathematics to describe its observed behavior is likely to continue for a long time.

In this course, we will familiarize ourselves with some advanced mathematical concepts that are at the basis of the modern theories of particle and condensed matter physics, as well as of some advanced topics in quantum mechanics. Although the subject is mathematical, we will by no means apply the rigor of a mathemati-

cian. Instead, we will concentrate on those aspects that are most relevant in physics applications. Consequently, we will avoid very abstract mathematical constructions in favor of a hands-on practitioner's approach. While we will aim at understanding the subject at a deep level, we will not attempt rigorous proofs, and we will sometimes even limit ourselves to applying specific procedures, for example, when we reduce products of $SU(n)$ representations, without proving their validity. When we discuss Lie algebras and Lie groups, as well as the topology of the corresponding group manifolds, we will focus on those aspects that are most important in physical applications. A systematic mathematical exposition of these subjects would go far beyond the scope of the current course. For example, the concept of the Killing vector field of a Lie group manifold is not developed, although it does have physical applications. Although the presented material is neither mathematically rigorous nor complete, it will hopefully provide the reader with a solid basis for better understanding the quantum field theories that underlie modern particle physics as well as some theories in condensed matter physics.

It is a most fascinating aspect of physics that, on all lengths scales, Nature can indeed be described mathematically. The most fundamental building blocks of matter, elementary particles such as electrons and photons, as well as quarks and gluons are precise embodiments of abstract mathematical concepts. Even on much coarser length scales the plateaus in the resistivity of a quantum Hall sample are most precisely quantized in units of Planck's quantum h and the elementary electric charge e. Even the evolution of the entire cosmos is controlled rather precisely by the differential equations derived from general relativity assuming the Friedmann-Lemaitre-Robertson-Walker metric. As I argue in an appendix to these lecture notes, it is the locality of space and time and the existence of vast hierarchies of distance scales that leads to the success of physics and thus "explains" why physics exists. In this course, we will familiarize ourselves with some important advanced concepts of theoretical physics, in order to be better prepared to understand the modern theories of particle and condensed matter physics, as well as some aspects of advanced quantum mechanics. This may even inspire us to "invent" our own new mathematics or to be creative when trying to describe some new aspect of Nature mathematically. Mathematics is a universal language spoken by curious minds and capable brains, and used for various purposes, in particular, to communicate the abstract beauty of Nature. Let us learn more about this language and let us keep in mind that theoretical physics is a man- and women-made endeavor that relies on our curiosity as much as on our mathematical capabilities.

Chapter 2

From Mechanics to Quantum Field Theory

This chapter provides a brief summary of the mathematical structure of quantum field theory. Classical field theories are discussed as a generalization of point mechanics to systems with infinitely many degrees of freedom — a given number per space point. Similarly, quantum field theories are just quantum mechanical systems with infinitely many degrees of freedom. In the same way as point mechanics systems, classical field theories can be quantized with path integral methods. The quantization of field theories at finite temperature leads to path integrals in Euclidean time. This provides us with an analogy between quantum field theory and classical statistical mechanics. We also mention the lattice regularization which has recently provided a mathematically satisfactory formulation of the standard model beyond perturbation theory.

2.1 From Point Mechanics to Classical Field Theory

Point mechanics describes the dynamics of classical nonrelativistic point particles. The coordinates of the particles represent a finite number of degrees of freedom. In the simplest case — a single particle moving in one spatial dimension — we are dealing with a single degree of freedom: the x-coordinate of the particle. The dynamics of a particle of mass m moving in an external potential $V(x)$ is described

by Newton's equation

$$m\partial_t^2 x = ma = F(x) = -\frac{dV(x)}{dx}. \tag{2.1.1}$$

Once the initial conditions are specified, this ordinary second order differential equation determines the particle's path $x(t)$, i.e. its position as a function of time. Newton's equation results from the variational principle to minimize the action

$$S[x] = \int dt\ L(x, \partial_t x), \tag{2.1.2}$$

over the space of all paths $x(t)$. The action is a functional (a function whose argument is itself a function) that results from the time integral of the Lagrange function

$$L(x, \partial_t x) = \frac{m}{2}(\partial_t x)^2 - V(x). \tag{2.1.3}$$

The Euler-Lagrange equation

$$\partial_t \frac{\delta L}{\delta(\partial_t x)} - \frac{\delta L}{\delta x} = 0, \tag{2.1.4}$$

is nothing but Newton's equation.

Classical field theories are a generalization of point mechanics to systems with infinitely many degrees of freedom — a given number for each space point $\vec{x}$. In this case, the degrees of freedom are the field values $\phi(\vec{x})$, where ϕ is some generic field. In case of a neutral scalar field, ϕ is simply a real number representing one degree of freedom per space point. A charged scalar field, on the other hand, is described by a complex number and hence represents two degrees of freedom per space point. The scalar Higgs field $\phi^a(\vec{x})$ (with $a \in \{1, 2\}$) in the standard model is a complex doublet, i.e. it has four real degrees of freedom per space point. An Abelian gauge field $A_i(\vec{x})$ (with a spatial direction index $i \in \{1, 2, 3\}$) — for example, the photon field in electrodynamics — is a neutral vector field with 3 real degrees of freedom per space point. One of these degrees of freedom is redundant due to the $U(1)_{em}$ gauge symmetry. Hence, an Abelian gauge field has two physical degrees of freedom per space point which correspond to the two polarization states of the massless photon. Note that the time-component $A_0(\vec{x})$ does not represent a physical degree of freedom. It is just a Lagrange multiplier field that enforces the Gauss law. A non-Abelian gauge field $A_i^a(\vec{x})$ is charged and has an additional index a. For example, the gluon field in chromodynamics with a color index $a \in \{1, 2, ..., 8\}$ represents $2 \times 8 = 16$ physical degrees of freedom per space point, again because of some redundancy due to the $SU(3)_c$ color gauge symmetry. The field that represents the W- and Z-bosons in the standard model has an index $a \in \{1, 2, 3\}$ and transforms under the gauge group $SU(2)_L$. Thus, it represents $2 \times 3 = 6$ physical degrees of

freedom. However, in contrast to the photon, the W- and Z-bosons are massive due to the Higgs mechanism and have three (not just two) polarization states. The extra degree of freedom is provided by the Higgs field.

The analog of Newton's equation in field theory is the classical field equation of motion. For example, for a neutral scalar field this is the Klein-Gordon equation

$$\partial_\mu \partial^\mu \phi = -\frac{dV(\phi)}{d\phi}. \tag{2.1.5}$$

Again, after specifying appropriate initial conditions it determines the classical field configuration $\phi(x)$, i.e. the values of the field ϕ at all space-time points $x = (t, \vec{x})$. Hence, the role of time in point mechanics is played by space-time in field theory, and the role of the point particle coordinates is now played by the field values. As before, the classical equation of motion results from minimizing the action

$$S[\phi] = \int d^4x \; \mathcal{L}(\phi, \partial_\mu \phi). \tag{2.1.6}$$

The integral over time in eq.(2.1.2) is now replaced by an integral over space-time and the Lagrange function of point mechanics gets replaced by the Lagrange density function (or Lagrangian)

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi). \tag{2.1.7}$$

A simple interacting field theory is the ϕ^4 theory with the potential

$$V(\phi) = \frac{m^2}{2}\phi^2 + \frac{\lambda}{4!}\phi^4. \tag{2.1.8}$$

Here m is the mass of the scalar field and λ is the coupling strength of its self-interaction. Note that the mass term corresponds to a harmonic oscillator potential in the point mechanics analog, while the interaction term corresponds to an anharmonic perturbation. As before, the Euler-Lagrange equation

$$\partial_\mu \frac{\delta L}{\delta(\partial_\mu \phi)} - \frac{\delta L}{\delta \phi} = 0, \tag{2.1.9}$$

is the classical equation of motion, in this case the Klein-Gordon equation. The analogies between point mechanics and field theory are summarized in table 2.1.

2.2 The Path Integral in Real Time

The quantization of field theories is most conveniently performed using the path integral approach. Here we first discuss the path integral in quantum mechanics

Point Mechanics	Field Theory
time t	space-time $x = (t, \vec{x})$
particle coordinate x	field value ϕ
particle path $x(t)$	field configuration $\phi(x)$
action $S[x] = \int dt\ L(x, \partial_t x)$	action $S[\phi] = \int d^4x\ \mathcal{L}(\phi, \partial_\mu \phi)$
Lagrange function $L(x, \partial_t x) = \frac{m}{2}(\partial_t x)^2 - V(x)$	Lagrangian $\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi)$
equation of motion $\partial_t \frac{\delta L}{\delta(\partial_t x)} - \frac{\delta L}{\delta x} = 0$	field equation $\partial_\mu \frac{\delta L}{\delta(\partial_\mu \phi)} - \frac{\delta L}{\delta \phi} = 0$
Newton's equation $\partial_t^2 x = -\frac{dV(x)}{dx}$	Klein-Gordon equation $\partial_\mu \partial^\mu \phi = -\frac{dV(\phi)}{d\phi}$
kinetic energy $\frac{m}{2}(\partial_t x)^2$	kinetic energy $\frac{1}{2}\partial_\mu \phi \partial^\mu \phi$
harmonic oscillator potential $\frac{m}{2}\omega^2 x^2$	mass term $\frac{m^2}{2}\phi^2$
anharmonic perturbation $\frac{\lambda}{4!}x^4$	self-interaction term $\frac{\lambda}{4!}\phi^4$

Table 2.1: *The dictionary that translates point mechanics into the language of field theory.*

— quantized point mechanics — using the real time formalism. A mathematically more satisfactory formulation uses an analytic continuation to so-called Euclidean time. This will be discussed in the next section.

The real time evolution of a quantum system described by a Hamilton operator H is given by the time-dependent Schrödinger equation

$$i\hbar \partial_t |\Psi(t)\rangle = H|\Psi(t)\rangle. \tag{2.2.1}$$

For a time-independent Hamilton operator the time evolution operator is given by

$$U(t', t) = \exp(-\frac{i}{\hbar} H(t' - t)), \tag{2.2.2}$$

such that

$$|\Psi(t')\rangle = U(t', t)|\Psi(t)\rangle. \tag{2.2.3}$$

Let us consider the transition amplitude $\langle x'|U(t', t)|x\rangle$ of a nonrelativistic point particle that starts at position x at time t and arrives at position x' at time t'. Using

$$\langle x|\Psi(t)\rangle = \Psi(x, t) \tag{2.2.4}$$

we obtain

$$\Psi(x', t') = \int dx\ \langle x'|U(t', t)|x\rangle \Psi(x, t), \tag{2.2.5}$$

i.e. $\langle x'|U(t',t)|x\rangle$ acts as a propagator for the wave function. The propagator is of physical interest because it contains information about the energy spectrum. When we consider propagation from an initial position x back to the same position we find

$$\begin{aligned}\langle x|U(t',t)|x\rangle &= \langle x|\exp(-\frac{i}{\hbar}H(t'-t))|x\rangle \\ &= \sum_n |\langle x|n\rangle|^2 \exp(-\frac{i}{\hbar}E_n(t'-t)). \end{aligned} \tag{2.2.6}$$

We have inserted a complete set, $\sum_n |n\rangle\langle n| = \mathbf{1}$, of energy eigenstates $|n\rangle$ with

$$H|n\rangle = E_n|n\rangle. \tag{2.2.7}$$

Hence, according to eq.(2.2.6), the Fourier transform of the propagator yields the energy spectrum as well as the energy eigenstates $\langle x|n\rangle$.

Inserting a complete set of position eigenstates we arrive at

$$\begin{aligned}\langle x'|U(t',t)|x\rangle &= \langle x'|\exp(-\frac{i}{\hbar}H(t'-t_1+t_1-t))|x\rangle \\ &= \int dx_1 \langle x'|\exp(-\frac{i}{\hbar}H(t'-t_1))|x_1\rangle \\ &\times \langle x_1|\exp(-\frac{i}{\hbar}H(t_1-t))|x\rangle \\ &= \int dx_1 \langle x'|U(t',t_1)|x_1\rangle\langle x_1|U(t_1,t)|x\rangle. \end{aligned} \tag{2.2.8}$$

It is obvious that we can repeat this process an arbitrary number of times. This is exactly what we do in the formulation of the path integral. Let us divide the time interval $[t,t']$ into N elementary time steps of size ε such that

$$t'-t = N\varepsilon. \tag{2.2.9}$$

Inserting a complete set of position eigenstates at the intermediate times $t_i, i \in \{1,2,...,N-1\}$ we obtain

$$\begin{aligned}\langle x'|U(t',t)|x\rangle &= \int dx_1 \int dx_2 ... \int dx_{N-1} \langle x'|U(t',t_{N-1})|x_{N-1}\rangle ... \\ &\times \langle x_2|U(t_2,t_1)|x_1\rangle\langle x_1|U(t_1,t)|x\rangle. \end{aligned} \tag{2.2.10}$$

In the next step we concentrate on one of the factors and we consider a single nonrelativistic point particle moving in an external potential $V(x)$ such that

$$H = \frac{p^2}{2m} + V(x). \tag{2.2.11}$$

Using the Baker-Campbell-Haussdorff formula and neglecting terms of order ε^2 we find

$$\begin{aligned}\langle x_{i+1}|U(t_{i+1},t_i)|x_i\rangle &= \langle x_{i+1}|\exp(-\frac{i\varepsilon p^2}{2m\hbar})\exp(-\frac{i\varepsilon}{\hbar}V(x))|x_i\rangle \\ &= \frac{1}{2\pi}\int dp\langle x_{i+1}|\exp(-\frac{i\varepsilon p^2}{2m\hbar})|p\rangle\langle p|\exp(-\frac{i\varepsilon}{\hbar}V(x))|x_i\rangle \\ &= \frac{1}{2\pi}\int dp\exp(-\frac{i\varepsilon p^2}{2m\hbar})\exp(-\frac{i}{\hbar}p(x_{i+1}-x_i)) \\ &\times \exp(-\frac{i\varepsilon}{\hbar}V(x_i)). \end{aligned} \tag{2.2.12}$$

The integral over p is ill-defined because the integrand is a very rapidly oscillating function. To make the expression well-defined we replace the time step ε by $\varepsilon - ia$, i.e. we go out into the complex time plane. After doing the integral we take the limit $a \to 0$. Still one should keep in mind that the definition of the path integral required an analytic continuation in time. One finds

$$\langle x_{i+1}|U(t_{i+1},t_i)|x_i\rangle = \sqrt{\frac{m}{2\pi i\hbar\varepsilon}}\exp(\frac{i}{\hbar}\varepsilon[\frac{m}{2}(\frac{x_{i+1}-x_i}{\varepsilon})^2 - V(x_i)]). \tag{2.2.13}$$

Inserting this back into the expression for the propagator we obtain

$$\langle x'|U(t',t)|x\rangle = \int \mathcal{D}x\exp(\frac{i}{\hbar}S[x]). \tag{2.2.14}$$

The action has been identified in the time continuum limit as

$$\begin{aligned} S[x] &= \int dt\ [\frac{m}{2}(\partial_t x)^2 - V(x)] \\ &= \lim_{\varepsilon\to 0}\sum_i \varepsilon[\frac{m}{2}(\frac{x_{i+1}-x_i}{\varepsilon})^2 - V(x_i)]. \end{aligned} \tag{2.2.15}$$

The integration measure is defined as

$$\int \mathcal{D}x = \lim_{\varepsilon\to 0}\sqrt{\frac{m}{2\pi i\hbar\varepsilon}}^N \int dx_1 \int dx_2 ... \int dx_{N-1}. \tag{2.2.16}$$

This means that we integrate over the possible particle positions for each intermediate time t_i. In this way we integrate over all possible paths of the particle starting at x and ending at x'. Each path is weighted with an oscillating phase factor $\exp(\frac{i}{\hbar}S[x])$ determined by the action. The classical path of minimum action has the smallest oscillations, and hence the largest contribution to the path integral. In the classical limit $\hbar \to 0$ only that contribution survives.

2.3 The Path Integral in Euclidean Time

As we have seen, it requires a small excursion into the complex time plane to make the path integral mathematically well-defined. Now we will make a big step into that plane and actually consider purely imaginary so-called Euclidean time. The physical motivation for this, however, comes from quantum statistical mechanics. Let us consider the quantum statistical partition function

$$Z = \mathrm{Tr}\exp(-\beta H), \tag{2.3.1}$$

where $\beta = 1/T$ is the inverse temperature. It is mathematically equivalent to the time interval we discussed in the real time path integral. In particular, the operator $\exp(-\beta H)$ turns into the time evolution operator $U(t', t)$ if we identify

$$\beta = \frac{i}{\hbar}(t' - t). \tag{2.3.2}$$

In this sense the system at finite temperature corresponds to a system propagating in purely imaginary (Euclidean) time. By dividing the Euclidean time interval into N time steps, i.e. by writing $\beta = Na/\hbar$, and again by inserting complete sets of position eigenstates we now arrive at the Euclidean time path integral

$$Z = \int \mathcal{D}x \exp(-\frac{1}{\hbar}S_E[x]). \tag{2.3.3}$$

The action now takes the Euclidean form

$$\begin{aligned} S_E[x] &= \int dt\ [\frac{m}{2}(\partial_t x)^2 + V(x)] \\ &= \lim_{a\to 0}\sum_i a[\frac{m}{2}(\frac{x_{i+1} - x_i}{a})^2 + V(x_i)]. \end{aligned} \tag{2.3.4}$$

In contrast to the real time case the measure now involves N integrations

$$\int \mathcal{D}x = \lim_{a\to 0}\sqrt{\frac{m}{2\pi\hbar a}}^N \int dx_1 \int dx_2 ... \int dx_N. \tag{2.3.5}$$

The extra integration over $x_N = x'$ is due to the trace in eq.(2.3.1). Note that there is no extra integration over $x_0 = x$ because the trace implies periodic boundary conditions in the Euclidean time direction, i.e. $x_0 = x_N$.

The Euclidean path integral allows us to evaluate thermal expectation values. For example, let us consider an operator $\mathcal{O}(x)$ that is diagonal in the position state basis. We can insert this operator in the path integral and thus compute its expectation value

$$\langle\mathcal{O}(x)\rangle = \frac{1}{Z}\mathrm{Tr}[\mathcal{O}(x)\exp(-\beta H)] = \frac{1}{Z}\int \mathcal{D}x\ \mathcal{O}(x(0))\exp(-\frac{1}{\hbar}S_E[x]). \tag{2.3.6}$$

Since the theory is translation invariant in Euclidean time one can place the operator anywhere in time, e.g. at $t = 0$ as done here. When we perform the low temperature limit, $\beta \to \infty$, the thermal fluctuations are switched off and only the quantum ground state $|0\rangle$ (the vacuum) contributes to the partition function, i.e. $Z \sim \exp(-\beta E_0)$. In this limit the path integral is formulated in an infinite Euclidean time interval, and describes the vacuum expectation value

$$\langle \mathcal{O}(x) \rangle = \langle 0|\mathcal{O}(x)|0\rangle = \lim_{\beta\to\infty} \frac{1}{Z} \int \mathcal{D}x \; \mathcal{O}(x(0)) \exp(-\frac{1}{\hbar} S_E[x]). \tag{2.3.7}$$

It is also interesting to consider 2-point functions of operators at different instances in Euclidean time

$$\begin{aligned} \langle \mathcal{O}(x(0))\mathcal{O}(x(t)) \rangle &= \frac{1}{Z} \mathrm{Tr}[\mathcal{O}(x) \exp(-Ht) \mathcal{O}(x) \exp(Ht) \exp(-\beta H)] \\ &= \frac{1}{Z} \int \mathcal{D}x \; \mathcal{O}(x(0))\mathcal{O}(x(t)) \exp(-\frac{1}{\hbar} S_E[x]). \end{aligned} \tag{2.3.8}$$

Again, we consider the limit $\beta \to \infty$, but we also separate the operators in time, i.e. we also let $t \to \infty$. Then the leading contribution is $|\langle 0|\mathcal{O}(x)|0\rangle|^2$. Subtracting this, and thus forming the connected 2-point function, one obtains

$$\lim_{\beta,t\to\infty} \langle \mathcal{O}(x(0))\mathcal{O}(x(t)) \rangle - |\langle \mathcal{O}(x) \rangle|^2 = |\langle 0|\mathcal{O}(x)|1\rangle|^2 \exp(-(E_1 - E_0)t). \tag{2.3.9}$$

Here $|1\rangle$ is the first excited state of the quantum system with an energy E_1. The connected 2-point function decays exponentially at large Euclidean time separations. The decay is governed by the energy gap $E_1 - E_0$. In a quantum field theory E_1 corresponds to the energy of the lightest particle. Its mass is determined by the energy gap $E_1 - E_0$ above the vacuum. Hence, in Euclidean field theory particle masses are determined from the exponential decay of connected 2-point correlation functions.

2.4 Spin Models in Classical Statistical Mechanics

So far we have considered quantum systems both at zero and at finite temperature. We have represented their partition functions as Euclidean path integrals over configurations on a time lattice of length β. We will now make a completely new start and study classical discrete systems at finite temperature. We will see that their mathematical description is very similar to the path integral formulation of quantum systems. Still, the physical interpretation of the formalism is drastically

different in the two cases. In the next section we will set up a dictionary that allows us to translate quantum physics language into the language of classical statistical mechanics.

For simplicity, let us concentrate on simple classical spin models. Here the word spin does not mean that we deal with quantized angular momenta. All we do is work with classical variables that can point in specific directions. The simplest spin model is the Ising model with classical spin variables $s_x = \pm 1$. (Again, these do not represent the quantum states up and down of a quantum mechanical angular momentum 1/2.) More complicated spin models with an $O(N)$ spin rotational symmetry are the XY model ($N = 2$) and the Heisenberg model ($N = 3$). The spins in the XY model are 2-component unit-vectors, while the spins in the Heisenberg model have three components. In all these models the spins live on the sites of a d-dimensional spatial lattice. The lattice is meant to be a crystal lattice (so typically $d = 3$) and the lattice spacing has a physical meaning. This is in contrast to the Euclidean time lattice that we have introduced to make the path integral mathematically well-defined, and that we finally send to zero in order to reach the Euclidean time continuum limit. The Ising model is characterized by its classical Hamilton function (not a quantum Hamilton operator) which simply specifies the energy of any configuration of spins. The Ising Hamilton function is a sum of nearest neighbor contributions

$$\mathcal{H}[s] = J \sum_{\langle xy \rangle} s_x s_y - \mu B \sum_x s_x, \tag{2.4.1}$$

with a ferromagnetic coupling constant $J < 0$ that favors parallel spins, plus a coupling to an external magnetic field B. The classical partition function of this system is given by

$$Z = \int \mathcal{D}s \ \exp(-\mathcal{H}[s]/T) = \prod_x \sum_{s_x = \pm 1} \exp(-\mathcal{H}[s]/T). \tag{2.4.2}$$

The sum over all spin configurations corresponds to an independent summation over all possible orientations of individual spins. Thermal averages are computed by inserting appropriate operators. For example, the magnetization is given by

$$\langle s_x \rangle = \frac{1}{Z} \prod_x \sum_{s_x = \pm 1} s_x \exp(-\mathcal{H}[s]/T). \tag{2.4.3}$$

Similarly, the spin correlation function is defined by

$$\langle s_x s_y \rangle = \frac{1}{Z} \prod_x \sum_{s_x = \pm 1} s_x s_y \exp(-\mathcal{H}[s]/T). \tag{2.4.4}$$

At large distances the connected spin correlation function typically decays exponentially

$$\langle s_x s_y \rangle - \langle s \rangle^2 \sim \exp(-|x-y|/\xi), \tag{2.4.5}$$

where ξ is the so-called correlation length. At general temperatures the correlation length is typically just a few lattice spacings. When one models real materials, the Ising model would generally be a great oversimplification, because real magnets, for example, not only have nearest neighbor couplings. Still, the details of the Hamilton function at the scale of the lattice spacing are not always important. There is a critical temperature T_c at which ξ diverges and universal behavior arises. At this temperature a second order phase transition occurs. Then the details of the model at the scale of the lattice spacing are irrelevant for the long range physics that takes place at the scale of ξ. In fact, at their critical temperatures some real materials behave just like the simple Ising model. This is why the Ising model is so interesting. It is just a very simple member of a large universality class of different models, which all share the same critical behavior. This does not mean that they have the same values of their critical temperatures. However, their magnetization goes to zero at the critical temperature with the same power of $T_c - T$, i.e. their critical exponents are identical.

2.5 Quantum Mechanics versus Statistical Mechanics

We notice a close analogy between the Euclidean path integral for a quantum mechanical system and a classical statistical mechanics system like the Ising model. The path integral for the quantum system is defined on a 1-dimensional Euclidean time lattice, just like an Ising model can be defined on a d-dimensional spatial lattice. In the path integral we integrate over all paths, i.e. over all configurations $x(t)$, while in the Ising model we sum over all spin configurations s_x. Paths are weighted by their Euclidean action $S_E[x]$ while spin configurations are weighted with their Boltzmann factors depending on the classical Hamilton function $\mathcal{H}[s]$. The prefactor of the action is $1/\hbar$, and the prefactor of the Hamilton function is $1/T$. Indeed $\hbar$ determines the strength of quantum fluctuations, while the temperature T determines the strength of thermal fluctuations. The kinetic energy $\frac{1}{2}((x_{i+1} - x_i)/a)^2$ in the path integral is analogous to the nearest neighbor spin coupling $s_x s_{x+1}$, and the potential term $V(x_i)$ is analogous to the coupling $\mu B s_x$ to an external magnetic field. The magnetization $\langle s_x \rangle$ corresponds to the vacuum expectation value of an operator $\langle \mathcal{O}(x) \rangle$ and the spin-spin correlation function $\langle s_x s_y \rangle$ corresponds to the 2-point correlation function $\langle \mathcal{O}(x(0)) \mathcal{O}(x(t)) \rangle$. The inverse correlation length $1/\xi$ is analogous to the energy gap $E_1 - E_0$ (and hence to a particle mass in a Euclidean

Quantum mechanics	Classical statistical mechanics
Euclidean time lattice	d-dimensional spatial lattice
elementary time step a	crystal lattice spacing
particle position x	classical spin variable s
particle path $x(t)$	spin configuration s_x
path integral $\int \mathcal{D}x$	sum over configurations $\prod_x \sum_{s_x}$
Euclidean action $S_E[x]$	classical Hamilton function $\mathcal{H}[s]$
Planck's constant $\hbar$	temperature T
quantum fluctuations	thermal fluctuations
kinetic energy $\frac{1}{2}(\frac{x_{i+1}-x_i}{a})^2$	neighbor coupling $s_x s_{x+1}$
potential energy $V(x_i)$	external field energy $\mu B s_x$
weight of a path $\exp(-\frac{1}{\hbar}S_E[x])$	Boltzmann factor $\exp(-\mathcal{H}[s]/T)$
vacuum expectation value $\langle\mathcal{O}(x)\rangle$	magnetization $\langle s_x\rangle$
2-point function $\langle\mathcal{O}(x(0))\mathcal{O}(x(t))\rangle$	correlation function $\langle s_x s_y\rangle$
energy gap $E_1 - E_0$	inverse correlation length $1/\xi$
continuum limit $a \to 0$	critical behavior $\xi \to \infty$

Table 2.2: *The dictionary that translates quantum mechanics into the language of classical statistical mechanics.*

quantum field theory). Finally, the Euclidean time continuum limit $a \to 0$ corresponds to a second order phase transition where $\xi \to \infty$. The lattice spacing in the path integral is an artifact of our mathematical description which we send to zero while the physics remains constant. In classical statistical mechanics, on the other hand, the lattice spacing is physical and hence fixed, while the correlation length ξ goes to infinity at a second order phase transition. All this is summarized in the dictionary of table 2.2.

2.6 The Transfer Matrix

The analogy between quantum mechanics and classical statistical mechanics suggests that there is an analog of the quantum Hamilton operator in the context of classical statistical mechanics. This operator is the so-called transfer matrix. The Hamilton operator induces infinitesimal translations in time. In classical statistical mechanics, on the other hand, the analog of continuous time is a 1-dimensional spatial lattice. Hence, the transfer matrix cannot induce infinitesimal space translations. Instead it induces translations by the smallest possible distance — namely by one lattice spacing. For a quantum mechanical system the transfer matrix transports us by one

lattice spacing in Euclidean time, and it is given by

$$T = \exp(-\frac{a}{\hbar}H). \tag{2.6.1}$$

Now we want to construct the transfer matrix for the 1-dimensional Ising model without an external magnetic field. The corresponding partition function is given by

$$Z = \prod_x \sum_{s_x=\pm 1} \exp(\beta J \sum_x s_x s_{x+1}). \tag{2.6.2}$$

The transfer matrix obeys

$$Z = \mathrm{Tr} T^N, \tag{2.6.3}$$

where N is the number of lattice points, and its matrix elements are given by the Boltzmann factor corresponding to a nearest neighbor pair by

$$\langle s_{x+1}|T|s_x\rangle = \exp(\beta J s_x s_{x+1}). \tag{2.6.4}$$

This is a 2×2 matrix. The eigenvalues of the transfer matrix can be written as $\exp(-E_0)$ and $\exp(-E_1)$. The energy gap then determines the inverse correlation length as

$$1/\xi = E_1 - E_0. \tag{2.6.5}$$

It is instructive to compute ξ as a function of β to locate the critical point of the 1-d Ising model.

Here we will do the corresponding calculation for the 1-d xy-model. In the xy-model the spins are unit vectors $(\cos\varphi_x, \sin\varphi_x)$ in the xy-plane that are attached to the points x of a d-dimensional lattice. Here we consider $d = 1$, i.e. we study a chain of xy-spins. The standard Hamilton function of the xy-model is given by

$$\mathcal{H}[\varphi] = J \sum_{\langle xy\rangle} (1 - \cos(\varphi_{x+1} - \varphi_x)). \tag{2.6.6}$$

In complete analogy to the Ising model the transfer matrix is now given by

$$\langle \varphi_{x+1}|T|\varphi_x\rangle = \exp(-\beta J(1 - \cos(\varphi_{x+1} - \varphi_x)), \tag{2.6.7}$$

which is a matrix with an uncountable number of rows and columns, because there is a continuum of values for φ_x and φ_{x+1}. Still, we can ask about the eigenvalues of this matrix. For this purpose we consider the Fourier representation

$$\langle \varphi_{x+1}|T|\varphi_x\rangle = \sum_{m\in\mathbb{Z}} \langle \varphi_{x+1}|m\rangle \exp(-\beta J) I_m(\beta J) \langle m|\varphi_x\rangle, \tag{2.6.8}$$

where

$$\langle \varphi_x|m\rangle = \exp(im\varphi_x), \tag{2.6.9}$$

are the eigenvectors of the transfer matrix. The eigenvalues are given in terms of modified Bessel functions

$$\exp(-E_m) = \exp(-\beta J) I_m(\beta J). \tag{2.6.10}$$

The energy gap between the ground state and an excited state is given by

$$E_m - E_0 = \log \frac{I_0(\beta J)}{I_m(\beta J)}, \tag{2.6.11}$$

which is nonzero for finite β. In the zero temperature limit $\beta \to \infty$ we have

$$\frac{I_0(\beta J)}{I_m(\beta J)} \sim 1 + \frac{m^2}{2\beta J}, \tag{2.6.12}$$

such that

$$\xi = 1/(E_1 - E_0) \sim 2\beta J \to \infty. \tag{2.6.13}$$

Hence, there is a critical point at zero temperature. In the language of quantum mechanics this implies the continuum limit of a Euclidean lattice theory corresponding to a quantum mechanical problem. In the continuum limit the energies corresponding to the eigenvalues of the transfer matrix take the form

$$E_m - E_0 \sim \frac{m^2}{2\beta J}. \tag{2.6.14}$$

These energies are in lattice units (the lattice spacing was put to 1). Hence, to extract physics we need to consider energy ratios and we find

$$\frac{E_m - E_0}{E_1 - E_0} \sim m^2. \tag{2.6.15}$$

These are the appropriate energy ratios of a quantum rotor — a particle that moves on a circle. Indeed the xy-spins describe an angle, which can be interpreted as the position of the quantum particle. Also the eigenvectors of the transfer matrix are just the energy eigenfunctions of a quantum rotor. Hence, we just solved the Schrödinger equation with a discrete Euclidean time step using the transfer matrix instead of the Hamilton operator. The fact that energy ratios approach physically meaningful constants in the continuum limit is known as scaling. Of course, the discretization introduces an error as long as we are not in the continuum limit. For example, at finite β the energy ratio is

$$\frac{E_m}{E_1} = \frac{\log(I_0(\beta J)/I_m(\beta J)}{\log(I_0(\beta J)/I_1(\beta J)}, \tag{2.6.16}$$

which is different from the continuum answer m^2. This cut-off effect due to a finite lattice spacing is known as a scaling violation.

2.7 Lattice Field Theory

So far we have restricted ourselves to quantum mechanical problems and to classical statistical mechanics. The former were defined by a path integral on a 1-d Euclidean time lattice, while the latter involved spin models on a d-dimensional spatial lattice. When we quantize field theories on the lattice, we formulate the theory on a d-dimensional space-time lattice, i.e. usually the lattice is 4-dimensional. Just as we integrate over all configurations (all paths) $x(t)$ of a quantum particle, we now integrate over all configurations $\phi(x)$ of a quantum field defined at any Euclidean space-time point $x = (\vec{x}, x_4)$. Again the weight factor in the path integral is given by the action. Let us illustrate this for a free neutral scalar field $\phi(x) \in R$. Its Euclidean action is given by

$$S_E[\phi] = \int d^4x \, [\frac{1}{2}\partial_\mu\phi\partial_\mu\phi + \frac{m^2}{2}\phi^2]. \tag{2.7.1}$$

Interactions can be included, for example, by adding a $\frac{\lambda}{4!}\phi^4$ term to the action. The Feynman path integral for this system is formally written as

$$Z = \int \mathcal{D}\phi \exp(-S_E[\phi]). \tag{2.7.2}$$

(Note that we have put $\hbar = c = 1$.) The integral is over all field configurations, which is a divergent expression if no regularization is imposed. One can make the expression mathematically well-defined by using dimensional regularization of Feynman diagrams. This approach is, however, limited to perturbation theory. The lattice allows us to formulate field theory beyond perturbation theory, which is very essential for strongly interacting theories like QCD, but also for the standard model in general. For example, due to the heavy mass of the top quark, the Yukawa coupling between the Higgs and top quark field is rather strong. The above free scalar field theory, of course, does not really require a nonperturbative treatment. We use it only to illustrate the lattice quantization method in a simple setting. On the lattice the continuum field $\phi(x)$ is replaced by a lattice field Φ_x, which is restricted to the points x of a d-dimensional space-time lattice. From now on we will work in lattice units, i.e. we put $a = 1$. The above continuum action can be approximated by discretizing the continuum derivatives such that

$$S_E[\Phi] = \sum_{x,\mu} \frac{1}{2}(\Phi_{x+\hat{\mu}} - \Phi_x)^2 + \sum_x \frac{m^2}{2}\Phi_x^2. \tag{2.7.3}$$

Here $\hat{\mu}$ is the unit vector in the μ-direction. The integral over all field configurations now becomes a multiple integral over all values of the field at all lattice points

$$Z = \prod_x \int_{-\infty}^{\infty} d\Phi_x \exp(-S_E[\Phi]). \tag{2.7.4}$$

For a free field theory the partition function is just a Gaussian integral. In fact, one can write the lattice action as

$$S_E[\Phi] = \frac{1}{2}\sum_{x,y} \Phi_x \mathcal{M}_{xy} \Phi_y, \tag{2.7.5}$$

where the matrix $\mathcal{M}$ describes the couplings between lattice points. Diagonalizing this matrix by a unitary transformation $\mathcal{U}$ one has

$$\mathcal{M} = \mathcal{U}^\dagger \mathcal{D} \mathcal{U}. \tag{2.7.6}$$

Introducing

$$\Phi'_x = \mathcal{U}_{xy}\Phi_y \tag{2.7.7}$$

one obtains

$$Z = \prod_x \int d\Phi'_x \exp(-\frac{1}{2}\sum_x \Phi'_x \mathcal{D}_{xx} \Phi'_x) = (2\pi)^{N/2} \mathrm{det}\mathcal{D}^{-1/2}, \tag{2.7.8}$$

where N is the number of lattice points.

To extract the energy values of the corresponding quantum Hamilton operator we need to study the 2-point function of the lattice field

$$\langle \Phi_x \Phi_y \rangle = \frac{1}{Z}\int \mathcal{D}\Phi \ \Phi_x \Phi_y \exp(-S_E[\Phi]). \tag{2.7.9}$$

This is most conveniently done by introducing a source field in the partition function, such that

$$Z[J] = \int \mathcal{D}\Phi \exp(-S_E[\Phi] + \sum_x J_x \Phi_x). \tag{2.7.10}$$

Then the connected 2-point function is given by

$$\langle \Phi_x \Phi_y \rangle - \langle \Phi \rangle^2 = \frac{\partial^2 \log Z[J]}{\partial J_x \partial J_y}|_{J=0}. \tag{2.7.11}$$

The Boltzmann factor characterizing the problem with the external sources is given by the exponent

$$\frac{1}{2}\Phi \mathcal{M} \Phi - J\Phi = \frac{1}{2}\Phi' \mathcal{M} \Phi' - \frac{1}{2} J \mathcal{M}^{-1} J. \tag{2.7.12}$$

Here we have introduced

$$\Phi' = \Phi - \mathcal{M}^{-1} J. \tag{2.7.13}$$

Integrating over Φ' in the path integral we obtain

$$Z[J] = (2\pi)^{N/2} \mathrm{det}\mathcal{D}^{-1/2} \exp(\frac{1}{2} J \mathcal{M}^{-1} J), \tag{2.7.14}$$

and hence

$$\langle \Phi_x \Phi_y \rangle = \frac{1}{2} \mathcal{M}^{-1}_{xy}. \tag{2.7.15}$$

It is instructive to invert the matrix $\mathcal{M}$ by going to Fourier space, i.e. by writing

$$\Phi_x = \frac{1}{(2\pi)^d} \int_B d^d p \; \Phi(p) \exp(ipx). \tag{2.7.16}$$

The momentum space of the lattice is given by the Brillouin zone $B =]-\pi, \pi]^d$. For the 2-point function in momentum space one then finds

$$\langle \Phi(-p)\Phi(p) \rangle = [\sum_\mu (2\sin(p_\mu/2))^2 + m^2]^{-1}. \tag{2.7.17}$$

This is the lattice version of the continuum propagator

$$\langle \Phi(-p)\Phi((p) \rangle = (p^2 + m^2)^{-1}. \tag{2.7.18}$$

From the lattice propagator we can deduce the energy spectrum of the lattice theory. For this purpose we construct a lattice field with definite spatial momentum $\vec{p}$ located in a specific time slice

$$\Phi(\vec{p})_t = \sum_x \Phi_{\vec{x},t} \exp(-i\vec{p} \cdot \vec{x}), \tag{2.7.19}$$

and we consider its 2-point function

$$\langle \Phi(-\vec{p})_0 \Phi(\vec{p})_t \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dp_d \langle \Phi(-p)\Phi(p) \rangle \exp(ip_d t). \tag{2.7.20}$$

Inserting the lattice propagator of eq.(2.7.17) one can perform the integral. One encounters a pole in the propagator when $p_d = iE$ with

$$(2\sinh(E/2))^2 = \sum_i (2\sin(p_i/2))^2 + m^2. \tag{2.7.21}$$

The 2-point function then takes the form

$$\langle \Phi(-\vec{p})_0 \Phi(\vec{p})_t \rangle = C \exp(-Et), \tag{2.7.22}$$

i.e. it decays exponentially with slope E. This allows us to identify E as the energy of the lattice scalar particle with spatial momentum $\vec{p}$. In general, E differs from the correct continuum dispersion relation

$$E^2 = \vec{p}^2 + m^2. \tag{2.7.23}$$

Only in the continuum limit, i.e. when E, $\vec{p}$ and m are small in lattice units, the lattice dispersion relation agrees with the one of the continuum theory.

Chapter 3

Lie Groups and Lie Algebras

In 1873 the Norwegian mathematician Sophus Lie discovered the general concept of non-Abelian continuous symmetry groups and their associated algebras, known as Lie algebras and Lie groups. The simplest continuous non-Abelian (i.e. non-commuting) symmetry is $SO(3)$ — the rotation group in a 3-dimensional Euclidean space, consisting of real orthogonal 3×3 matrices with determinant 1. The associated algebra is characterized by the angular momentum commutation relations among the three generators of the $so(3)$ algebra. This algebra is equivalent to the algebra $su(2)$ which generates the group $SU(2)$ of complex unitary 2×2 matrices with determinant 1. The group manifold of $SU(2)$ is the 3-dimensional sphere S^3, while the group manifold of $SO(3)$ is the coset space $S^3/\mathbb{Z}(2)$. Here $\mathbb{Z}(2) = \{1, -1\}$ is the so-called center of the group $SU(2)$, such that $SO(3) = SU(2)/\mathbb{Z}(2)$. $SU(2)$ is the so-called universal covering group of $SO(3)$. The so-called simple compact Lie algebras have been completely classified by Elie Cartan in 1894. There are the orthogonal algebras $so(n)$ (with $n \geq 3$), the unitary algebras $su(n)$ (with $n \geq 2$), the symplectic algebras $sp(n)$ (with $n \geq 1$), as well as the exceptional algebras $g(2)$, $f(4)$, $e(6)$, $e(7)$, and $e(8)$, which were discovered by Wilhelm Killing in 1890. Correspondingly, there are the Lie groups $SO(n)$, $SU(n)$, $Sp(n)$, $G(2)$, $F(4)$, $E(6)$, $E(7)$, and $E(8)$. The $SO(n)$ groups have a universal covering group known as $Spin(n)$.

3.1 Definition of a Lie Algebra

A Lie algebra is spanned by its generators T^a with $a \in \{1, 2, \ldots, n_G\}$, where n_G denotes the number of generators. A general element of the algebra is given by a

linear combination

$$H = \sum_{a=1}^{n_G} \omega^a T^a = \omega^a T^a, \tag{3.1.1}$$

with real-valued parameters $\omega^a \in \mathbb{R}$. Here we have used Einstein's convention of summing over repeated indices. In this case, we do not distinguish between upper and lower (or co- and contra-variant) indices. Different elements of the algebra can be added together, i.e.

$$H_1 + H_2 = H, \quad H_1 = \omega_1^a T^a, \quad H_2 = \omega_2^a T^a, \quad \omega_1^a + \omega_2^a = \omega^a, \tag{3.1.2}$$

such that H again belongs to the algebra. The product $H_1 H_2$, on the other hand, does in general not belong to the algebra. Instead, the concept of "multiplication" in a Lie algebra is represented by the commutator

$$[H_1, H_2] = H_1 H_2 - H_2 H_1 = \omega_1^a \omega_2^b [T^a, T^b]. \tag{3.1.3}$$

The structure of a Lie algebra is characterized by its structure constants $f_{abc} \in \mathbb{R}$ which determine the commutation relations between the generators

$$[T^a, T^b] = i f_{abc} T^c. \tag{3.1.4}$$

In particular, $i[T^a, T^b] = -f_{abc} T^c$ and thus $i[H_1, H_2]$ is again a member of the algebra. Due to the anti-symmetry of the commutator the structure constants satisfy $f_{bac} = -f_{abc}$. The Jacobi identity

$$[[T^a, T^b], T^c] + [[T^b, T^c], T^a] + [[T^c, T^a], T^b] = 0, \tag{3.1.5}$$

implies that the f_{abc} are even totally anti-symmetric against all permutations of the indices. The generators of a Lie algebra are Hermitean, i.e.

$$T^{a\dagger} = T^a, \tag{3.1.6}$$

such that the commutator

$$[T^a, T^b]^\dagger = (T^a T^b - T^b T^a)^\dagger = T^{b\dagger} T^{a\dagger} - T^{a\dagger} T^{b\dagger} = T^b T^a - T^a T^b = -[T^a, T^b] \tag{3.1.7}$$

is anti-Hermitean, and $i[T^a, T^b]$ is indeed again Hermitean. It is a non-trivial feature of Lie algebras that the commutator of two generators is again a linear combination of generators. This structure can be realized only with very specific sets of generators.

3.2 Simple and Semi-Simple Lie Algebras

Important internal symmetries in particle physics are associated with so-called simple and semi-simple Lie algebras. By definition, a semi-simple Lie algebra does not have any invariant Abelian sub-algebras. First of all, a sub-algebra is generated by a subset S^a of generators (which are linear combinations of the T^a) that is closed under commutation, i.e.

$$[S^a, S^b] = i f_{abc} S^c. \tag{3.2.1}$$

An invariant sub-algebra obeys the additional requirement that

$$[T^a, S^b] = i f_{abc} S^c, \tag{3.2.2}$$

for all generators T^a of the algebra and all generators S^b of the sub-algebra. A Lie algebra is called simple if it has no invariant sub-algebra. A sub-algebra is Abelian if its generators commute with each other, i.e. $[S^a, S^b] = 0$. A Lie algebra is called semi-simple if it has no invariant Abelian sub-algebra. A semi-simple Lie algebra may still have non-Abelian invariant sub-algebras.

The number of generators of the maximal Abelian sub-algebra of a semi-simple Lie algebra (which cannot be an invariant sub-algebra) determines the rank r of the algebra. In the following, we will focus on compact Lie algebras whose associated Lie groups have a compact group manifold. It turns out that compact Lie algebras are semi-simple.

3.3 Representations of Lie Algebras

A Lie algebra has many possible representations, which can be viewed as concrete realizations of its generators, which we have, until now, defined only abstractly and implicitly through their commutation relations. As we will see, the generators can, for example, be represented by finite-dimensional matrices. In fact, every semi-simple Lie algebra has an n_G-dimensional representation, which is known as the adjoint representation. In the adjoint representation, the generators are realized as $n_G \times n_G$ matrices whose elements can be chosen as

$$T^a_{bc} = -i f_{abc}, \tag{3.3.1}$$

such that

$$[T^a, T^b]_{de} = T^a_{df} T^b_{fe} - T^b_{df} T^a_{fe} = -f_{adf} f_{bfe} + f_{bdf} f_{afe}. \tag{3.3.2}$$

The commutation relation $[T^a, T^b]_{de} = i f_{abc} T^c_{de} = f_{abc} f_{cde}$ thus implies the consistency condition

$$-f_{adf} f_{bfe} + f_{bdf} f_{afe} = f_{abc} f_{cde}, \tag{3.3.3}$$

which follows directly from the Jacobi identity (3.1.5).

The conjugate $\widetilde{T}$ of a representation is defined as

$$i\widetilde{T}^a = (iT^a)^* = -i(T^a)^T \;\Rightarrow\; \widetilde{T}^a = -(T^a)^T \tag{3.3.4}$$

where $*$ denotes complex conjugation and T denotes transpose. The conjugate is again a representation of the algebra because

$$\begin{aligned}[\widetilde{T}^a, \widetilde{T}^b] &= [(T^a)^T, (T^b)^T] = (T^a)^T(T^b)^T - (T^b)^T(T^a)^T = (T^bT^a - T^aT^b)^T \\ &= -[T^a, T^b]^T = -if_{abc}(T^c)^T = if_{abc}\widetilde{T}^c. \end{aligned} \tag{3.3.5}$$

A representation is called real, if $\widetilde{T}^a = T^a$. It is easy to show that the adjoint representation is real. A representation is called pseudo-real if $\widetilde{T}^a$ and T^a are unitarily equivalent, i.e. if there is a unitary transformation U such that $U\widetilde{T}^aU^\dagger = T^a$. Representations that are in-equivalent to their conjugate representation are known as complex representations. Among the compact semi-simple Lie algebras, only $su(n)$ with $n \geq 3$, $so(4n+2)$ with $n \geq 1$, and $e(6)$ have complex representations.

3.4 The Lie Algebra $so(3)$ and its Representations

The simplest simple Lie algebra has three generators (i.e. $n_G = 3$) which obey the commutation relations

$$[T^a, T^b] = i\varepsilon_{abc}T^c, \tag{3.4.1}$$

where ε_{abc} is the totally anti-symmetric Levi-Civita tensor. It is straightforward to show that the consistency condition eq.(3.3.3) is indeed satisfied. The above algebra is well-known from quantum mechanics where the generators T^a are the three components of an angular momentum vector. Angular momentum is conserved as a consequence of the isotropy of 3-dimensional Euclidean coordinate space. The rotations in this space are described by real orthogonal 3×3 matrices with determinant 1. Under multiplication, these matrices form the group $SO(3)$. The corresponding Lie algebra is known as $so(3)$. One can show that $so(3)$ is simple, i.e. it has no invariant sub-algebra.

The quantum mechanical orbital angular momentum vector of a single particle

$$\vec{L} = \vec{r} \times \vec{p} = -i\hbar\, \vec{r} \times \vec{\nabla} \tag{3.4.2}$$

provides a concrete realization of the commutation relations of the $so(3)$ algebra when we identify $T^a = L^a$ with $a \in \{1, 2, 3\}$. As we know, the quantum mechanical angular momentum operator acts in the Hilbert space of square integrable wave

functions, which is infinite-dimensional. Hence, the representation of the $so(3)$ generators by the operator $\vec{L}$ is infinite-dimensional.

As we know from quantum mechanics, general wave functions can be decomposed into linear combinations of angular momentum eigenstates $|lm\rangle$ which obey

$$\vec{L}^{\,2}|lm\rangle = \hbar^2 l(l+1)|lm\rangle, \quad L^3|lm\rangle = \hbar m|lm\rangle, \quad l \in \mathbb{N}, \quad m \in \{-l, -l+1, \ldots, l\}. \tag{3.4.3}$$

The operator $\vec{L}^{\,2} = L^a L^a$ commutes with all components of the angular momentum vector, i.e. $[\vec{L}^{\,2}, L^a] = 0$. This is characteristic of a so-called Casimir operator of a Lie algebra, which consists of sums of products of generators and commutes with all generators T^a. The Casimir operators themselves do not belong to the Lie algebra. According to a theorem due to Racah, the number of independent Casimir operators of a semi-simple Lie algebra is given by the rank. Since the different components of an angular momentum vector do not commute with each other, $so(3)$ has rank 1. Consequently, $\vec{L}^{\,2}$ is the only independent Casimir operator. Obviously, any polynomial of $\vec{L}^{\,2}$ also commutes with all the generators. However, those operators are not independent of $\vec{L}^{\,2}$.

If we consider a particle moving on the surface of a 2-dimensional sphere S^2, in spherical coordinates its position can be described by the angles θ and φ. Correspondingly, its wave function can be expressed as

$$\Psi(\theta, \varphi) = \langle \theta, \varphi | \Psi \rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} \langle \theta, \varphi | lm \rangle, \tag{3.4.4}$$

where $c_{lm} \in \mathbb{C}$ and $\langle \theta, \varphi | lm \rangle = Y_{lm}(\theta, \varphi)$ are the spherical harmonics. The set of all states $|lm\rangle$ span the infinite-dimensional Hilbert space of wave functions on the sphere S^2. However, already the states $|lm\rangle$ with fixed $l \in \mathbb{N}$ and with $m \in \{-l, -l+1, \ldots, l\}$ alone span a finite $(2l+1)$-dimensional representation of the $so(3)$ algebra. The infinite-dimensional representation through the angular momentum operators $\vec{L}$ is reducible and decomposes into a sum of irreducible representations characterized by the quantum number l. A representation is called irreducible if the algebra cannot be realized on any subset of states that are involved. It turns out that the eigenvalues of the independent Casimir operators uniquely characterize the irreducible representations of a semi-simple Lie algebra.

The trivial irreducible representation of the $so(3)$ algebra is 1-dimensional and corresponds to $l = 0$ (with $m = 0$) consisting of the single state $|00\rangle$. In this trivial subspace of the Hilbert space, the generators are simply given by

$$T^1 = T^2 = T^3 = 0, \tag{3.4.5}$$

which obviously satisfy the commutation relations. A non-trivial 3-dimensional representation corresponds to $l = 1$ and is spanned by the three states $|11\rangle$, $|10\rangle$,

and $|1-1\rangle$. In this case, the $so(3)$ generators are represented by concrete 3×3 matrices

$$T^1=\frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix},\quad T^2=\frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix},\quad T^3=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{3.4.6}$$

It is straightforward to convince oneself that these matrices indeed obey the commutation relations of $so(3)$. The dimension of this representation is equal to the number of generators $n_G=3$. Indeed, this representation is equivalent to the adjoint representation $T^a_{bc}=-i\varepsilon_{abc}$. Although the representation matrices are not identical, they span the same algebra, because they are related to each other by a unitary transformation.

As we know from quantum mechanics, not only orbital angular momentum but also spin obeys the angular momentum commutation relations. In particular, in contrast to orbital angular momentum, spin can also be quantized in half-integer units. For example, the Pauli matrices $\vec{\sigma}$ give rise to the representation

$$T^1=\frac{1}{2}\sigma^1=\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix},\ T^2=\frac{1}{2}\sigma^2=\frac{1}{2}\begin{pmatrix}0&-i\\i&0\end{pmatrix},\ T^3=\frac{1}{2}\sigma^3=\frac{1}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{3.4.7}$$

This representation is usually associated with the algebra $su(2)$, which, however, is identical with the one of $so(3)$ (and also with the one of $sp(1)$). As we will see, while the algebras $so(3)$ and $su(2)$ are identical, the corresponding groups $SO(3)$ and $SU(2)$ are still different. In the following, we will associate general representations with integer or half-integer angular momentum with the algebra $su(2)$. On the other hand, when we want to explicitly restrict ourselves to integer representations only, we will associate them with the algebra $so(3)$. In general, an irreducible $(2j+1)$-dimensional representation of the $su(2)$ algebra is characterized by its angular momentum $j\in\{0,\frac{1}{2},1,\frac{3}{2},\ldots\}$. The spin $j=\frac{1}{2}$ representation is the smallest non-trivial representation, also known as the fundamental representation. It is easy to show that this representation is pseudo-real. In fact, all $su(2)$ representations are real or pseudo-real. Consequently, $su(2)$ has no complex representation. For illustrative purposes, we also write down the 4×4 matrices of the spin $j=\frac{3}{2}$ representation

$$T^1=\frac{1}{2}\begin{pmatrix}0&\sqrt{3}&0&0\\\sqrt{3}&0&2&0\\0&2&0&\sqrt{3}\\0&0&\sqrt{3}&0\end{pmatrix},\ T^2=\frac{1}{2}\begin{pmatrix}0&-i\sqrt{3}&0&0\\i\sqrt{3}&0&-2i&0\\0&2i&0&-i\sqrt{3}\\0&0&i\sqrt{3}&0\end{pmatrix},$$
$$T^3=\frac{1}{2}\begin{pmatrix}3&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-3\end{pmatrix}. \tag{3.4.8}$$

It is easy to convince oneself that $T^aT^a = \frac{3}{2}(\frac{3}{2}+1)\mathbf{1}$.

When we couple the orbital angular momentum $\vec{L}$ and the spin $\vec{S}$ of a particle to its total angular momentum $\vec{J} = \vec{L} + \vec{S}$, we are working with the direct Kronecker product of two Lie algebras $so(3) \times su(2)$. While $so(3)$ and $su(2)$ (which, in fact, are the same algebra) are both simple, their product is semi-simple. Indeed, typical examples of semi-simple Lie algebras are direct products of simple ones.

3.5 The Groups $SU(2)$ and $SO(3)$

Let us now consider the $SU(2)$ and $SO(3)$ groups associated with the $su(2) = so(3)$ algebra. The group $U(2)$ consists of all unitary 2×2 matrices U, i.e.

$$UU^\dagger = U^\dagger U = \mathbf{1}. \tag{3.5.1}$$

The group $SU(2)$ is a subgroup of $U(2)$ consisting of those unitary 2×2 matrices that have determinant 1. It is straightforward to show that these matrices indeed form a group under matrix multiplication. The elements of $SU(2)$ can be associated with elements of the $su(2)$ algebra by exponentiation in the $j = \frac{1}{2}$ representation, i.e.

$$U = \exp(i\omega^a T^a) = \exp\left(\frac{i}{2}\vec{\omega}\cdot\vec{\sigma}\right) = \cos\left(\frac{\omega}{2}\right)\mathbf{1} + \frac{i}{\omega}\sin\left(\frac{\omega}{2}\right)\vec{\omega}\cdot\vec{\sigma}. \tag{3.5.2}$$

Here $\omega = |\vec{\omega}|$ and $\omega^a \in \mathbb{R}$. A general $SU(2)$ element can also be expressed as

$$U = \begin{pmatrix} A & B \\ -B^* & A^* \end{pmatrix}, \tag{3.5.3}$$

with $A = A_r + iA_i, B = B_r + iB_i \in \mathbb{C}$ and $|A|^2 + |B|^2 = 1$. The two complex numbers A and B with this length constraint correspond to four real numbers (the real and imaginary parts of A and B) with the constraint $A_r^2 + A_i^2 + B_r^2 + B_i^2 = 1$. This implies that the group elements of $SU(2)$ can be viewed as points on a 3-dimensional sphere S^3 (embedded in a 4-dimensional Euclidean space). Thus the $SU(2)$ group manifold is the sphere S^3.

Every group has a subgroup known as the center which consists of those group elements that commute with all other group elements. The center of the group $SU(2)$ consists of the unit-matrix $\mathbf{1}$ and the matrix $-\mathbf{1}$ (which is unitary and has determinant 1) and obviously commutes with all group elements. These two elements of the center form the Abelian group $\mathbb{Z}(2)$.

Similarly, the group $SO(3)$ consists of the real orthogonal 3×3 matrices with determinant 1, i.e.

$$OO^T = O^T O = \mathbf{1}, \quad \det O = 1. \tag{3.5.4}$$

Again, it is easy to convince oneself that these matrices form a group under matrix multiplication. I this case, the group elements can be obtained by exponentiating the algebra elements in the 3-dimensional adjoint representation $T^a_{bc} = -i\varepsilon_{abc}$ such that

$$O = \exp(i\omega^a T^a) = \exp \begin{pmatrix} 0 & \omega^3 & -\omega^2 \\ -\omega^3 & 0 & \omega^1 \\ \omega^2 & -\omega^1 & 0 \end{pmatrix}. \tag{3.5.5}$$

In contrast to $SU(2)$, the center of $SO(3)$ is trivial and consists only of the 3×3 unit-matrix $\mathbf{1}$. In particular, the matrix $-\mathbf{1}$ no longer has determinant 1 and thus does not belong to the group. Indeed the non-trivial center of $SU(2)$ is related to the fact that there are two types of $su(2)$ representations: those with integer and those with half-integer spin, while the $so(3)$ algebra only contains the inter spin representations.

Since the commutation relations of the $su(2)$ and $so(3)$ algebra are the same, it is not surprising that the $SU(2)$ and $SO(3)$ groups are also closely related. Indeed, it is instructive to convince oneself that the adjoint representation of $SO(3)$ is related to the fundamental representation of $SU(2)$ by

$$O_{ab} = \frac{1}{2}\mathrm{Tr}(U\sigma^a U^\dagger \sigma^b). \tag{3.5.6}$$

In this way, the unit-element $U = \mathbf{1}$ is mapped to

$$O_{ab} = \frac{1}{2}\mathrm{Tr}(\sigma^a \sigma^b) = \delta_{ab}, \tag{3.5.7}$$

which corresponds to the unit-element $O = \mathbf{1}$ of $SO(3)$. The inverse $U^\dagger$ of U is mapped to

$$\frac{1}{2}\mathrm{Tr}(U^\dagger \sigma^a U \sigma^b) = \frac{1}{2}\mathrm{Tr}(U\sigma^b U^\dagger \sigma^a) = O_{ba} = O^T_{ab}, \tag{3.5.8}$$

which corresponds to the inverse of O. It is instructive to convince oneself that the group structure $U_1 U_2 = U$ is maintained by the mapping such that $O_1 O_2 = O$.

Since there is a map from the group $SU(2)$ to $SO(3)$, one might expect that the two groups are identical. This is, however, not the case, because the mapping is not one to one. In fact, both U and $-U$ are mapped to the same group element O of $SO(3)$. In other words, $SO(3)$ is insensitive to the center $\mathbb{Z}(2)$ of $SU(2)$, which is consistent with the fact that $SU(2)$ has the center $\mathbb{Z}(2)$ while $SO(3)$ has a trivial center (consisting just of the 3×3 unit-matrix $\mathbf{1}$). Since U and $-U$ are

anti-podal points on the $SU(2)$ group manifold S^3, the $SO(3)$ manifold is the coset space $S^3/\mathbb{Z}(2)$ in which anti-podal points are identified with each other. $SU(2)$ is known as the universal covering group of $SO(3)$.

3.6 The Unitary Group $SU(n)$ and its Algebra

The unitary $n \times n$ matrices with determinant 1 form a group under matrix multiplication — the special unitary group $SU(n)$. This follows immediately from

$$UU^\dagger = U^\dagger U = 1,\ \mathrm{det}U = 1, \quad \mathrm{det}UV = \mathrm{det}U\mathrm{det}V = 1. \tag{3.6.1}$$

Associativity (i.e. $(UV)W = U(VW)$) holds for all matrices, a unit-element exists (the unit matrix $\mathbf{1}$), the inverse is $U^{-1} = U^\dagger$, and finally the group property

$$(UV)^\dagger UV = V^\dagger U^\dagger UV = 1,\ UV(UV)^\dagger = UVV^\dagger U^\dagger = 1 \tag{3.6.2}$$

also holds. The group $SU(n)$ is non-Abelian because in general $UV \neq VU$. Each element $U \in SU(n)$ can be represented as

$$U = \exp(iH), \tag{3.6.3}$$

where H is Hermitean and traceless. The matrices H form the $su(n)$ algebra. It has n^2-1 free parameters, and hence n^2-1 generators T^a, among which $n-1$ commute with each other. Thus the rank of $su(n)$ is $r = n-1$. Correspondingly, $su(n)$ has $n-1$ independent Casimir operators. The simplest non-trivial representation of $su(n)$ is the fundamental n-dimensional representation.

The group manifold of $SU(n)$ is a product of spheres

$$SU(n) = S^3 \times S^5 \times \ldots \times S^{2n-1}. \tag{3.6.4}$$

In order to show this, we decompose a general $SU(n)$ matrix U into an element W of the subgroup $SU(n-1)$ and an element V of the coset $SU(n)/SU(n-1)$

$$U = VW,\ W = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \tilde{U}_{11} & \tilde{U}_{12} & \ldots & \tilde{U}_{1\ n-1} \\ 0 & \tilde{U}_{21} & \tilde{U}_{22} & \ldots & \tilde{U}_{1\ n-1} \\ . & . & . & & . \\ . & . & . & & . \\ 0 & \tilde{U}_{n-1\ 1} & \tilde{U}_{n-1\ 2} & \ldots & \tilde{U}_{n-1\ n-1} \end{pmatrix}, \tag{3.6.5}$$

where the embedded matrix $\tilde{U}$ is in $SU(n-1)$. It is indirectly defined by

$$V = \begin{pmatrix} U_{11} & -U_{21}^* & -\frac{U_{31}^*(1+U_{11})}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*(1+U_{11})}{1+U_{11}^*} \\ U_{21} & \frac{1+U_{11}^*-|U_{21}|^2}{1+U_{11}} & -\frac{U_{31}^*U_{21}}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*U_{21}}{1+U_{11}^*} \\ U_{31} & -\frac{U_{21}^*U_{31}}{1+U_{11}} & \frac{1+U_{11}^*-|U_{31}|^2}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*U_{31}}{1+U_{11}^*} \\ . & . & . & & . \\ . & . & . & & . \\ U_{n1} & -\frac{U_{21}^*U_{n1}}{1+U_{11}} & -\frac{U_{31}^*U_{n1}}{1+U_{11}^*} & \cdots & \frac{1+U_{11}^*-|U_{n1}|^2}{1+U_{11}^*} \end{pmatrix} \in SU(n). \qquad (3.6.6)$$

One should convince oneself that V is indeed an $SU(n)$ matrix, and that the resulting matrix $\tilde{U}$ is indeed in $SU(n-1)$. The matrix V is constructed entirely from the elements $U_{11}, U_{21}, \ldots, U_{n1}$ of the first column of the matrix U, which is normalized to $|U_{11}|^2 + |U_{21}|^2 + \ldots + |U_{n1}|^2 = 1$. This implies that the matrix V takes values on the sphere S^{2n-1} and hence the coset manifold is

$$SU(n)/SU(n-1) = S^{2n-1}. \qquad (3.6.7)$$

By successively factoring out odd-dimensional spheres, one reduces $SU(n)$ all the way down to $SU(2)$, whose group manifold is S^3, which thus proves eq.(3.6.4).

The center of $SU(n)$ is the cyclic group $\mathbb{Z}(n) = \{\exp(2\pi im/n)\mathbf{1}, m = 0, 1, \ldots, n-1\}$ consisting of the unit-matrix $\mathbf{1}$ multiplied by a complex n-th root $\exp(2\pi im/n)$ of 1. These matrices obviously commute with all other group elements. In addition, they are unitary and have determinant 1, and thus indeed belong to $SU(n)$.

3.7 The Group $SU(3)$ and its Algebra

Let us consider the group $SU(3)$ and its algebra $su(3)$ in some detail. First of all, $su(3)$ has $3^2 - 1 = 8$ generators $T^a = \frac{1}{2}\lambda^a$, where λ^a are the Gell-Mann matrices

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$
$$\lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \qquad (3.7.1)$$

abc	123	147	156	246	257	345	367	458	678
f^{abc}	1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$

Table 3.1: *Non-zero values of the structure constants f_{abc} of the Lie algebra $su(3)$. The structure constants are completely anti-symmetric against permutation of the indices a, b, and c.*

abc	118	146	157	228	247	256	338	344
d^{abc}	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$
abc	355	366	377	448	558	668	778	888
d^{abc}	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$

Table 3.2: *Non-zero coefficients d_{abc} of the Lie algebra $su(3)$. These coefficients are symmetric against permutation of the indices a, b, and c.*

The corresponding structure constants of the $su(3)$ algebra are listed in table 2.1. The three generators λ^1, λ^2, and λ^3 form an $su(2)$ sub-algebra. Since λ^3 and λ^8 commute, and since there is no other independent linear combination of generators that commutes with these two, the rank of $su(3)$ is $r = 2$. According to Racah's theorem, this implies that $su(3)$ has two independent Casimir operators. One of them is quadratic in the generators and simply given by $C_1 = T^a T^a$. The other Casimir operator is $C_2 = d_{abc} T^a T^b T^c$, where the coefficients d_{abc} are listed in table 2.2. The structure constants f_{abc} as well as the coefficients d_{abc} can be expressed as

$$f_{abc} = \frac{1}{4i}\mathrm{Tr}([\lambda^a, \lambda^b]\lambda^c), \quad d_{abc} = \frac{1}{4}\mathrm{Tr}(\{\lambda^a, \lambda^b\}\lambda^c). \tag{3.7.2}$$

The anti-commutator of two Gell-Mann matrices can be expressed as

$$\{\lambda^a, \lambda^b\} = \frac{4}{3}\delta_{ab}\mathbf{1} + 2d_{abc}\lambda^c. \tag{3.7.3}$$

In analogy to the raising and lowering operators of spin in the $su(2)$ algebra, we introduce the following shift operators

$$T_\pm = T^1 \pm iT^2, \quad V_\pm = T^4 \pm iT^5, \quad U_\pm = T^6 \pm iT^7. \tag{3.7.4}$$

Besides these, there are the diagonal generators

$$T^3 = \frac{1}{2}\lambda^3, \quad Y = \frac{1}{\sqrt{3}}\lambda^8. \tag{3.7.5}$$

One then obtains the following commutation relations

$$
\begin{aligned}
&[T^3, T_\pm] = \pm T_\pm, \quad [T_+, T_-] = 2T^3, \\
&[T^3, V_\pm] = \pm\frac{1}{2}V_\pm, \quad [V_+, V_-] = \frac{3}{2}Y + T^3, \\
&[T^3, U_\pm] = \mp\frac{1}{2}U_\pm, \quad [U_+, U_-] = \frac{3}{2}Y - T^3, \\
&[Y, T^3] = [Y, T_\pm] = 0, \quad [Y, V_\pm] = \pm V_\pm, \quad [Y, U_\pm] = \pm U_\pm, \\
&[T_+, V_+] = [T_+, U_-] = [U_+, V_+] = 0, \\
&[T_+, V_-] = -U_-, \quad [T_+, U_+] = V_+, \quad [U_+, V_-] = T_-.
\end{aligned}
\tag{3.7.6}
$$

Since the generators T^3 and Y commute with each other, we can diagonalize them simultaneously and characterize the states of an $su(3)$ multiplet by the corresponding eigenvalues. The generators T^1, T^2, and T^3 generate an $su(2)$ sub-algebra of $su(3)$. Since these three generators all commute with Y, we can also simultaneously diagonalize the $su(2)$ Casimir operator $C = T^1T^1 + T^2T^2 + T^3T^3 = T(T+1)\mathbf{1}$. The states of an $su(3)$ multiplet can thus be further distinguished by the value of T. Hence, we characterize states of an irreducible $su(3)$ representation $\{\Gamma\}$ by Y, T, and T^3, and we denote them as $|\{\Gamma\}YTT^3\rangle$. For example, (using the flavor notation of up, down, and strange quarks) the states of the 3-dimensional fundamental representation $\{3\}$ are given by

$$
|u\rangle = |\{3\}\frac{1}{3}\frac{1}{2}\frac{1}{2}\rangle, \ |d\rangle = |\{3\}\frac{1}{3}\frac{1}{2}-\frac{1}{2}\rangle, \ |s\rangle = |\{3\}-\frac{2}{3}00\rangle. \tag{3.7.7}
$$

Similarly, the states of the conjugate representation $\{\overline{3}\}$ (representing anti-quarks) are given by

$$
|\overline{u}\rangle = |\{\overline{3}\}-\frac{1}{3}\frac{1}{2}-\frac{1}{2}\rangle, \ |\overline{d}\rangle = |\{\overline{3}\}-\frac{1}{3}\frac{1}{2}\frac{1}{2}\rangle, \ |\overline{s}\rangle = |\{\overline{3}\}\frac{2}{3}00\rangle. \tag{3.7.8}
$$

The raising and lowering operators $T_\pm$ shift the value of T^3 by ± 1, i.e.

$$
T_\pm|\{\Gamma\}YTT^3\rangle = \sqrt{T(T+1) - T^3(T^3 \pm 1)}|\{\Gamma\}YTT^3 \pm 1\rangle. \tag{3.7.9}
$$

The operators $V_\pm$ and $U_\pm$ also act as raising an lowering operators. In particular, we obtain

$$
\begin{aligned}
YV_\pm|\{\Gamma\}YTT^3\rangle &= ([Y, V_\pm] + V_\pm Y)|\{\Gamma\}YTT^3\rangle \\
&= (\pm V_\pm + V_\pm Y)|\{\Gamma\}YTT^3\rangle = (Y \pm 1)V_\pm|\{\Gamma\}YTT^3\rangle \\
T^3V_\pm|\{\Gamma\}YTT^3\rangle &= ([T^3, V_\pm] + V_\pm T^3)|\{\Gamma\}YTT^3\rangle \\
&= (\pm\frac{1}{2}V_\pm + V_\pm T^3)|\{\Gamma\}YTT^3\rangle = (T^3 \pm \frac{1}{2})V_\pm|\{\Gamma\}YTT^3\rangle
\end{aligned}
$$

$$\begin{aligned} YU_\pm|\{\Gamma\}YTT^3\rangle &= ([Y,U_\pm]+U_\pm Y)|\{\Gamma\}YTT^3\rangle \\ &= (\pm U_\pm + U_\pm Y)|\{\Gamma\}YTT^3\rangle = (Y\pm 1)U_\pm|\{\Gamma\}YTT^3\rangle \\ T^3U_\pm|\{\Gamma\}YTT^3\rangle &= ([T^3,U_\pm]+U_\pm T^3)|\{\Gamma\}YTT^3\rangle \\ &= (\mp\frac{1}{2}U_\pm + U_\pm T^3)|\{\Gamma\}YTT^3\rangle = (T^3\mp\frac{1}{2})U_\pm|\{\Gamma\}YTT^3\rangle. \end{aligned} \tag{3.7.10}$$

This implies that

$$\begin{aligned} V_\pm|\{\Gamma\}YTT^3\rangle &= \sum_{T'} C_{T'YTT^3}|\{\Gamma\}Y\pm 1,T',T^3\pm\frac{1}{2}\rangle, \\ U_\pm|\{\Gamma\}YTT^3\rangle &= \sum_{T'} C'_{T'YTT^3}|\{\Gamma\}Y\pm 1,T',T^3\mp\frac{1}{2}\rangle, \end{aligned} \tag{3.7.11}$$

where $C_{T'YTT^3}$ and $C'_{T'YTT^3}$ are coefficients that could be determined by a straightforward calculation. It should be noted that T' differs from T by a half-integer.

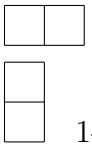
The weight diagrams of some $su(3)$ representations are illustrated in the figures. David Speiser, a physics professor at Leuven University and nephew of the mathematician Andreas Speiser who was a professor in Zürich, has invented a simple scheme that allows us to couple $su(3)$ representations. In his scheme, the different $su(3)$ representations are associated with the points of a hexagonal grid that is divided into six sectors associated with alternating + and - signs, and separated by empty lines that carry no representations. When the origin of a weight diagram of an $su(3)$ representation is placed on top of an $su(3)$ representation in a + sector of the grid, one can read off the reduction of the product of the two $su(3)$ representations into a sum of irreducible representations. This is done by listing all representations on the grid that are covered by the corresponding states in the weight diagram, taking into account both the multiplicity of the state and the sign of the sector of the grid. Points of the weight diagram that fall on top of the empty lines that separate the sectors do not contribute to the reduction.

This simple and elegant scheme generalizes to all compact semi-simple Lie algebras. For algebras with rank r the corresponding grid as well as the corresponding weight diagrams are r-dimensional. Hence, the scheme is most practical for the algebras of rank $r = 2$, which are $su(3)$, $so(4) = su(2)\oplus su(2)$, $so(5) = sp(2)$, and $g(2)$.

3.8 The Permutation Group S_N

Let us consider the permutation symmetry of N objects — for example the fundamental representations of $SU(n)$. Their permutations form the group S_N. The

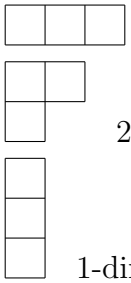
permutation group has $N!$ elements — all permutations of N objects. The group S_2 has two elements: the identity and the pair permutation. The representations of S_2 are represented by Young tableaux



1-dimensional symmetric representation,

1-dimensional antisymmetric representation. (3.8.1)

To describe the permutation properties of three objects we need the group S_3. It has $3! = 6$ elements: the identity, 3 pair permutations, and 2 cyclic permutations. The group S_3 has three irreducible representations

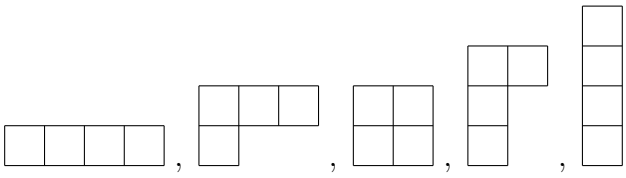


1-dimensional symmetric representation,

2-dimensional representation of mixed symmetry,

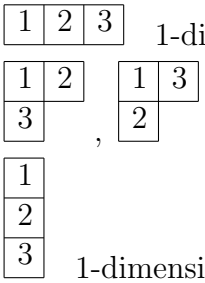
1-dimensional antisymmetric representation. (3.8.2)

The representations of the group S_N are given by the Young tableaux with N boxes. The boxes are arranged in left-bound rows, such that no row is longer than the one above it. For example, for the representations of S_4 one finds



(3.8.3)

The dimension of a representation is determined as follows. The boxes of the corresponding Young tableau are enumerated from 1 to N such that the numbers grow as one reads each row from left to right, and each column from top to bottom. The number of possible enumerations determines the dimension of the representation. For example, for S_3 one obtains



1-dimensional,

2-dimensional,

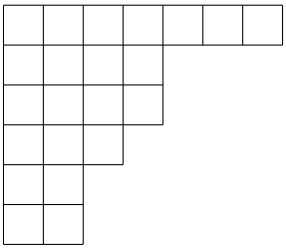
1-dimensional. (3.8.4)

The squares of the dimensions of all representations add up to the order of the group, i.e.

$$\sum_{\Gamma} d_{\Gamma}^2 = N! \ . \tag{3.8.5}$$

In particular, for S_2 we have $1^2 + 1^2 = 2 = 2!$ and for S_3 one obtains $1^2 + 2^2 + 1^2 = 6 = 3!$.

A general Young tableau can be characterized by the number of boxes m_i in its i-th row. For example the Young tableau

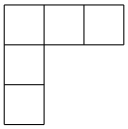


(3.8.6)

has $m_1 = 7$, $m_2 = 4$, $m_3 = 4$, $m_4 = 3$, $m_5 = 2$ and $m_6 = 2$. The dimension of the corresponding representation is given by

$$d_{m_1,m_2,\ldots,m_n} = N! \frac{\prod_{i<k}(l_i - l_k)}{l_1! l_2! \ldots l_n!}, \ l_i = m_i + n - i. \tag{3.8.7}$$

Applying this formula to the following Young tableau from S_5



(3.8.8)

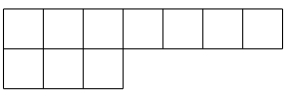
with $m_1 = 3$, $m_2 = 1$, $m_3 = 1$ and $n = 5$ yields $l_1 = 3+3-1 = 5$, $l_2 = 1+3-2 = 2$, $l_3 = 1 + 3 - 3 = 1$ and hence

$$d_{3,1,1} = 5! \frac{(l_1 - l_2)(l_1 - l_3)(l_2 - l_3)}{l_1! l_2! l_3!} = 5! \frac{3 \cdot 4 \cdot 1}{5!2!1!} = 6. \tag{3.8.9}$$

The permuted objects can be the fundamental representations of $SU(n)$. For $SU(2)$ we identify

 $= \{2\}.$ (3.8.10)

To each Young tableau with no more than two rows one can associate an $SU(2)$ representation. Such a Young tableau is characterized by m_1 and m_2, e.g.



(3.8.11)

has $m_1 = 7$ and $m_2 = 3$. The corresponding $SU(2)$ representation has

$$S = \frac{1}{2}(m_1 - m_2), \tag{3.8.12}$$

which is also denoted by $\{m_1 - m_2 + 1\}$. The above Young tableau hence represents $S = 2$ — a spin quintet $\{5\}$. Young tableaux with more than two rows have no realization in $SU(2)$ since among just two distinguishable objects no more than two can be combined anti-symmetrically.

3.9 Coupling of $su(n)$ Representations

The simplest nontrivial representation of $SU(n)$ is the fundamental representation. It is n-dimensional and can be identified with the Young tableau $\square$. Every irreducible representation of $SU(n)$ can be obtained from coupling N fundamental representations. In this way each $SU(n)$ representation is associated with a Young tableau with N boxes, which characterizes the permutation symmetry of the fundamental representations in the coupling. Since the fundamental representation is n-dimensional, there are n different fundamental properties (e.g. u and d in the $SU(2)$ isospin or u, d, and s in $SU(3)$ flavor symmetry of quarks). Hence, we can maximally anti-symmetrize n objects, and the Young tableaux for $SU(n)$ representations are therefore restricted to no more than n rows.

The dimension of an $SU(n)$ representation can be obtained from the corresponding Young tableau by filling it with factors as follows

n	$n+1$	$n+2$	$n+3$	$n+4$	$n+5$	$n+6$
$n-1$	n	$n+1$	$n+2$			
$n-2$	$n-1$	n	$n+1$			
$n-3$	$n-2$	$n-1$				
$n-4$	$n-3$					
$n-5$	$n-4$					

. (3.9.1)

The dimension of the $SU(n)$ representation is given as the product of all factors di-

vided by $N!$ and multiplied with the S_N dimension $d_{m1,m_2,\ldots,m_n}$ of the Young tableau

$$\begin{aligned} D^n_{m1,m_2,\ldots,m_n} &= \frac{(n+m_1-1)!}{(n-1)!}\frac{(n+m_2-2)!}{(n-2)!}\cdots\frac{m_n!}{0!}\frac{1}{N!}N!\frac{\prod_{i<k}(l_i-l_k)}{l_1!l_2!\ldots l_n!} \\ &= \frac{\prod_{i<k}(m_i-m_k-i+k)}{(n-1)!(n-2)!\ldots 0!}. \end{aligned} \tag{3.9.2}$$

We see that the dimension of a representation depends only on the differences $q_i = m_i - m_{i+1}$. In particular, for $SU(2)$ we find

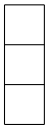
$$D^2_{m_1,m_2} = \frac{m_1-m_2-1+2}{1!0!} = m_1 - m_2 + 1 = q_1 + 1 \tag{3.9.3}$$

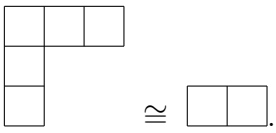
in agreement with our previous result. For a rectangular Young tableau with n rows, e.g. in $SU(2)$ for

, (3.9.4)

all $q_i = 0$, and we obtain

$$D^n_{m,m,\ldots,m} = \frac{\prod_{i<k}(m_i-m_k-i+k)}{(n-1)!(n-2)!\ldots 0!} = \frac{(n-1)!(n-2)!\ldots 0!}{(n-1)!(n-2)!\ldots 0!} = 1, \tag{3.9.5}$$

and therefore a singlet. This shows that in $SU(3)$  corresponds to a singlet. It also explains why the dimension of an $SU(n)$ representation depends only on the differences q_i. Without changing the dimension we can couple a representation with a singlet, and hence we can always add a rectangular Young tableau with n rows to any $SU(n)$ representation. For example in $SU(3)$

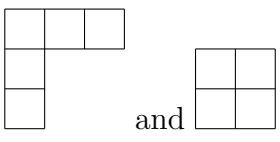


. (3.9.6)

We obtain the conjugate representation of a given representation by replacing m_i and q_i with

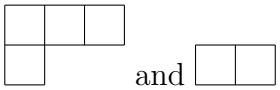
$$\bar{m}_i = m_1 - m_{n-i+1},\ \bar{q}_i = \bar{m}_i - \bar{m}_{i+1} = m_{n-i} - m_{n-i+1} = q_{n-i}. \tag{3.9.7}$$

Geometrically the Young tableau of a representation and its conjugate representation (after rotation) fit together to form a rectangular Young tableau with n rows. For example, in $SU(3)$

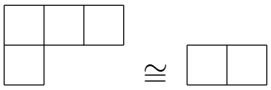


and (3.9.8)

are conjugate representations. In $SU(2)$ each representation is unitarily equivalent to its conjugate representation (i.e. the representations are real or pseudo-real). For example

 (3.9.9)

are conjugate representations, but

 (3.9.10)

This is not the case for higher n. The dimension of a representation and its conjugate representation are identical

$$D^n_{\bar{m}_1,\bar{m}_2,\ldots,\bar{m}_n} = D^n_{m_1,m_2,\ldots,m_n}. \tag{3.9.11}$$

For general n the adjoint representation is given by $q_1 = q_{n-1} = 1$, $q_i = 0$ otherwise, and it is identical with its conjugate representation. The dimension of the adjoint representation is

$$D^n_{2,1,1,\ldots,1,0} = n^2 - 1. \tag{3.9.12}$$

Next we want to discuss a method to couple $SU(n)$ representations by operating on their Young tableaux. Two Young tableaux with N and M boxes are coupled by forming an external product. In this way we generate Young tableaux with $N + M$ boxes that can then be translated back into $SU(n)$ representations. The external product is built as follows. The boxes of the first row of the second Young tableau are labeled with 'a', the boxes of the second row with 'b', etc. Then the boxes labeled with 'a' are added to the first Young tableau in all possible ways that lead to new allowed Young tableau, such that two 'a' never end up in the same column. Then the 'b' boxes are added to the resulting Young tableaux in the same way. Now each of the resulting tableaux is read row-wise from top-right to bottom-left. Whenever a 'b' or 'c' appears before the first 'a', or a 'c' occurs before the first 'b' etc., the corresponding Young tableau is deleted. The remaining tableaux form the reduction of the external product.

We now want to couple N fundamental representations of $SU(n)$. In Young tableau language this reads

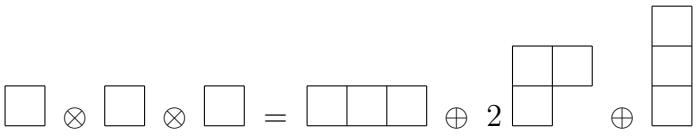
$$\{n\} \otimes \{n\} \otimes \ldots \otimes \{n\} = \square \otimes \square \otimes \ldots \otimes \square\,. \tag{3.9.13}$$

In this way we generate all irreducible representations of S_N, i.e. all Young tableaux with N boxes. Each Young tableau is associated with an $SU(n)$ multiplet. It occurs

in the product as often as the dimension of the corresponding S_N representation indicates, i.e. $d_{m_1,m_2,\ldots,m_n}$ times. Hence we can write

$$\{n\} \otimes \{n\} \otimes \ldots \otimes \{n\} = \sum_{\Gamma} d_{m_1,m_2,\ldots,m_n} \{D^n_{m_1,m_2,\ldots,m_n}\}. \tag{3.9.14}$$

The sum goes over all Young tableaux with N boxes. For example



(3.9.15)

Translated into $SU(n)$ language this reads

$$\begin{aligned} \{n\} \otimes \{n\} \otimes \{n\} &= \left\{ \frac{n(n+1)(n+2)}{6} \right\} \oplus 2 \left\{ \frac{(n-1)n(n+1)}{3} \right\} \\ &\oplus \left\{ \frac{(n-2)(n-1)n}{6} \right\}. \end{aligned} \tag{3.9.16}$$

The dimensions test

$$\frac{n(n+1)(n+2)}{6} + 2\frac{(n-1)n(n+1)}{3} + \frac{(n-2)(n-1)n}{6} = n^3 \tag{3.9.17}$$

confirms this result. In $SU(2)$ this corresponds to

$$\{2\} \otimes \{2\} \otimes \{2\} = \{4\} \oplus 2\{2\} \oplus \{0\}, \tag{3.9.18}$$

and in $SU(3)$

$$\{3\} \otimes \{3\} \otimes \{3\} = \{10\} \oplus 2\{8\} \oplus \{1\}. \tag{3.9.19}$$

3.10 Coupling of $\{3\}$ and $\{\overline{3}\}$ in $su(3)$

In order to further illustrate the coupling of $su(n)$ representations, we now couple the fundamental $\{3\}$ representation of $su(3)$ with its conjugate representation $\{\overline{3}\}$. From the considerations of the previous section, we obtain

$$\{3\} \otimes \{\overline{3}\} = \{1\} \oplus \{8\}. \tag{3.10.1}$$

We now want to explicitly construct the singlet state and all states of the octet, by calculating the corresponding Clebsch-Gordan coefficients. First of all, we relate the states in the triplet by shift operators

$$T_-|u\rangle = |d\rangle,\ T_+|d\rangle = |u\rangle,\ U_-|d\rangle = |s\rangle,\ U_+|s\rangle = |d\rangle,\ V_-|u\rangle = |s\rangle,\ V_+|s\rangle = |u\rangle. \tag{3.10.2}$$

The representation $\{\overline{3}\}$ is obtained from $\{3\}$ by conjugation, which is realized by $\widetilde{T}^a = -(T^a)^T$, such that

$$\begin{aligned}&\widetilde{T}_-|\overline{d}\rangle = -|\overline{u}\rangle,\ \widetilde{T}_+|\overline{u}\rangle = -|\overline{d}\rangle,\ \widetilde{U}_-|\overline{s}\rangle = -|\overline{d}\rangle,\ \widetilde{U}_+|\overline{s}\rangle = -|\overline{u}\rangle,\\ &\widetilde{V}_-|\overline{s}\rangle = -|\overline{u}\rangle,\ \widetilde{V}_+|\overline{s}\rangle = -|\overline{d}\rangle.\end{aligned} \tag{3.10.3}$$

We now construct the state in the octet that has the larges value of T^3. This is the state

$$|\{8\}011\rangle = |\{3\}\frac{1}{3}\frac{1}{2}\frac{1}{2}\rangle|\{\overline{3}\} - \frac{1}{3}\frac{1}{2}\frac{1}{2}\rangle = u\overline{d}. \tag{3.10.4}$$

Starting from this state, we can now reach the other ones in the multiplet by applying the appropriate shift operations. In this way, we obtain

$$-\sqrt{2}|\{8\}010\rangle = (T_- + \widetilde{T}_-)u\overline{d} = -u\overline{u} + d\overline{d} \ \Rightarrow\ |\{8\}010\rangle = \frac{1}{\sqrt{2}}(u\overline{u} - d\overline{d}), \tag{3.10.5}$$

as well as

$$\sqrt{2}|\{8\}01-1\rangle = (T_- + \widetilde{T}_-)\frac{1}{\sqrt{2}}(u\overline{u} - d\overline{d}) = \sqrt{2}d\overline{u} \ \Rightarrow\ |\{8\}010\rangle = d\overline{u}. \tag{3.10.6}$$

By applying the shift operator U_+ one finds

$$-|\{8\}1\frac{1}{2}\frac{1}{2}\rangle = (U_+ + \widetilde{U}_+)u\overline{d} = -u\overline{s} \ \Rightarrow\ |\{8\}1\frac{1}{2}\frac{1}{2}\rangle = u\overline{s}, \tag{3.10.7}$$

and in complete analogy we obtain

$$|\{8\}1\frac{1}{2}-\frac{1}{2}\rangle = d\overline{s},\ |\{8\}-1\frac{1}{2}\frac{1}{2}\rangle = s\overline{d},\ |\{8\}-1\frac{1}{2}-\frac{1}{2}\rangle = s\overline{u}. \tag{3.10.8}$$

We are still lacking the state $|\{8\}000\rangle$, which we reach by applying

$$(V_- + \widetilde{V}_-)|\{8\}1\frac{1}{2}\frac{1}{2}\rangle = (V_- + \widetilde{V}_-)u\overline{s} = s\overline{s} - u\overline{u} = \alpha|\{8\}000\rangle + \beta|\{8\}010\rangle. \tag{3.10.9}$$

Also demanding normalization and orthogonality, i.e. $\langle\{8\}000|\{8\}010\rangle = 0$, one then obtains

$$|\{8\}000\rangle = \frac{1}{\sqrt{6}}(u\overline{u} + d\overline{d} - 2s\overline{s}). \tag{3.10.10}$$

The last remaining state represents the singlet $|\{1\}000\rangle$. Again demanding normalization and orthogonality, one finally obtains

$$|\{1\}000\rangle = \frac{1}{\sqrt{3}}(u\overline{u} + d\overline{d} + s\overline{s}). \tag{3.10.11}$$

3.11 The Orthogonal Group $SO(n)$ and its Algebra

The real-valued $n\times n$ orthogonal matrices O with determinant 1 obey $OO^T = O^TO = \mathbf{1}$ and form the group $SO(n)$ under matrix multiplication. The corresponding $so(n)$ algebra consists of the purely imaginary traceless Hermitean $n \times n$ matrices. There are $n_G = n(n-1)/2$ such matrices. The algebra $so(4) = su(2) \times su(2)$ is the direct Kronecker product of two $su(2)$ algebras and thus semi-simple (but not simple).

As we discussed before, the group $SO(3)$ has a trivial center, while its universal covering group $SU(2)$ has the non-trivial center $\mathbb{Z}(2)$. The universal covering group of $SO(n)$ is called $Spin(n)$, such that $Spin(3) = SU(2)$. Similarly, the universal covering group of $SO(4)$ is $Spin(4) = SU(2) \times SU(2)$, which has the center $\mathbb{Z}(2) \times \mathbb{Z}(2)$. The center of $SO(4)$ itself, on the other hand, is just $\mathbb{Z}(2)$ and consists on the 4×4 unit-matrix $\mathbf{1}$ and $-\mathbf{1}$. Since for $n = 5$ the matrix $-\mathbf{1}$ does not have determinant 1, the group $SO(5)$ has a trivial center, while its universal covering group $Spin(5)$ has the center $\mathbb{Z}(2)$. The $so(6)$ algebra has $n_G = 6 \cdot 5/2 = 15$ generators. This is the same number as for $su(4)$ which has $4^2 - 1$ generators. Indeed, one can show that the algebras of $so(6)$ and $su(4)$ are identical. As it was also the case for $so(3)$ and $su(2)$, the corresponding groups $SO(6)$ and $SU(4)$ are still different. In particular, the center of $SO(6)$ is $\mathbb{Z}(2)$ while the center of $SU(4)$ is $\mathbb{Z}(4)$. Indeed, the universal covering group of $SO(6)$ is $Spin(6) = SU(4)$, and $SU(4)/\mathbb{Z}(2) = SO(6)$. The higher $so(n)$ algebras (with $n \geq 7$) are not equivalent to an $su(n)$ algebra. The center of $Spin(n)$ is $\mathbb{Z}(2)$ for odd n, $\mathbb{Z}(2) \times \mathbb{Z}(2)$ for $n = 4, 8, 12, \ldots$, and $\mathbb{Z}(4)$ for $n = 6, 10, 14, \ldots$.

The group manifold of $Spin(n)$ is the product of spheres

$$Spin(n) = S^1 \times S^2 \times \ldots \times S^{n-1}. \tag{3.11.1}$$

Based on the so-called Hopf fibration one can show that

$$S^3/S^1 = S^2, \quad S^7/S^3 = S^4, \quad S^{15}/S^7 = S^8, \tag{3.11.2}$$

such that at least locally

$$S^3 = S^1 \times S^2, \quad S^7 = S^3 \times S^4, \quad S^{15} = S^7 \times S^8. \tag{3.11.3}$$

Consequently, one obtains

$$\begin{aligned} &Spin(3) = S^1 \times S^2 = S^3 = SU(2), \\ &Spin(6) = S^1 \times S^2 \times S^3 \times S^4 \times S^5 = S^3 \times S^5 \times S^7 = SU(4). \end{aligned} \tag{3.11.4}$$

The $n(n-1)/2$-dimensional adjoint representation of $so(n)$ transforms as an anti-symmetric tensor under rotations in n dimensions. Similarly, there is an $[n(n+1)/2-1]$-dimensional representation that corresponds to a symmetric traceless tensor. In addition, $so(n)$ has an n-dimensional vector representation. Since in three dimensions the vector cross product again generates a vector instead of an anti-symmetric tensor, for $so(3)$ the vector representation is equivalent to the adjoint. The $so(n)$ algebras also have spinor representations. While $so(3) = su(2)$ only has a single 2-dimensional spinor representation $\{2\}$, which corresponds to an ordinary spin $\frac{1}{2}$, $so(4) = su(2) \times su(2)$ has two 2-dimensional spinor representations $\{2\} \times \{1\}$ and $\{1\} \times \{2\}$, which are both pseudo-real. The algebra $so(5)$ has a single 4-dimensional fundamental spinor representation, while $so(6) = su(4)$ has two in-equivalent 4-dimensional spinor representations, which correspond to the fundamental representation $\{4\}$ of $su(4)$ and its conjugate $\{\overline{4}\}$. In fact, the $so(n)$ algebras with $n = 6, 10, 14, \ldots$ are the only ones that have complex representations. Continuing this scheme, the algebra $so(7)$ has a single 8-dimensional spinor representation, while $so(8)$ has two in-equivalent 8-dimensional pseudo-real spinor representations (in addition to its 8-dimensional vector representation). Similarly, $so(9)$ has one 16-dimensional spinor representations, while $so(10)$ has two in-equivalent 16-dimensional complex spinor representations, which are conjugate to each other. One of these representations plays an important role in grand unified theories of the electroweak and strong interaction. It contains one generation of chiral Standard Model fermions (including quarks and leptons).

3.12 The Symplectic Group $Sp(n)$ and its Algebra

The group $Sp(n)$ is a subgroup of $SU(2n)$ which leaves the skew-symmetric matrix

$$J = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix} = i\sigma^2 \otimes \mathbf{1}, \tag{3.12.1}$$

invariant. Here $\mathbf{1}$ is the $n \times n$ unit-matrix and σ^2 is the imaginary Pauli matrix. The elements $U \in SU(2n)$ that belong to the subgroup $Sp(n)$ satisfy the constraint

$$U^* = JUJ^\dagger. \tag{3.12.2}$$

Consequently, U and U^* are related by the unitary transformation J. Hence the $2n$-dimensional fundamental representation of $Sp(n)$ is pseudo-real. The matrix J itself also belongs to $Sp(n)$. It is easy to show that the matrices that obey the constraint eq.(3.12.2) form a group under matrix multiplication. The constraint implies the following form of a generic $Sp(n)$ matrix

$$U = \begin{pmatrix} W & X \\ -X^* & W^* \end{pmatrix}, \tag{3.12.3}$$

where W and X are complex $n \times n$ matrices. Since U must still be an element of $SU(2n)$, these matrices must satisfy $WW^\dagger + XX^\dagger = \mathbf{1}$ and $WX^T = XW^T$. Note that the eigenvalues of U come in complex conjugate pairs. Since center elements are multiples of the unit-matrix, in this case eq.(3.12.3) immediately implies $W = W^*$. Hence, the center of $Sp(n)$ is $\mathbb{Z}(2)$ for all n. In fact, $Sp(n)$ is its own universal covering group and thus acquires no further central extension.

Writing $U = \exp(iH)$, where H is a Hermitean traceless matrix, eq.(3.12.2) implies that the generators H of the algebra $sp(n)$ satisfy the constraint

$$H^* = -JHJ^\dagger = JHJ. \tag{3.12.4}$$

This relation leads to the following generic form,

$$H = \begin{pmatrix} A & B \\ B^* & -A^* \end{pmatrix}, \tag{3.12.5}$$

where A and B are $n \times n$ matrices. The Hermiticity condition $H = H^\dagger$ implies $A = A^\dagger$ and $B = B^T$. Note that, since A is Hermitean, H is automatically traceless. The Hermitean $n \times n$ matrix A has n^2 degrees of freedom and the complex symmetric $n \times n$ matrix B has $(n+1)n$ degrees of freedom. Hence the dimension of the $sp(n)$ algebra is $n_G = n^2 + (n+1)n = (2n+1)n$. There are n independent diagonal generators of the maximal Abelian Cartan sub-algebra. Hence the rank of $Sp(n)$ is $r = n$.

The $n = 1$ case is equivalent to $so(3)$, while the $n = 2$ case is equivalent to $so(5)$. Since the group $Sp(n)$ has the center $\mathbb{Z}(2)$ while $SO(3)$ and $SO(5)$ have a trivial center, the group $Sp(1)$ corresponds to the universal covering group $Spin(3) = SU(2)$ of $SO(3)$, and the group $Sp(2)$ is the universal covering group $Spin(5)$ of $SO(5)$. Although both $sp(n)$ and $so(2n+1)$ have the same number of $(2n+1)n$ generators, the two algebras are in-equivalent for $n \geq 3$.

Since $sp(2)$ has rank $r = 2$, the weight diagrams of its representations can be drawn in a 2-d plane. The weight diagrams of the fundamental representation $\{4\}$, the $so(5)$ vector representation $\{5\}$, and the adjoint representation $\{10\}$ are depicted in figures 3.12, 3.12, and 3.12, respectively. It is instructive to work out the Speiser scheme for the algebra $sp(2) = so(5)$ in analogy to the $su(3)$ scheme discussed before.

figure=rep4.eps,width=6cm

Figure 3.1: *The weight diagram for the fundamental* $\{4\}$ *representation of* $sp(2)$.

figure=rep5.eps,width=6cm

Figure 3.2: *The weight diagram for the* $\{5\}$ *representation of* $sp(2)$ *(the vector representation of* $so(5)$*).*

figure=rep10.eps,width=6cm

Figure 3.3: *The weight diagram for the adjoint* $\{10\}$ *representation of* $sp(2)$ *(and of* $so(5)$*).*

The group manifold of $Sp(n)$ is the product of spheres

$$Sp(n) = S^3 \times S^7 \times \ldots \times S^{4n-1}, \tag{3.12.6}$$

which implies

$$Sp(1) = S^3 = SU(2), \quad Sp(2) = S^3 \times S^7 = S^1 \times S^2 \times S^3 \times S^4 = Spin(5). \tag{3.12.7}$$

On the other hand, since $S^5 \times S^6 \neq S^{11}$ we have

$$Sp(3) = S^3 \times S^7 \times S^{11} = S^1 \times S^2 \times S^3 \times S^4 \times S^{11} \neq S^1 \times S^2 \times S^3 \times S^4 \times S^5 \times S^6 = Spin(7). \tag{3.12.8}$$

3.13 The Exceptional Group $G(2)$ and its Algebra

In this section we discuss some basic properties of the Lie group $G(2)$ — the simplest among the exceptional groups $G(2)$, $F(4)$, $E(6)$, $E(7)$ and $E(8)$ — which do not fit into the main sequences $SU(n)$, $Spin(n)$, and $Sp(n)$. While there is only one non-Abelian semi-simple Lie algebra of rank 1 — namely the one of $su(2) = so(3) = sp(1)$ — there are four of rank $r = 2$. These rank 2 algebras generate the groups $SU(3)$, $SU(2) \times SU(2) = Spin(4)$, $Spin(5) = Sp(2)$, and $G(2)$, which have 8, 6, 10, and 14 generators, respectively. The group $G(2)$ is of particular interest because it has a trivial center and still is its own universal covering group.

It is natural to construct $G(2)$ as a subgroup of $SO(7)$ which has rank 3 and 21 generators. The 7×7 real orthogonal matrices O of the group $SO(7)$ have determinant 1 and obey the constraint

$$O_{ab} O_{ac} = \delta_{bc}. \tag{3.13.1}$$

The $G(2)$ subgroup contains those matrices that, in addition, satisfy the cubic constraint

$$T_{abc} = T_{def} O_{da} O_{eb} O_{fc}. \tag{3.13.2}$$

Here T is a totally anti-symmetric tensor whose non-zero elements follow by anti-symmetrization from

$$T_{127} = T_{154} = T_{163} = T_{235} = T_{264} = T_{374} = T_{576} = 1. \tag{3.13.3}$$

The tensor T also defines the multiplication rules for octonions. Eq.(3.13.3) implies that eq.(3.13.2) represents 7 non-trivial constraints which reduce the 21 degrees of freedom of $SO(7)$ to the 14 parameters of $G(2)$. It should be noted that $G(2)$ inherits the reality properties of $SO(7)$: all its representations are real.

We make the following choice for the first 8 generators of $G(2)$ in the 7-dimensional fundamental representation

$$\Lambda^a = \frac{1}{\sqrt{2}} \begin{pmatrix} \lambda^a & 0 & 0 \\ 0 & -\lambda^{a*} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.13.4}$$

Here λ^a (with $a \in \{1, 2, \ldots, 8\}$) are the usual 3×3 Gell-Mann generators of $su(3)$ which indeed is a sub-algebra of $g(2)$. We have chosen the standard normalization $\mathrm{Tr}\lambda^a\lambda^b = \mathrm{Tr}\Lambda^a\Lambda^b = 2\delta_{ab}$. This representation contains the complex representations $\{3\}$ and $\{\overline{3}\}$ of $su(3)$. However, it is unitarily equivalent to a real representation. In the chosen basis of the generators it is manifest that under $SU(3)$ subgroup transformations the 7-dimensional representation decomposes into

$$\{7\} = \{3\} \oplus \{\overline{3}\} \oplus \{1\}. \tag{3.13.5}$$

Since $g(2)$ has rank $r = 2$, only two generators can be diagonalized simultaneously. In our choice of basis these are the $su(3)$ sub-algebra generators Λ^3 and Λ^8. Consequently, just as for $su(3)$, the weight diagrams of $g(2)$ representations can be drawn in a 2-dimensional plane. For example, the weight diagram of the fundamental representation is shown in figure 3.13. One notes that it is indeed a superposition of

file=figura1.eps,width=60mm,angle=0

Figure 3.4: *The weight diagram of the 7-dimensional fundamental representation of* $g(2)$ *(rescaled by a factor* $\sqrt{2}$*).*

the weight diagrams of a $\{3\}$, $\{\overline{3}\}$, and $\{1\}$ of $su(3)$. Since all $g(2)$ representation are real, the $\{7\}$ representation is equivalent to its complex conjugate. It should be

noted that the $\{3\} \oplus \{\overline{3}\}$ contained in the $\{7\}$ representation of of $g(2)$ corresponds to a real reducible 6-dimensional representation of $su(3)$.

As usual,

$$\begin{aligned}
T^+ &= \frac{1}{\sqrt{2}}(\Lambda^1 + i\Lambda^2) = |1\rangle\langle 2| - |5\rangle\langle 4|, \\
T^- &= \frac{1}{\sqrt{2}}(\Lambda^1 - i\Lambda^2) = |2\rangle\langle 1| - |4\rangle\langle 5|, \\
U^+ &= \frac{1}{\sqrt{2}}(\Lambda^4 + i\Lambda^5) = |2\rangle\langle 3| - |6\rangle\langle 5|, \\
U^- &= \frac{1}{\sqrt{2}}(\Lambda^4 - i\Lambda^5) = |3\rangle\langle 2| - |5\rangle\langle 6|, \\
V^+ &= \frac{1}{\sqrt{2}}(\Lambda^6 + i\Lambda^7) = |1\rangle\langle 3| - |6\rangle\langle 4|, \\
V^- &= \frac{1}{\sqrt{2}}(\Lambda^6 - i\Lambda^7) = |3\rangle\langle 1| - |4\rangle\langle 6|, \qquad (3.13.6)
\end{aligned}$$

define $su(3)$ shift operations between the different states $|1\rangle$, $|2\rangle$,...,$|7\rangle$ in the fundamental representation. The remaining 6 generators of $G(2)$ also define shifts

$$\begin{aligned}
X^+ &= \frac{1}{\sqrt{2}}(\Lambda^9 + i\Lambda^{10}) = |2\rangle\langle 4| - |1\rangle\langle 5| - \sqrt{2}|7\rangle\langle 3| - \sqrt{2}|6\rangle\langle 7|, \\
X^- &= \frac{1}{\sqrt{2}}(\Lambda^9 - i\Lambda^{10}) = |4\rangle\langle 2| - |5\rangle\langle 1| - \sqrt{2}|3\rangle\langle 7| - \sqrt{2}|7\rangle\langle 6|, \\
Y^+ &= \frac{1}{\sqrt{2}}(\Lambda^{11} + i\Lambda^{12}) = |6\rangle\langle 1| - |4\rangle\langle 3| - \sqrt{2}|2\rangle\langle 7| - \sqrt{2}|7\rangle\langle 5|, \\
Y^- &= \frac{1}{\sqrt{2}}(\Lambda^{11} - i\Lambda^{12}) = |1\rangle\langle 6| - |3\rangle\langle 4| - \sqrt{2}|7\rangle\langle 2| - \sqrt{2}|5\rangle\langle 7|, \\
Z^+ &= \frac{1}{\sqrt{2}}(\Lambda^{13} + i\Lambda^{14}) = |3\rangle\langle 5| - |2\rangle\langle 6| - \sqrt{2}|7\rangle\langle 1| - \sqrt{2}|4\rangle\langle 7|, \\
Z^- &= \frac{1}{\sqrt{2}}(\Lambda^{13} - i\Lambda^{14}) = |5\rangle\langle 3| - |6\rangle\langle 2| - \sqrt{2}|1\rangle\langle 7| - \sqrt{2}|7\rangle\langle 4|. \qquad (3.13.7)
\end{aligned}$$

The generators themselves transform under the 14-dimensional adjoint representation of $g(2)$ whose weight diagram is shown in figure 3.13. From this diagram one sees that under an $SU(3)$ subgroup transformation the adjoint representation of $g(2)$ decomposes into

$$\{14\} = \{8\} \oplus \{3\} \oplus \{\overline{3}\}. \qquad (3.13.8)$$

Again, it is instructive to work out the Speiser scheme for the algebra $g(2)$.

Let us now discuss the center of $G(2)$. It is interesting to note that the maximal Abelian (Cartan) subgroup of both $G(2)$ and $SU(3)$ is $U(1)^2$ which must contain

file=figura2.eps,width=60mm,angle=0

Figure 3.5: *The weight diagram of the 14-dimensional adjoint representation of* $g(2)$ *(rescaled by a factor* $\sqrt{2}$*).*

the center in both cases. Since $G(2)$ contains $SU(3)$ as a subgroup its center cannot be bigger than $\mathbb{Z}(3)$ (the center of $SU(3)$) because the potential center elements of $G(2)$ must commute with all $G(2)$ matrices (not just with the elements of the $SU(3)$ subgroup). In the fundamental representation of $G(2)$ the center elements of the $SU(3)$ subgroup are given by

$$Z = \begin{pmatrix} z\mathbf{1} & 0 & 0 \\ 0 & z^*\mathbf{1} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.13.9}$$

where $\mathbf{1}$ is the 3×3 unit matrix and $z \in \{1, \exp(\pm 2\pi i/3\}$ is an element of $\mathbb{Z}(3)$. By construction, the three 7×7 matrices Z commute with the 8 generators of the $SU(3)$ subgroup of $G(2)$. However, an explicit calculation shows that this is not the case for the remaining 6 generators. Consequently, the center of $G(2)$ is trivial and contains only the identity. The above argument applies to any representation of $G(2)$. In other words, the universal covering group of $G(2)$ is $G(2)$ itself and still it has a trivial center.

In $SU(3)$ the non-trivial center $\mathbb{Z}(3)$ gives rise to the concept of triality. For example, the trivial representation $\{1\}$ and the adjoint representation $\{8\}$ of $SU(3)$ have trivial triality, while the fundamental $\{3\}$ and anti-fundamental $\{\overline{3}\}$ have non-trivial opposite trialities. Since its center is trivial, the concept of triality does not extend to $G(2)$. In particular, as one can see from eqs.(3.13.5,3.13.8), $G(2)$ representations decompose into mixtures of $SU(3)$ representations with different trialities. This has interesting consequences for the results of $G(2)$ tensor decompositions. For example, in contrast to the $SU(3)$ case, the product of two fundamental representations

$$\{7\} \otimes \{7\} = \{1\} \oplus \{7\} \oplus \{14\} \oplus \{27\}, \tag{3.13.10}$$

contains both the trivial and the adjoint representation. Some further tensor product decompositions are given by

$$\begin{aligned} &\{7\} \otimes \{14\} = \{7\} \oplus \{27\} \oplus \{64\}, \\ &\{14\} \otimes \{14\} = \{1\} \oplus \{14\} \oplus \{27\} \oplus \{77\} \oplus \{77'\}. \end{aligned} \tag{3.13.11}$$

Here $\{77\}$ and $\{77'\}$ are two in-equivalent 77-dimensional representations of $g(2)$.

The group manifold of $G(2)$ is the product of the group manifold of $SU(3)$ with a 6-dimensional sphere S^6, i.e.

$$G(2) = SU(3) \times S^6 = S^3 \times S^5 \times S^6. \tag{3.13.12}$$

From this one obtains

$$SO(7) = S^1 \times S^2 \times S^3 \times S^4 \times S^5 \times S^6 = G(2) \times S^1 \times S^2 \times S^4 = G(2) \times S^3 \times S^4 = G(2) \times S^7. \tag{3.13.13}$$

Chapter 4

Galilei, Lorentz, and Poincaré Algebras

A non-relativistic system without a preferred rest frame corresponds to a representation of the Galilei algebra. Similarly, a relativistic system provides us with a representation of the Poincaré algebra. We consider the Galilei and Poincaré algebras in one, two, and three spatial dimensions, first without specifying a particular physical system. Then we consider systems of free particles with and without spin. We also review the concept of a center of energy — a relativistic generalization of the center of mass.

4.1 The Galilei Algebra in One, Two, and Three Dimensions

The Galilei algebra describes the symmetries of a space-time which is spatially Euclidean with absolute Newtonian time and without an a priori preferred rest frame. Euclidean space is homogeneous and isotropic, which gives rise to momentum and angular momentum as the infinitesimal generators of spatial translations and rotations. Translation invariance in time gives rise to the Hamiltonian as an infinitesimal generator of time translations. In addition, although, in this case, there is no symmetry between space and time, there are generators of Galileian boosts, which are related to the center of mass coordinate. When matter condenses in Galileian space-time, its center of mass singles out a preferred rest frame, which amounts to a spontaneous breakdown of Galileian boost invariance.

In one spatial dimension the Galilei algebra includes the Hamilton operator H,

the momentum operator P, and the boost operator K. The corresponding commutation relations are

$$[P, H] = 0,\ [K, H] = iP,\ [K, P] = i\mathcal{M}. \tag{4.1.1}$$

Here $\mathcal{M}$ represents the total mass of the system. Energy H, momentum P, as well as the mass operator $\mathcal{M}$ (which is proportional to the identity $\mathbf{1}$) generate an invariant Abelian sub-algebra. Hence, the Galilei algebra is not semi-simple. Obviously, the mass operator commutes with all generators of the algebra, and thus plays the role of a rather trivial Casimir operator.

In two spatial dimensions, in addition to the Hamiltonian there are now two momenta P_i ($i \in \{1, 2\}$), one angular momentum operator J that generates spatial rotations, as well as two Galileian boosts K_i ($i \in \{1, 2\}$). The corresponding commutation relations are

$$\begin{aligned}&[P_i, H] = 0,\ [J, H] = 0,\ [K_i, H] = iP_i,\ [P_i, P_j] = 0,\ [J, P_i] = i\epsilon_{ij}P_j,\\ &[K_i, P_j] = i\delta_{ij}\mathcal{M},\ [J, K_i] = i\epsilon_{ij}K_j,\ [K_i, K_j] = 0.\end{aligned} \tag{4.1.2}$$

In three spatial dimensions the Galilei algebra consists of the Hamiltonian H, three momenta P_i, three angular momenta J_i, and three Galileian boosts K_i ($i \in \{1, 2, 3\}$), with the commutators

$$\begin{aligned}&[P_i, H] = 0,\ [J_i, H] = 0,\ [K_i, H] = iP_i,\ [P_i, P_j] = 0,\ [J_i, P_j] = i\epsilon_{ijk}P_k,\\ &[K_i, P_j] = i\delta_{ij}\mathcal{M},\ [J_i, J_j] = i\epsilon_{ijk}J_k,\ [J_i, K_j] = i\epsilon_{ijk}K_k,\ [K_i, K_j] = 0.\end{aligned} \tag{4.1.3}$$

Introducing the center of mass coordinate

$$R_i = \frac{K_i}{\mathcal{M}}, \tag{4.1.4}$$

one obtains

$$[R_i, H] = i\frac{P_i}{\mathcal{M}} = iV_i = i\dot{R}_i, \tag{4.1.5}$$

which represents the Heisenberg equation of motion for the center of mass, with $V_i = \dot{R}_i$ being its velocity. In addition,

$$[R_i, P_j] = i\delta_{ij},\ [R_i, R_j] = 0, \tag{4.1.6}$$

which identifies the center of mass and the total momentum as canonically conjugate variables.

4.2 Representation of the Galilei Algebra with Non-Relativistic Interacting Particles

Let us consider a system of N free spinless non-relativistic particles of mass m_a with positions $\vec{r}_a$ and momenta $\vec{p}_a$ ($a \in \{1, 2, ..., N\}$) the operators of the Galilei algebra can then be represented as

$$H = \sum_{a=1}^{N} \left(m_a + \frac{\vec{p}_a{}^2}{2m_a} \right) + \sum_{a<b} V(|\vec{r}_a - \vec{r}_b|),$$
$$\vec{P} = \sum_{a=1}^{N} \vec{p}_a, \ \vec{J} = \sum_{a=1}^{N} \vec{r}_a \times \vec{p}_a, \ \vec{K} = \sum_{a=1}^{N} m_a \vec{r}_a, \ \mathcal{M} = \sum_{a=1}^{N} m_a. \quad (4.2.1)$$

Using the canonical commutation relations

$$[r_{ai}, p_{bj}] = i\delta_{ab}\delta_{ij}, \quad (4.2.2)$$

it is straightforward to show that the operators from above indeed obey the commutation relations of the Galilei algebra. This is not the most general form of the potential. In particular, besides 2-body forces, one could also include 3-body or many-body forces.

Let us also discuss particles with spins $\vec{S}_a$, obeying the commutation relation

$$[S_{ai}, S_{bj}] = i\delta_{ab}\varepsilon_{ijk}S_{ak}. \quad (4.2.3)$$

In this case, one can represent the Galilei algebra as

$$H = \sum_{a=1}^{N} \left(m_a + \frac{\vec{p}_a{}^2}{2m_a} \right) + \sum_{a<b} V(\vec{r}_a - \vec{r}_b, \vec{S}_a, \vec{S}_b),$$
$$\vec{P} = \sum_{a=1}^{N} \vec{p}_a, \ \vec{J} = \sum_{a=1}^{N} (\vec{r}_a \times \vec{p}_a + \vec{S}_a), \ \vec{K} = \sum_{a=1}^{N} m_a \vec{r}_a, \ \mathcal{M} = \sum_{a=1}^{N} m_a. \quad (4.2.4)$$

In this case, one can include spin-dependent forces. It is instructive to investigate the constraints on a spin-dependent 2-body potential $V(\vec{r}_a - \vec{r}_b, \vec{S}_a, \vec{S}_b)$, that result from the Galileian symmetry.

4.3 The Poincaré Algebra in One, Two, and Three Dimensions

In one spatial dimension the Poincaré algebra has three generators, the Hamilton operator H that generates time translations, the momentum operator P that generates

spatial translations, and the boost operator K that generates space-time rotations. The corresponding commutation relations are

$$[P, H] = 0,\ [K, H] = iP,\ [K, P] = iH. \tag{4.3.1}$$

The quadratic Casimir operator,

$$\mathcal{M}^2 = H^2 - P^2, \tag{4.3.2}$$

commutes with all generators and represents the rest mass squared. Energy and momentum are components of a Lorentz vector $P^\mu = (H, P)$, and K is an element of an anti-symmetric tensor

$$M_{\mu\nu} = \begin{pmatrix} 0 & K \\ -K & 0 \end{pmatrix}. \tag{4.3.3}$$

The commutation relations can then be expressed as

$$[P_\mu, P_\nu] = 0,\ [M_{\mu\nu}, P_\rho] = i(g_{\nu\rho}P_\mu - g_{\mu\rho}P_\nu),\ [M_{\mu\nu}, M_{\rho\sigma}] = 0, \tag{4.3.4}$$

where $g_{\mu\nu}$ is the metric of Minkowski space-time.

In two spatial dimensions the Poincaré algebra has six generators. In addition to the Hamiltonian there are now two momenta P_i ($i \in \{1, 2\}$), one angular momentum operator J that generates spatial rotations, as well as two boosts K_i ($i \in \{1, 2\}$). The corresponding commutation relations are

$$\begin{aligned} &[P_i, H] = 0,\ [J, H] = 0,\ [K_i, H] = iP_i,\ [P_i, P_j] = 0,\ [J, P_i] = i\epsilon_{ij}P_j, \\ &[K_i, P_j] = i\delta_{ij}H,\ [J, K_i] = i\epsilon_{ij}K_j,\ [K_i, K_j] = -i\epsilon_{ij}J. \end{aligned} \tag{4.3.5}$$

Again, the rest mass

$$\mathcal{M}^2 = H^2 - \vec{P}^2, \tag{4.3.6}$$

is a Casimir operator. In this case, we can write

$$M_{\mu\nu} = \begin{pmatrix} 0 & K_1 & K_2 \\ -K_1 & 0 & J \\ -K_2 & -J & 0 \end{pmatrix}, \tag{4.3.7}$$

which then implies

$$\begin{aligned} &[P_\mu, P_\nu] = 0,\ [M_{\mu\nu}, P_\rho] = i(g_{\nu\rho}P_\mu - g_{\mu\rho}P_\nu), \\ &[M_{\mu\nu}, M_{\rho\sigma}] = i(g_{\mu\sigma}M_{\nu\rho} - g_{\mu\rho}M_{\nu\sigma} - g_{\nu\sigma}M_{\mu\rho} + g_{\nu\rho}M_{\mu\sigma}). \end{aligned} \tag{4.3.8}$$

Finally, in three spatial dimensions the Poincaré algebra has ten generators: the Hamiltonian H, three momenta P_i, three angular momenta J_i, and three boosts K_i ($i \in \{1,2,3\}$), with the commutators

$$\begin{aligned}&[P_i, H] = 0,\ [J_i, H] = 0,\ [K_i, H] = iP_i,\ [P_i, P_j] = 0,\ [J_i, P_j] = i\epsilon_{ijk}P_k,\\ &[K_i, P_j] = i\delta_{ij}H,\ [J_i, J_j] = i\epsilon_{ijk}J_k,\ [J_i, K_j] = i\epsilon_{ijk}K_k,\ [K_i, K_j] = -i\epsilon_{ijk}J_k.\end{aligned} \tag{4.3.9}$$

Once again, one Casimir operator is

$$\mathcal{M}^2 = H^2 - \vec{P}^2. \tag{4.3.10}$$

Now angular momenta and boosts combine to

$$M_{\mu\nu} = \begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & J_3 & -J_2 \\ -K_2 & -J_3 & 0 & J_1 \\ -K_3 & J_2 & -J_1 & 0 \end{pmatrix}, \tag{4.3.11}$$

which implies

$$\begin{aligned}&[P_\mu, P_\nu] = 0,\ [M_{\mu\nu}, P_\rho] = i(g_{\nu\rho}P_\mu - g_{\mu\rho}P_\nu),\\ &[M_{\mu\nu}, M_{\rho\sigma}] = i(g_{\mu\sigma}M_{\nu\rho} - g_{\mu\rho}M_{\nu\sigma} - g_{\nu\sigma}M_{\mu\rho} + g_{\nu\rho}M_{\mu\sigma}).\end{aligned} \tag{4.3.12}$$

One can also define the Pauli-Lubanski vector

$$I_\mu = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma, \tag{4.3.13}$$

whose components are given by

$$I_0 = P_iJ_i,\ I_i = \varepsilon_{ijk}K_jP_k - HJ_i, \tag{4.3.14}$$

and which indeed transforms as a 4-vector

$$[P_\mu, I_\nu] = 0,\ [M_{\mu\nu}, I_\rho] = i(g_{\nu\rho}I_\mu - g_{\mu\rho}I_\nu). \tag{4.3.15}$$

whose components obey the commutation relations

$$[I_\mu, I_\nu] = i\varepsilon_{\mu\nu\rho\sigma}I^\rho P^\sigma. \tag{4.3.16}$$

By construction, $I_\mu P^\mu = 0$, however,

$$I_\mu I^\mu = I_0^2 - I_i^2 = -\mathcal{M}^2\vec{J}^2 \tag{4.3.17}$$

is another Casimir operator, representing the total spin in the rest frame.

4.4 Free Spinless Particles

For a system of N free spinless relativistic particles of rest mass m_a with positions $\vec{r}_a$ and momenta $\vec{p}_a$ ($a \in \{1, 2, ..., N\}$) the operators of the Poincaré algebra can be represented as

$$H = \sum_{a=1}^{N} \sqrt{\vec{p}_a{}^2 + m_a^2}, \;\; \vec{P} = \sum_{a=1}^{N} \vec{p}_a, \;\; \vec{J} = \sum_{a=1}^{N} \vec{r}_a \times \vec{p}_a,$$
$$\vec{K} = \sum_{a=1}^{N} \frac{1}{2}(\vec{r}_a \sqrt{\vec{p}_a{}^2 + m_a^2} + \sqrt{\vec{p}_a{}^2 + m_a^2}\; \vec{r}_a). \tag{4.4.1}$$

Using the canonical commutation relations

$$[r_{ai}, p_{bj}] = i\delta_{ab}\delta_{ij}, \tag{4.4.2}$$

it is straightforward to show that the operators from above indeed obey the commutation relations of the Poincaré algebra.

The resulting relativistic generalization of the free particle Schrödinger equation takes the form

$$i\partial_t \Psi = H\Psi = \sum_{a=1}^{N} \sqrt{-\Delta_a + m_a^2}\; \Psi, \tag{4.4.3}$$

It should be noted that the Hamiltonian in eq.(4.4.3) is by construction positive definite. Consequently, it only describes particles and no anti-particles (negative energy states). In the context of field theory this would clearly be an undesirable feature. In fact, upon second quantization a system of indistinguishable particles of mass $m_a = m$ described by eq.(4.5.2) would turn into a field theory with a non-local (and thus unacceptable) Lagrange density

$$\mathcal{L}[\Psi] = i\Psi^* \partial_t \Psi + \Psi^* \sqrt{-\Delta + m^2}\; \Psi. \tag{4.4.4}$$

In field theory a natural path to follow is the one first taken by Klein and Gordon, which leads to the local Lagrangian

$$\mathcal{L}[\Psi] = \Psi^* [\partial_t^2 - \Delta + m^2] \Psi. \tag{4.4.5}$$

The Klein-Gordon scalar field theory describes both particles and anti-particles and thus has negative energy states. As such it is perfectly well suited as a quantum field theory, but it is, in fact, unacceptable in the context of relativistic point particle quantum mechanics with a finite number of degrees of freedom. As long as we do not want to turn the particle system into a quantum field theory by second quantization, eq.(4.4.3) is perfectly acceptable from a theoretical point of view. In particular, its

non-locality is perfectly consistent with the principles of relativity and causality. It is a separate issue that Nature has chosen to make particles as quanta of fields. This implies that the field theoretical description with infinitely many degrees of freedom and with anti-particles is phenomenologically successful. Relativistic point particle mechanics with its finite number of degrees of freedom — and hence without anti-particles — is perfectly consistent from a theoretical point of view. Interestingly, as was shown by Leutwyler, it is essentially limited to non-interacting particle systems.

4.5 Free Particles with Spin

Dirac's discovery of the wave equation for the electron — which turned out to have negative energy solutions — led him to the prediction of the positron. Since a single particle interpretation of the Dirac equation is problematical, it was necessary to extend it to quantum field theory by second quantization, which then culminated in the development of QED, QCD and the standard model. Although there is no doubt that Nature chose to make particles as quanta of fields, we like to investigate relativistic quantum mechanics which has a finite number of particles built in from the start. In such a theory anti-particles are undesirable because we insist on the standard quantum mechanical interpretation of the wave function. Insisting on a positive definite Hamiltonian, we have already obtained a modification of the Klein-Gordon equation for spinless particles. We will now apply the same logic in order to derive a modification of the Dirac equation for particles with spin. We point out again that what we will get cannot be turned into a local field theory upon second quantization. If this were our goal, the Dirac equation were the way to go. However, just like the Klein-Gordon equation, the Dirac equation is not an equation of relativistic quantum mechanics, but belongs to quantum field theory. In relativistic quantum mechanics one is led down another path.

We now consider a system of N free particles with spin $\vec{S}_a$ ($a \in \{1, 2, ..., N\}$). The corresponding wave function thus has $\prod_{a=1}^{N}(2S_a+1)$ components. The commutation relations of the spins are the usual ones

$$[S_{ai}, S_{bj}] = i\delta_{ab}\epsilon_{ijk}S_{ak}. \tag{4.5.1}$$

It is again straightforward to show that the Poincaré algebra can be represented as

$$H = \sum_{a=1}^{N} \sqrt{\vec{p}_a{}^2 + m_a^2},\ \vec{P} = \sum_{a=1}^{N} \vec{p}_a,\ \vec{J} = \sum_{a=1}^{N} [\vec{r}_a \times \vec{p}_a + \vec{S}_a],$$
$$\vec{K}_a = \sum_{a=1}^{N} [\frac{1}{2}(\vec{r}_a \sqrt{\vec{p}_a{}^2 + m_a^2} + \sqrt{\vec{p}_a{}^2 + m_a^2}\ \vec{r}_a) + \frac{\vec{p}_a \times \vec{S}_a}{m_a + \sqrt{\vec{p}_a{}^2 + m_a^2}}]. \tag{4.5.2}$$

This form of the boost operator first appeared in [**?**]. The resulting analog of the free particle Schrödinger equation is again eq.(4.4.3), but now for a multi-component wave function.

4.6 The Center of Energy

For non-relativistic particle systems it is useful to separate the center of mass motion from the relative motion. For relativistic systems various generalizations of a center of mass have been discussed [**?**, **?**]. Here we consider the concept of a center of energy $\vec{R}$ as a generalization of the center of mass to relativistic systems. For a general relativistic system we define

$$\vec{R} = \frac{1}{2}(H^{-1}\vec{K} + \vec{K}H^{-1}), \tag{4.6.1}$$

which is Hermitean by construction. For a system of free non-relativistic particles this expression reduces to

$$R_i = \frac{\sum_{a=1}^{N} m_a \vec{r}_a}{\sum_{a=1}^{N} m_a}, \tag{4.6.2}$$

which is just the center of mass. Using the Poincaré algebra we obtain

$$[R_i, P_j] = \frac{1}{2}(H^{-1}[K_i, P_j] + [K_i, P_j]H^{-1}) = i\delta_{ij}\frac{1}{2}(H^{-1}H + HH^{-1}) = i\delta_{ij}, \tag{4.6.3}$$

i.e. the center of energy is indeed the canonically conjugate momentum to the total momentum $\vec{P}$. The commutator with the total angular momentum is

$$[R_i, J_j] = \frac{1}{2}(H^{-1}[K_i, J_j] + [K_i, J_j]H^{-1}) = i\epsilon_{ijk}\frac{1}{2}(H^{-1}K_k + K_kH^{-1}) = i\epsilon_{ijk}R_k, \tag{4.6.4}$$

which just confirms that $\vec{R}$ transforms as a vector under spatial rotations. We also obtain

$$[R_i, H] = \frac{1}{2}(H^{-1}[K_i, H] + [K_i, H]H^{-1}) = i\frac{1}{2}(H^{-1}P_i + P_iH^{-1}) = iV_i. \tag{4.6.5}$$

The velocity of the center of energy,

$$\vec{V} = \vec{P}H^{-1}, \tag{4.6.6}$$

is a conserved quantity, because both H and $\vec{P}$ are conserved. It enters the Heisenberg equation of motion for $\vec{R}$ as

$$\partial_t \vec{R} = -i[\vec{R}, H] = \vec{V}, \tag{4.6.7}$$

which implies that — just like the center of mass in non-relativistic mechanics — the center of energy moves with a constant velocity.

Despite these similarities with the non-relativistic center of mass, the relativistic center of energy also has some perhaps unexpected features. It is straightforward to show that

$$[R_i, R_j] = i(R_i P_j - R_j P_i - \epsilon_{ijk} J_k) H^{-2}. \tag{4.6.8}$$

Hence, the various components of the position of the center of energy are not simultaneously measurable with arbitrary precision. Introducing the angular momentum of the center of energy,

$$\vec{\mathcal{J}} = \vec{R} \times \vec{P}, \tag{4.6.9}$$

one obtains the usual commutation relations,

$$[\mathcal{J}_i, \mathcal{J}_j] = i\epsilon_{ijk} \mathcal{J}_k, \tag{4.6.10}$$

and one finds

$$[\mathcal{J}_i, H] = \epsilon_{ijk} [R_j, H] P_k = i\epsilon_{ijk} P_j H^{-1} P_k = 0, \tag{4.6.11}$$

i.e. $\vec{\mathcal{J}}$ is conserved. Furthermore one obtains

$$[\mathcal{J}_i, P_j] = i\epsilon_{ijk} P_k, \ [\mathcal{J}_i, J_j] = i\epsilon_{ijk} \mathcal{J}_k. \tag{4.6.12}$$

The relative angular momentum

$$\vec{j} = \vec{J} - \vec{\mathcal{J}}, \tag{4.6.13}$$

is conserved because both the total angular momentum $\vec{J}$ and the angular momentum of the center of energy $\vec{\mathcal{J}}$ are conserved. The commutation relations of the relative angular momentum are given by

$$[j_i, j_j] = [J_i - \mathcal{J}_i, J_j - \mathcal{J}_j] = i\epsilon_{ijk}(J_k - 2\mathcal{J}_k + \mathcal{J}_k) = i\epsilon_{ijk} j_k. \tag{4.6.14}$$

Eq.(4.6.8) can now be written as

$$[R_i, R_j] = -i\epsilon_{ijk} j_k H^{-2}. \tag{4.6.15}$$

Hence, two components of the position of the center of energy are simultaneously measurable only if the relative angular momentum $\vec{j}$ vanishes. It should be noted that the center of energy can be modified to what has been called the "center of spin" [**?**, **?**]

$$\vec{Q} = \vec{R} - \frac{\vec{P} \times \vec{j}}{H^2 - \vec{P}^2 + H\sqrt{H^2 - \vec{P}^2}}, \tag{4.6.16}$$

which obeys $[Q_i, P_j] = i\delta_{ij}$ as well as $[Q_i, Q_j] = 0$.

Chapter 5

Abelian and Non-Abelian Gauge Fields

In this chapter we introduce the concept of non-Abelian gauge fields. Abelian gauge fields are familiar from classical electrodynamics. The quantization of Abelian gauge fields leads to quantum electrodynamics (QED), with photons being the field quanta that mediate the electromagnetic interaction between charged particles. QED is embedded in the Standard Model of particle physics, which also incorporates the weak and the strong interactions. The gauge group of the Standard Model is $SU(3)_c \times SU(2)_L \times U(1)_Y$, which thus contains non-Abelian local symmetries. The concept of non-Abelian local symmetry was first described in an unpublished letter of Wolfgang Pauli to Abraham Pais. The first paper introducing $SU(2)$ gauge theories is the ground-breaking work of Chen-Ning Yang and Richard Mills in 1954. It still took until the end of the 1960s to realize that the fundamental forces between elementary particles are indeed governed by non-Abelian gauge fields. The quanta of the $SU(3)_c$ gauge field are the gluons that mediate the strong interaction between quarks. In contrast to photons, due to the non-Abelian nature of the group $SU(3)$, gluons not only couple to quarks but also to other gluons. In other words, non-Abelian gauge fields themselves carry "charge", while Abelian gauge fields are neutral. Remarkably, non-Abelian local symmetries also arise in other physical systems, including the Pauli equation and the adiabatic Berry phase. Here we introduce non-Abelian gauge fields in the context of quantum mechanics rather than quantum field theory.

5.1 From Abelian $U(1)$ to Non-Abelian $SU(2)$ Gauge Fields

In appendix E we have summarized the relativistic formulation of classical electrodynamics, and in appendix D we have coupled a non-relativistic charged particle to an external electromagnetic field, both at the classical and at the quantum level. The corresponding Schrödinger equation takes the form

$$iD_t\Psi(x) = -\frac{1}{2M}\vec{D}\cdot\vec{D}\Psi(x), \tag{5.1.1}$$

with the covariant derivatives given by

$$D_t\Psi(x) = \partial_t\Psi(x) - ie\Phi(x)\Psi(x), \quad \vec{D}\Psi(x) = \vec{\nabla}\Psi(x) + ie\vec{A}(x)\Psi(x). \tag{5.1.2}$$

Here x represents the space-time point $(\vec{x}, t)$. The Schrödinger equation is invariant under the Abelian gauge transformations

$$\Phi(x)' = \Phi(x) + \partial_t\varphi(x), \quad \vec{A}(x)' = \vec{A}(x) - \vec{\nabla}\varphi(x), \quad \Psi(x)' = \exp(ie\varphi(x))\Psi(x). \tag{5.1.3}$$

Under this transformation, both sides of the Schrödinger equation change by a factor $\exp(ie\varphi(x))$. Canceling this factor out, the equation remains invariant. This is in contrast to the Schrödinger equation of a free particle

$$i\partial_t\Psi(x) = -\frac{1}{2M}\vec{\nabla}\cdot\vec{\nabla}\Psi(x), \tag{5.1.4}$$

which is invariant only against global $U(1)$ phase rotations

$$\Psi(x)' = \exp(i\phi)\Psi(x). \tag{5.1.5}$$

We now want to develop the concept of non-Abelian gauge fields by considering a quantum mechanical particle whose wave function is described by a 2-component spinor

$$\Psi(x) = \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \end{pmatrix}. \tag{5.1.6}$$

We are familiar with 2-component Pauli spinors representing the spin degree of freedom. Here we do not necessarily identify the two components with spin. The corresponding Schrödinger equation of a free particle then takes the form

$$i\partial_t \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \end{pmatrix} = -\frac{1}{2M}\vec{\nabla}\cdot\vec{\nabla}\begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \end{pmatrix}, \tag{5.1.7}$$

which is again invariant against the global phase transformations of eq.(5.1.5). In addition, the Schrödinger equation is invariant under global $SU(2)$ transformations

$$\Psi(x)' = U\Psi(x), \quad U \in SU(2). \tag{5.1.8}$$

We now want to promote the $SU(2)$ invariance to a local symmetry, i.e. we want to modify the Schrödinger equation such that it becomes invariant against $SU(2)$ gauge transformations $U(x)$. First of all, we then have

$$\partial_\mu \Psi(x)' = \partial_\mu U(x)\Psi(x) + U(x)\partial_\mu \Psi(x). \tag{5.1.9}$$

In analogy to Abelian gauge theory, we would like to construct a covariant derivative that transforms as

$$D_\mu \Psi(x)' = U(x) D_\mu \Psi(x). \tag{5.1.10}$$

For this purpose, we make the ansatz

$$D_\mu \Psi(x) = (\partial_\mu + W_\mu(x))\Psi(x). \tag{5.1.11}$$

The appropriate transformation behavior of eq.(5.1.10) emerges if

$$W_\mu(x)' = U(x)(W_\mu(x) + \partial_\mu)U(x)^\dagger, \tag{5.1.12}$$

which is the gauge transformation behavior of a non-Abelian vector potential. Up to a factor i (which makes $W_\mu(x)$ anti-Hermitean) the gauge field itself takes values in the $su(2)$ algebra, i.e.

$$W_\mu(x) = igW_\mu^a(x)T^a, \quad T^a = \frac{1}{2}\sigma^a. \tag{5.1.13}$$

Here g is the non-Abelian charge (the analog of the fundamental electric charge e) which controls the strength of the coupling between non-Abelian charges. An $SU(2)$ gauge covariant version of the Schrödinger equation then takes the form

$$iD_t\Psi(x) = -\frac{1}{2M}\vec{D}\cdot\vec{D}\Psi(x), \quad \Psi(x) = \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \end{pmatrix}. \tag{5.1.14}$$

5.2 $SU(2)_s \times U(1)_{em}$ Gauge Symmetry of the Pauli Equation

Up to corrections of order $1/M^3$ (where M is the electron mass) the Pauli equation (i.e. the non-relativistic reduction of the Dirac equation to its upper components) takes the form

$$i(\partial_t - ie\Phi + i\frac{e}{8M^2}\vec{\nabla}\cdot\vec{E} + i\frac{e}{2M}\vec{B}\cdot\vec{\sigma})\Psi = -\frac{1}{2M}(\vec{\nabla} + ie\vec{A} - i\frac{e}{4M}\vec{E}\times\vec{\sigma})^2\Psi. \tag{5.2.1}$$

Here $\Psi(x)$ is a 2-component Pauli spinor, $\vec{\sigma}$ are the Pauli matrices, $\Phi(x)$ and $\vec{A}(x)$ are the electromagnetic scalar and vector potentials, and

$$\vec{E}(x) = -\vec{\nabla}\Phi(x) - \partial_t \vec{A}(x), \quad \vec{B}(x) = \vec{\nabla} \times \vec{A}(x), \tag{5.2.2}$$

are the usual electromagnetic field strengths. The first two terms on the left-hand side of eq.(5.2.1) form the $U(1)_{em}$ covariant derivative familiar from QED. The third (Darwin) and fourth (Zeeman) term on the left-hand side represent relativistic corrections. The first two terms on the right-hand side again form an ordinary $U(1)_{em}$ covariant derivative, while the third term represents the relativistic spin-orbit coupling. The Pauli equation transforms covariantly under $U(1)_{em}$ gauge transformations

$$\Psi(x)' = \exp(ie\varphi(x))\Psi(x), \quad \Phi(x)' = \Phi(x) + \partial_t \varphi(x), \quad \vec{A}(x)' = \vec{A}(x) - \vec{\nabla}\varphi(x). \tag{5.2.3}$$

Obviously, it is also covariant under global spatial rotations

$${}^{O}\Psi(\vec{x},t) = U\Psi(O\vec{x},t), \quad {}^{O}\Phi(\vec{x},t) = \Phi(O\vec{x},t), \quad {}^{O}\vec{A}(\vec{x},t) = O^T\vec{A}(O\vec{x},t). \tag{5.2.4}$$

Here O is a general orthogonal 3×3 rotation matrix with

$$O_{ab} = \frac{1}{2}\mathrm{Tr}(U\sigma^a U^\dagger \sigma^b), \tag{5.2.5}$$

where $U \in SU(2)_s$ represents the rotation $O \in SO(3)$ in spinor space.

As was noted by Fröhlich and Studer, the Pauli equation for an electron in an external electromagnetic field indeed has more than just the electromagnetic $U(1)_{em}$ gauge symmetry. Quite remarkably the $SU(2)_s$ spin symmetry manifests itself as a local symmetry. Up to order M^{-3} the Pauli equation then takes the form

$$iD_t\Psi(x) = -\frac{1}{2M}\vec{D}\cdot\vec{D}\Psi(x), \tag{5.2.6}$$

with the $SU(2)_s \times U(1)_{em}$ covariant derivatives given by

$$D_t = \partial_t - ie\Phi(x) + W_0(x), \quad \vec{D} = \vec{\nabla} + ie\vec{A}(x) + \vec{W}(x). \tag{5.2.7}$$

Here $\Phi(x)$ and $\vec{A}(x)$ are the ordinary electromagnetic scalar and vector potentials. The components of the non-Abelian scalar and vector potential can be identified as the electromagnetic field strengths $\vec{E}(x)$ and $\vec{B}(x)$, i.e.

$$W_0^a(x) = \mu B^a(x), \quad W_i^a(x) = \frac{\mu}{2}\varepsilon_{iab}E^b(x). \tag{5.2.8}$$

The anomalous magnetic moment $\mu = g_e e/2M$ of the electron (where, up to tiny QED corrections, $g_e = 2$) appears as a non-Abelian gauge coupling. The Abelian

vector potential $A_\mu(x)$ is the usual one, except for a small additional contribution to the scalar potential due to the Darwin term,

$$A_0(x) = -\Phi(x) + \frac{1}{8M^2}\vec{\nabla}\cdot\vec{E}(x). \tag{5.2.9}$$

Hence, somewhat unexpectedly, the Pauli equation also transforms covariantly under local $SU(2)_s$ transformations

$$\Psi(x)' = U(x)\Psi(x), \quad W_\mu(x)' = U(x)(W_\mu(x) + \partial_\mu)U(x)^\dagger. \tag{5.2.10}$$

It should be pointed out that $SU(2)_s$ is not a gauge symmetry in the usual sense. In particular, the non-Abelian vector potential $W_\mu(x)$ is not an independent degree of freedom, but just given in terms of the external electromagnetic field strengths $\vec{E}(x)$ and $\vec{B}(x)$. The local $SU(2)_s$ symmetry is related to the global spatial rotations discussed before. In particular, global $SU(2)_s$ transformations take the form

$$\Psi(x)' = U\Psi(x), \quad W_\mu(x)' = UW_\mu(x)U^\dagger, \tag{5.2.11}$$

which implies

$$\vec{E}(x)' = O^T\vec{E}(x), \quad \vec{B}(x)' = O^T\vec{B}(x), \tag{5.2.12}$$

where the resulting 3×3 rotation matrix $O \in SO(3)$ is given by eq.(5.2.5). In contrast to a full spatial rotation, a global $SU(2)_s$ transformation does not rotate the argument $\vec{x}$ of the fields to $O\vec{x}$. Also the potentials $\Phi(x)$ and $\vec{A}(x)$ are unaffected by the global $SU(2)_s$ symmetry. Consequently, the $SU(2)_s$ symmetry is inconsistent with the Maxwell equations. Despite this, the local $SU(2)_s$ symmetry of the Pauli equation dictates how external electromagnetic fields couple to particles with spin. In particular, the local $SU(2)_s$ structure implies that in non-relativistic systems spin plays the role of an internal quantum number analogous to flavor in particle physics.

5.3 Non-Abelian $SU(2)$ Field Strength Tensor

Up to now, the gauge field A_μ appeared only as an external field. We have not yet introduced a kinetic term for it. From classical electrodynamics we indeed know such a term. We construct the field strength tensor

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x), \tag{5.3.1}$$

which is the obvious gauge invariant quantity to be built from first derivatives of $A_\mu(x)$,

$$\begin{aligned} F'_{\mu\nu}(x) &= \partial_\mu A'_\nu(x) - \partial_\nu A'_\mu(x) = \partial_\mu A_\nu(x) + \partial_\mu\partial_\nu\varphi(x) - \partial_\nu A_\mu(x) - \partial_\nu\partial_\mu\varphi(x) \\ &= F_{\mu\nu}(x). \end{aligned} \tag{5.3.2}$$

The Lagrangian of the free electromagnetic field reads

$$\mathcal{L} = \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{5.3.3}$$

In the classical limit this Lagrangian leads to the inhomogeneous Maxwell equations

$$\partial^\mu F_{\mu\nu} = 0, \tag{5.3.4}$$

while the homogeneous Maxwell equations are automatically implied by the use of the 4-vector potential A_μ.

The field strength tensor of a non-Abelian gauge field is given by

$$W_{\mu\nu} = D_\mu W_\nu - D_\nu W_\mu = \partial_\mu W_\nu - \partial_\nu W_\mu + [W_\mu, W_\nu], \tag{5.3.5}$$

and it transforms as

$$W'_{\mu\nu}(x) = U(x) W_{\mu\nu}(x) U(x)^\dagger. \tag{5.3.6}$$

We see that it is natural to add the commutator term to $W_{\mu\nu}$, since it transforms in the same way as the other terms. Moreover, it is consistent to use the covariant derivative also for the formulation of the field strength. Hence, we may consider this as the general form of a field strength. The case of a $U(1)$ gauge field that we discussed before in eq.(5.3.1) was just the special situation where the commutator vanishes. The presence of a commutator term in $W_{\mu\nu}$ has important consequences: in contrast to Abelian gauge fields, non-Abelian gauge fields are themselves charged. Hence they interact among each other, even if no other charged fields are present.

In analogy to the Abelian gauge theory, eq.(5.3.3), we write

$$\mathcal{L} = \frac{1}{4g^2} W^{a\mu\nu} W^a_{\mu\nu} = -\frac{1}{2g^2} \operatorname{Tr} W^{\mu\nu} W_{\mu\nu} \ , \tag{5.3.7}$$

which is indeed gauge invariant, and $W_{\mu\nu} = igW^a_{\mu\nu} \frac{\sigma^a}{2}$.

5.4 General Non-Abelian Gauge Fields

It is straightforward to generalize the structure of non-Abelian gauge theories to arbitrary compact semi-simple Lie groups beyond $SU(2)$. A general non-Abelian gauge field is anti-Hermitean and takes values in some Lie algebra with the generators T^a

$$G_\mu(x) = igG^a_\mu(x) T^a. \tag{5.4.1}$$

For example, the gluon field in the Standard Model is an $SU(3)$ non-Abelian gauge field with the generators $T^a = \frac{1}{2}\lambda^a$ given in terms of the Gell-Mann matrices. Under group-valued gauge transformations $U(x) \in G$ a non-Abelian vector potential transforms as

$$G_\mu(x)' = U(x)(G_\mu(x) + \partial_\mu)U(x)^\dagger. \tag{5.4.2}$$

The corresponding field strength, which is given by

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu + [G_\mu, G_\nu], \tag{5.4.3}$$

transforms as

$$G'_{\mu\nu}(x) = U(x)G_{\mu\nu}(x)U(x)^\dagger. \tag{5.4.4}$$

This renders the corresponding Lagrange density

$$\mathcal{L} = -\frac{1}{2g^2}\,\mathrm{Tr}\, G^{\mu\nu}G_{\mu\nu} \tag{5.4.5}$$

gauge invariant.

Chapter 6

Topology of Gauge Fields

Gauge transformations can be viewed as maps from space-time into the corresponding gauge group. Such maps can be classified topologically. Maps that can be deformed into each other continuously are topologically equivalent, and thus fall into the same equivalence class, a so-called homotopy class. The various homotopy classes are characterized by a topological winding number which takes values in the so-called homotopy group. The non-trivial topology of gauge fields has important consequences both in particle and in condensed matter physics.

6.1 Maps from S^d to S^n

Let us first consider maps from a circle S^1, which we could view, for example, as a compactified 1-d space $x \in [0, L]$ with periodic boundary conditions, into another circle S^1, which we can interpret as the group manifold of $U(1)$. Indeed, $U(1)$ gauge transformations $U(x) = \exp(i\varphi(x)) \in U(1)$ in a 1-d compactified space can be viewed as maps $x \to U(x)$ from S^1 to S^1. Such maps fall in topologically distinct equivalence classes. Two maps are equivalent if they can be continuously deformed into each other. The equivalence classes are known as homotopy classes, which are characterized by a topological winding number

$$\begin{aligned} n[U] &= \frac{1}{2\pi i}\int_{S^1} dx\; U(x)^*\partial_x U(x) = \frac{1}{2\pi}\int_0^L dx\; \partial_x\varphi(x) = \frac{1}{2\pi}[\varphi(L)-\varphi(0)] \\ &\in \Pi_1[U(1)] = \mathbb{Z}. \end{aligned} \tag{6.1.1}$$

Periodic boundary conditions mean that $U(L) = U(0)$ which implies $\varphi(L) - \varphi(0) = 2\pi n[U]$, where $n[U] \in \mathbb{Z}$ is an integer topological winding number that characterizes the homotopy class of the gauge transformation. The winding numbers $n[U]$

form the homotopy group $\Pi_1[U(1)] = \mathbb{Z}$. The winding numbers are additive under multiplication of the corresponding gauge transformations, i.e.

$$n[U] = n[U_1] + n[U_2], \tag{6.1.2}$$

where $U(x) = U_1(x)U_2(x)$.

The constant gauge transformation $U(x) = 1$ is topologically trivial, i.e. $n[U] = 0$. This gauge transformation maps all points x in the compactified space onto the same point $U(x) = 1$. All continuous deformations of this map also have $n[U] = 0$, and are thus again topologically trivial. Let us also consider the identity map $U(x) = \exp(2\pi i x/L)$, which covers each point in the image S^1 exactly once, and thus has $n[U] = 1$. Similarly, we can specify a member $U(x) = \exp(2\pi i n x/L)$, for each homotopy class characterized by $n[U] = n$. This map covers each point in the image S^1 exactly n times.

The sphere S^2 is not a group manifold. Hence we cannot interpret maps from S^2 to S^2 as gauge transformations. Such maps occur, for example, in the effective field theory for magnets. The direction of the magnetization at a point x is described by a 3-component unit-vector $\vec{e}(x) \in S^2$. When we compactify space to a 2-d sphere $x \in S^2$, the magnetization defines a map $x \to \vec{e}(x)$ from S^2 to S^2. Such maps again fall into homotopy classes characterized by a winding number

$$n[\vec{e}] = \frac{1}{8\pi} \int d^2x \; \varepsilon_{ij} \vec{e} \cdot (\partial_i \vec{e} \times \partial_j \vec{e}) \in \Pi_2[S^2] = \mathbb{Z}. \tag{6.1.3}$$

It is instructive to convince oneself that this expression indeed always takes integer values. The winding number $n[\vec{e}]$ counts topological excitations in the magnetization which are known as baby-Skyrmions. When we consider the magnetization field also as a function of Euclidean time, and if we compactify the corresponding 3-d space-time to the sphere S^3, the magnetization defines a map from S^3 to S^2. Such maps again fall into topological classes characterized by the so-called Hopf number $H[\vec{e}] \in \Pi_3[S^2] = \mathbb{Z}$. This has very interesting consequences in condensed matter physics, because it implies that in a 3-d space-time, particles need not necessarily be either bosons or fermions. In fact, particles can then be so-called anyons, with any fractional spin and any statistics characterized by a statistics parameter $\Theta \in]-\pi, \pi]$, where $\Theta = 0$ and π correspond to bosons and fermions, respectively. For anyons, a term $i\Theta H[\vec{e}]$ enters the Euclidean action. For example, the quasi-particle excitations in a fractional quantum Hall sample are anyons with fractional spin, statistics, and charge, which can be viewed as fractionalized electrons, e.g. with charge $-e/3$. However, we should not think of an individual electron dividing itself into three anyons. Fractionalization is a collective phenomenon that can happen only if a macroscopic number of electrons behaves coherently at extremely low temperatures.

Next, let us consider maps from S^3 to S^3, which is the group manifold of $SU(2)$. For example, gauge transformations $U(x) \in SU(2)$ are maps from S^3 to S^3, if

$x \in S^3$ is a point in a compactified 3-d space. In this case, the corresponding winding number is given by

$$n[U] = \frac{1}{24\pi^2} \int_{S^3} d^3x \; \varepsilon_{ijk} \mathrm{Tr}[(U\partial_i U^\dagger)(U\partial_j U^\dagger)(U\partial_k U^\dagger)] \in \Pi_3[SU(2)] = \mathbb{Z}. \quad (6.1.4)$$

Again, it is instructive to convince oneself that this expression always assumes integer values. As we will discuss below, the corresponding winding number gives rise to the so-called vacuum angle $\theta \in]-\pi, \pi]$ in non-Abelian gauge theories. The corresponding topologically non-trivial Euclidean gauge field configurations are known as instantons.

The pions in QCD are described by an effective field theory, whose field $U(x) \in SU(2)$ again maps a compactified space S^3 into the $SU(2)$ group manifold S^3. In this case, $n[U]$ counts topological Skyrmion excitations in the pion field, which represent baryons. When we consider a compactified Euclidean space-time S^4, the pion field defines a map from S^4 to S^3. Interestingly, $\Pi_4[S^3] = \mathbb{Z}(2)$, which implies that in a 4-d space-time particles are either bosons or fermions, and anyons are not possible in that case.

The homotopy group $\Pi_4[S^3] = \mathbb{Z}(2)$ also plays a role in the Standard Model of particle physics, where it characterizes the topology of $SU(2)_L$ gauge transformations governing the weak interactions. Since in 4-d there are topologically non-trivial $SU(2)_L$ gauge transformations, the Standard Model is affected by Witten's so-called "global" anomaly. An anomaly amounts to an explicit violation of gauge invariance by quantum effects, despite the fact that the classical Lagrangian is gauge invariant. Gauge anomalies render a theory mathematically inconsistent, and must thus be canceled. In the Standard Model, the global anomaly is canceled only if the number of left-handed fermion doublets is even. Since there is one left-handed lepton doublet (consisting of the left-handed neutrino and electron fields), there must be an odd number of left-handed quark doublets (consisting of the left-handed up and down quark fields), which implies that the number of quark colors N_c must be odd. Indeed, in Nature $N_c = 3$.

In general, maps from a sphere S^d to S^d are characterized by an integer winding number in the homotopy group $\Pi_d[S^d] = \mathbb{Z}$.

6.2 Homotopy Groups of Lie Group Manifolds

As we have discussed before, the group manifolds of compact Lie groups are products of spheres, at least locally. In particular, we have

$$SU(n) = S^3 \times S^5 \times \ldots \times S^{2n-1},$$

$$
\begin{aligned}
&Spin(n) = S^1 \times S^2 \times \ldots \times S^{n-1}, \\
&Sp(n) = S^3 \times S^7 \times \ldots \times S^{4n-1}, \\
&G(2) = SU(3) \times S^6 = S^3 \times S^5 \times S^6.
\end{aligned} \tag{6.2.1}
$$

It is interesting to note that all these group manifolds contain a factor S^3, which implies that

$$
\Pi_3[SU(n)] = \Pi_3[Spin(n)] = \Pi_3[Sp(n)] = \Pi_3[G(2)] = \mathbb{Z}. \tag{6.2.2}
$$

As a consequence, in 4-d all non-Abelian gauge theories have instantons and a corresponding vacuum angle $\theta \in]-\pi, \pi]$.

Interestingly, while $\Pi_4[SU(2)] = \mathbb{Z}(2)$, $\Pi_4[SU(n)] = \{0\}$ is topologically trivial for $n \geq 3$. This implies that in 4-d the corresponding $SU(n)$ gauge transformations can all be continuously deformed into each other. Hence, there is no global anomaly for $SU(n)$ with $n \geq 3$.

The low-energy effective theory of the Goldstone pions, kaons, and η-mesons in QCD is formulated in terms of matrix-valued fields $U(x) \in SU(3)$. Since $\Pi_3[SU(3)] = \mathbb{Z}$, again there are Skyrmions which represent baryons. This theory also has anomalies which are represented by the so-called Wess-Zumino-Novikov-Witten term

$$
S_{WZNW}[U] = \frac{1}{480\pi^3 i} \int_{H^5} d^5x \varepsilon_{\mu\nu\rho\sigma\lambda} \mathrm{Tr}[(U\partial_\mu U^\dagger)(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger)(U\partial_\lambda U^\dagger)]. \tag{6.2.3}
$$

Here H^5 is a 5-d hemi-sphere whose boundary $\partial H^5 = S^4$ is the compactified 4-d Euclidean space-time. Interestingly, the WZNW-term can be constructed only because $\Pi_4[SU(3)] = \{0\}$. It enters the Euclidean action as $2\pi i N_c S_{WZNW}[U]$. The prefactor, which turns out to be given by the number of colors N_c, must be quantized in integer units because $\Pi_5[U] = \mathbb{Z}$. In fact, the winding number

$$
\begin{aligned}
n[U] &= \frac{1}{480\pi^3 i} \int_{S^5} d^5x \varepsilon_{\mu\nu\rho\sigma\lambda} \mathrm{Tr}[(U\partial_\mu U^\dagger)(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger)(U\partial_\lambda U^\dagger)] \\
&\in \Pi_5[SU(3)] = \mathbb{Z},
\end{aligned} \tag{6.2.4}
$$

gives rise to an ambiguity of $S_{WZNW}[U]$, which cancels from the Euclidean path integral only when the prefactor is quantized in integer units.

6.3 The Topological Charge of a 2-d Abelian Gauge Field

Let us consider an Abelian gauge field $A_\mu(x)$ in 2-d Euclidean space-time. The corresponding action then takes the form

$$S[A] = \frac{1}{2e^2} \int d^2x \; F_{\mu\nu} F_{\mu\nu}, \tag{6.3.1}$$

where $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$ is the field strength, and e is the electric charge, which has the dimension of mass in 2-d. In two space-time dimensions $F_{\mu\nu}$ has only one non-trivial component $F_{12}(x) = E(x)$, which represents the electric field. The classical Euclidean equation of motion $\partial_\mu F_{\mu\nu}(x) = 0$ then implies $\partial_1 E(x) = \partial_2 E(x) = 0$, such that $E(x)$ is then a constant.

One can also define a dual object

$$q(x) = \frac{1}{4\pi} \varepsilon_{\mu\nu} F_{\mu\nu}(x) = \frac{1}{2\pi} E(x), \tag{6.3.2}$$

which is again given by the electric field, and plays the role of a topological charge density. The corresponding topological charge is given by

$$Q = \int d^2x \; q = \frac{1}{4\pi} \int d^2x \; \varepsilon_{\mu\nu} F_{\mu\nu}. \tag{6.3.3}$$

The topological charge density is a total divergence

$$q(x) = \frac{1}{4\pi} \varepsilon_{\mu\nu} F_{\mu\nu}(x) = \frac{1}{2\pi} \varepsilon_{\mu\nu} \partial_\mu A_\nu(x) = \partial_\mu \Omega_\mu^{(0)}(x), \tag{6.3.4}$$

where

$$\Omega_\mu^{(0)}(x) = \frac{1}{2\pi} \varepsilon_{\mu\nu} A_\nu(x) \tag{6.3.5}$$

is known as a 0-cochain. Using Gauss' theorem, we can the write

$$Q = \int d^2x \; q = \int d^2x \; \partial_\mu \Omega_\mu^{(0)} = \int_{S^1} d\sigma_\mu \; \Omega_\mu^{(0)} = \frac{1}{2\pi} \int_{S^1} d\sigma_\mu \; \varepsilon_{\mu\nu} A_\nu, \tag{6.3.6}$$

where $S^1 = \partial \mathbb{R}^2$ is a large circle at the boundary of the Euclidean space-time $\mathbb{R}^2$.

Let us now consider field configurations of finite action which implies that $F_{\mu\nu}(x)$ must vanish at space-time infinity. The field strength vanishes when the vector potential is gauge equivalent to

$$A'_\mu(x) = A_\mu(x) - \partial_\mu \varphi(x) = 0 \;\; \Rightarrow \;\; A_\mu(x) = \partial_\mu \varphi(x), \tag{6.3.7}$$

i.e. $A_\mu(x)$ is a pure gauge, at least at the boundary of space-time. Hence, the topological charge can be expressed as

$$Q = \int d^2x \; q = \int d^2x \; \partial_\mu \Omega_\mu^{(0)} = \int_{S^1} d\sigma_\mu \; \Omega_\mu^{(0)} = \frac{1}{2\pi} \int_{S^1} d\sigma_\mu \; \varepsilon_{\mu\nu} \partial_\nu \varphi \in \Pi_1[S^1] = \mathbb{Z}. \tag{6.3.8}$$

Here we have identified Q as the topological winding number $n[U]$ of the gauge transformation $U(x) = \exp(i\varphi(x))$ at space-time infinity.

6.4 The Topological Charge of a 4-d Non-Abelian Gauge Field

Let us now consider an non-Abelian gauge field $G_\mu(x) = igG_\mu^a(x)T^a$ in 4-d Euclidean space-time. The action then takes the form

$$S[G] = -\frac{1}{2g^2} \int d^2x \; \mathrm{Tr}(G_{\mu\nu} G_{\mu\nu}), \tag{6.4.1}$$

Here $G_{\mu\nu}(x) = \partial_\mu G_\nu(x) - \partial_\nu G_\mu(x) + [G_\mu(x), G_\nu(x)]$. We now define the so-called Chern-Pontryagin density

$$q(x) = -\frac{1}{32\pi^2} \varepsilon_{\mu\nu\rho\sigma} \mathrm{Tr}(G_{\mu\nu}(x) G_{\rho\sigma}(x)), \tag{6.4.2}$$

which is a total divergence, i.e.

$$q(x) = \partial_\mu \Omega_\mu^{(0)}(x). \tag{6.4.3}$$

Here $\Omega_\mu^{(0)}(x)$ is the so-called Chern-Simons density or 0-cochain, which is given by

$$\Omega_\mu^{(0)}(x) = -\frac{1}{8\pi^2} \varepsilon_{\mu\nu\rho\sigma} \mathrm{Tr}[G_\nu(x)(\partial_\rho G_\sigma(x) + \frac{2}{3} G_\rho(x) G_\sigma(x))]. \tag{6.4.4}$$

It is a good exercise to convince oneself that this satisfies eq.(6.4.3). The topological charge is defined as

$$Q = -\frac{1}{32\pi^2} \int d^4x \; \varepsilon_{\mu\nu\rho\sigma} \mathrm{Tr}(G_{\mu\nu} G_{\rho\sigma}) = \int d^4x \; q = \int d^4x \; \partial_\mu \Omega_\mu^{(0)} = \int_{S^3} d^3\sigma_\mu \; \Omega_\mu^{(0)}. \tag{6.4.5}$$

We have used Gauss' law to reduce the integral over Euclidean space-time to an integral over its boundary, which is topologically a 3-sphere S^3. We will restrict ourselves to gauge field configurations with a finite action. Hence, their field strength

should vanish at infinity, and consequently the gauge potential should then be a pure gauge (a gauge transformation of a zero field)

$$G_\mu(x) = U(x)\partial_\mu U(x)^\dagger. \tag{6.4.6}$$

Of course, this expression is only valid at space-time infinity. Inserting it in the expression for the 0-cochain we obtain

$$\begin{aligned} Q &= -\frac{1}{8\pi^2}\int_{S^3} d^3\sigma_\mu\ \varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(U\partial_\nu U^\dagger)(\partial_\rho(U\partial_\sigma U^\dagger) + \frac{2}{3}(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger))] \\ &= -\frac{1}{8\pi^2}\int_{S^3} d^3\sigma_\mu\ \varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[-(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger) \\ &+ \frac{2}{3}(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger)] \\ &= \frac{1}{24\pi^2}\int_{S^3} d^3\sigma_\mu\ \varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger)]. \end{aligned} \tag{6.4.7}$$

The gauge transformation $U(x)$ defines a map of the sphere S^3 at space-time infinity to the gauge group $SU(n)$

$$U : S^3 \to SU(n). \tag{6.4.8}$$

Such maps have topological properties. They fall into the homotopy class

$$\Pi_3[SU(n)] = \mathbb{Z}. \tag{6.4.9}$$

The third homotopy group of $SU(n)$ is given by the integers. This means that for each integer Q there is a class of maps that can be continuously deformed into one another, while maps with different Q are topologically distinct. The integer Q that characterizes the map topologically is the topological charge. Now we want to show that the above expression for Q is exactly that integer. For this purpose we decompose

$$U = VW,\ W = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \tilde{U}_{11} & \tilde{U}_{12} & \dots & \tilde{U}_{1,n-1} \\ 0 & \tilde{U}_{21} & \tilde{U}_{22} & \dots & \tilde{U}_{1,n-1} \\ . & . & . & & . \\ . & . & . & & . \\ 0 & \tilde{U}_{n-1,1} & \tilde{U}_{n-1,2} & \dots & \tilde{U}_{n-1,n-1} \end{pmatrix}, \tag{6.4.10}$$

where the embedded matrix $\tilde{U}$ is in $SU(n-1)$. It is indirectly defined by

$$V = \begin{pmatrix} U_{11} & -U_{21}^* & -\frac{U_{31}^*(1+U_{11})}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*(1+U_{11})}{1+U_{11}^*} \\ U_{21} & \frac{1+U_{11}^*-|U_{21}|^2}{1+U_{11}} & -\frac{U_{31}^*U_{21}}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*U_{21}}{1+U_{11}^*} \\ U_{31} & -\frac{U_{21}^*U_{31}}{1+U_{11}} & \frac{1+U_{11}^*-|U_{31}|^2}{1+U_{11}^*} & \cdots & -\frac{U_{n1}^*U_{31}}{1+U_{11}^*} \\ . & . & . & & . \\ . & . & . & & . \\ U_{n1} & -\frac{U_{21}^*U_{n1}}{1+U_{11}} & -\frac{U_{31}^*U_{n1}}{1+U_{11}^*} & \cdots & \frac{1+U_{11}^*-|U_{n1}|^2}{1+U_{11}^*} \end{pmatrix} \in SU(n). \qquad (6.4.11)$$

The matrix V is constructed entirely from the elements $U_{11}, U_{21}, \ldots, U_{n1}$ of the first column of the matrix U. One should convince oneself that V is indeed an $SU(n)$ matrix, and that the resulting matrix $\tilde{U}$ is indeed in $SU(n-1)$. The idea now is to reduce the expression for the topological charge from $SU(n)$ to $SU(n-1)$ by using the formula

$$\begin{aligned} &\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(VW)\partial_\nu(VW)^\dagger(VW)\partial_\rho(VW)^\dagger(VW)\partial_\sigma(VW)^\dagger] = \\ &\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(V\partial_\nu V^\dagger)(V\partial_\rho V^\dagger)(V\partial_\sigma V^\dagger) \\ &+\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(W\partial_\nu W^\dagger)(W\partial_\rho W^\dagger)(W\partial_\sigma W^\dagger)] \\ &+3\partial_\nu\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[((V\partial_\rho V^\dagger)(W\partial_\sigma W^\dagger)]. \end{aligned} \qquad (6.4.12)$$

Again, it is instructive to prove this formula. Applying the formula to the expression for the topological charge and using $U = VW$ we obtain

$$\begin{aligned} Q &= \frac{1}{24\pi^2}\int_{S^3} d^3\sigma_\mu\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(U\partial_\nu U^\dagger)(U\partial_\rho U^\dagger)(U\partial_\sigma U^\dagger)] \\ &= \frac{1}{24\pi^2}\int_{S^3} d^3\sigma_\mu\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(V\partial_\nu V^\dagger)(V\partial_\rho V^\dagger)(V\partial_\sigma V^\dagger) \\ &+ (W\partial_\nu W^\dagger)(W\partial_\rho W^\dagger)(W\partial_\sigma W^\dagger)]. \end{aligned} \qquad (6.4.13)$$

The ∂_ν term of the formula eq.(6.4.12) drops out using Gauss' law together with the fact that S^3 has no boundary. It follows that the topological charge of a product of two gauge transformations V and W is the sum of the topological charges of V and W. Since V only depends on $U_{11}, U_{21}, \ldots, U_{n1}$, it can be viewed as a map of S^3 into the sphere S^{2n-1}

$$V : S^3 \to S^{2n-1}. \qquad (6.4.14)$$

This is because $|U_{11}|^2 + |U_{21}|^2 + \ldots + |U_{n1}|^2 = 1$. Remarkably, the corresponding homotopy group is trivial for $n > 2$, i.e.

$$\Pi_3[S^{2n-1}] = \{0\}. \qquad (6.4.15)$$

All maps of S^3 into the higher dimensional sphere S^{2n-1} are topologically equivalent (they can be deformed into each other).

Since the map V is topologically trivial, its contribution to the topological charge vanishes. The remaining W term reduces to the $SU(n-1)$ contribution

$$Q = \frac{1}{24\pi^2}\int_{S^3} d^3\sigma_\mu\ \varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(\tilde{U}\partial_\nu\tilde{U}^\dagger)(\tilde{U}\partial_\rho\tilde{U}^\dagger)(\tilde{U}\partial_\sigma\tilde{U}^\dagger)]. \tag{6.4.16}$$

The separation of the V contribution works only if the decomposition of U into V and $\tilde{U}$ is non-singular. In fact, the expression for V is singular for $U_{11} = -1$. This corresponds to an $((n-1)^2-1)$-dimensional subspace of the (n^2-1)-dimensional $SU(n)$ group space. The map U itself covers a 3-d subspace of $SU(n)$. Hence it is arbitrarily improbable to hit a singularity (it is of measure zero). Since we have now reduced the $SU(n)$ topological charge to the $SU(n-1)$ case, we can go down all the way to $SU(2)$. It remains to be shown that the $SU(2)$ expression is actually an integer. First of all

$$\tilde{U} : S^3 \to SU(2) = S^3, \tag{6.4.17}$$

and indeed

$$\Pi_3[SU(2)] = \Pi_3[S^3] = \mathbb{Z}. \tag{6.4.18}$$

The topological charge specifies how often the $SU(2)$ group space (which is isomorphic to the 3-sphere) is covered by $\tilde{U}$ as we go along the boundary of Euclidean space-time (which is also topologically S^3). Let us parametrize the map $\tilde{U}$ as

$$\begin{aligned}\tilde{U}(x) &= \exp(i\vec{\alpha}(x)\cdot\vec{\sigma}) = \cos\alpha(x) + i\sin\alpha(x)\vec{e}_\alpha(x)\cdot\vec{\sigma},\\ \vec{e}_\alpha(x) &= (\sin\theta(x)\sin\varphi(x), \sin\theta(x)\cos\varphi(x), \cos\theta(x)).\end{aligned} \tag{6.4.19}$$

It is a good exercise to convince oneself that

$$\begin{aligned}&\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[(\tilde{U}(x)\partial_\nu\tilde{U}(x)^\dagger)(\tilde{U}(x)\partial_\rho\tilde{U}(x)^\dagger)(\tilde{U}(x)\partial_\sigma\tilde{U}(x)^\dagger)]\\ &= 12\sin^2\alpha(x)\sin\theta(x)\varepsilon_{\mu\nu\rho\sigma}\partial_\nu\alpha(x)\partial_\rho\theta(x)\partial_\sigma\varphi(x).\end{aligned} \tag{6.4.20}$$

This is exactly the volume element of a 3-sphere (and hence of the $SU(2)$ group space). Thus we now write

$$Q = \frac{1}{2\pi^2}\int_{S^3} d^3\sigma_\mu\ \sin^2\alpha\sin\theta\varepsilon_{\mu\nu\rho\sigma}\partial_\nu\alpha\partial_\rho\theta\partial_\sigma\varphi = \frac{1}{2\pi^2}\int_{S^3} d\tilde{U}. \tag{6.4.21}$$

The volume of the 3-sphere is given by $2\pi^2$. When the map $\tilde{U}$ covers the sphere Q times, the integral gives Q times the volume of S^3. This finally explains why the prefactor $1/32\pi^2$ was introduced in the original expression for the topological charge.

6.5 The Instanton in $SU(2)$

We have argued mathematically that gauge field configurations fall into topologically distinct classes. Now we want to construct concrete examples of topologically nontrivial field configurations. Here we consider instantons, which have $Q = 1$ and are solutions of the Euclidean classical field equations. The instanton occurs at a given instant in Euclidean time. Since these solutions do not exist in Minkowski space-time they have no direct interpretation in terms of real-time events. Also it is unclear which role they play in the quantum theory. Instantons describe tunneling processes between degenerate classical vacuum states. Their existence gives rise to the θ-vacuum structure of non-Abelian gauge theories.

Here we concentrate on $SU(2)$. This is sufficient, because we have seen that the $SU(n)$ topological charge can be reduced to the $SU(2)$ case. In this section we go back to an infinite space with a boundary sphere S^3, and we demand that the gauge field has finite action. Then at space-time infinity the gauge potential is in a pure gauge

$$G_\mu(x) = U(x)\partial_\mu U(x)^\dagger. \tag{6.5.1}$$

Provided the gauge field is otherwise smooth, the topology resides entirely in the map U. We want to construct a field configuration with topological charge $Q = 1$, i.e. one in which the map U covers the group space $SU(2) = S^3$ once, as we integrate over the boundary sphere S^3. The simplest map of this kind is the identity, i.e. each point at the boundary of space-time is mapped to the corresponding point in group space, such that

$$U(x) = \frac{x_0 + i\vec{x}\cdot\vec{\sigma}}{|x|},\ |x| = \sqrt{x_0^2 + |\vec{x}|^2}. \tag{6.5.2}$$

Next we want to extend the gauge field to the interior of space-time without introducing singularities. We cannot simply maintain the form of eq.(6.5.1) because U is singular at $x = 0$. To avoid this singularity we make the ansatz

$$G_\mu(x) = f(|x|)U(x)\partial_\mu U(x)^\dagger, \tag{6.5.3}$$

where $f(\infty) = 1$ and $f(0) = 0$. For any smooth function f with these properties the above gluon field configuration has $Q = 1$. Still, this does not mean that we have constructed an instanton. Instantons are field configurations with $Q \neq 0$ that are in addition solutions of the Euclidean classical equations of motion, i.e. they are minima of the Euclidean action

$$S[G] = -\frac{1}{2g^2}\int d^4x\ \mathrm{Tr}[G_{\mu\nu}G_{\mu\nu}]. \tag{6.5.4}$$

Let us consider the following integral

$$-\int d^4x\ \mathrm{Tr}[(G_{\mu\nu} \pm \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G_{\rho\sigma})(G_{\mu\nu} \pm \frac{1}{2}\varepsilon_{\mu\nu\kappa\lambda}G_{\kappa\lambda}) =$$

$$
\begin{aligned}
&- \int d^4x \ \mathrm{Tr}[G_{\mu\nu}G_{\mu\nu} \pm \varepsilon_{\mu\nu\rho\sigma}G_{\mu\nu}G_{\rho\sigma} + G_{\mu\nu}G_{\mu\nu}] \\
&= 4g^2 S[G] \pm 32\pi^2 Q[G]. \qquad (6.5.5)
\end{aligned}
$$

We have integrated a square. Hence it is obvious that

$$
S[G] \pm \frac{8\pi^2}{g^2} Q[G] \geq 0 \ \Rightarrow \ S[G] \geq \frac{8\pi^2}{g^2}|Q[G]|, \qquad (6.5.6)
$$

i.e. a topologically nontrivial field configuration costs at least a minimum action proportional to the topological charge. Instantons are configurations with minimum action, i.e. for them

$$
S[G] = \frac{8\pi^2}{g^2}|Q[G]|. \qquad (6.5.7)
$$

From the above argument it is clear that a minimum action configuration arises only if

$$
G_{\mu\nu}(x) = \pm\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G_{\rho\sigma}(x). \qquad (6.5.8)
$$

Configurations that obey this equation with a plus sign are called selfdual. The ones that obey it with a minus sign are called anti-selfdual. It is instructive to convince oneself that the above gluon field with

$$
f(|x|) = \frac{|x|^2}{|x|^2 + \rho^2} \qquad (6.5.9)
$$

is indeed an instanton for any value of ρ. The instanton configuration hence takes the form

$$
G_\mu(x) = \frac{|x|^2}{|x|^2 + \rho^2} U(x)\partial_\mu U(x)^\dagger. \qquad (6.5.10)
$$

There is a whole family of instantons with different radii ρ. As a consequence of scale invariance of the classical action they all have the same action $S[G] = 8\pi^2/g^2$.

Chapter 7

The Adiabatic Berry Phase

In this chapter we will consider quantum mechanical systems with a Hamiltonian that depends on some slowly varying external parameters, such that the system undergoes an adiabatic time evolution. In 1928 Born and Fock derived the adiabatic theorem in quantum mechanics. According to the theorem, a quantum system then evolves from an eigenstate of the initial Hamiltonian through the momentary eigenstates of the time-dependent Hamiltonian. When the Hamiltonian undergoes a periodic time evolution, the system ultimately returns to the initial eigenstate, at least up to a complex $U(1)$ phase, known as the Berry phase which was noticed by Michael Berry in 1983. When the eigenstate is n-fold degenerate, the Berry phase becomes a non-Abelian $U(n)$ matrix. There is an abstract Abelian or non-Abelian Berry gauge field in the space of slowly varying external parameters, whose parallel transport along a closed path in parameter space yields the Berry phase. Non-Abelian gauge fields even arise in the classical physics of falling cats.

7.1 Abelian Berry Phase of a Spin $\frac{1}{2}$ in a Magnetic Field

Let us consider a spin $\frac{1}{2}$ in a time-dependent magnetic field $\vec{B}(t)$. The time-dependent Hamiltonian then takes the form

$$H(t) = \mu \vec{B}(t) \cdot \vec{\sigma}, \tag{7.1.1}$$

where $\vec{\sigma}$ denotes the Pauli matrices, and μ is a magnetic moment.

The states of a spin $\frac{1}{2}$ can be parameterized as

$$|\vec{e}\rangle = a|\uparrow\rangle + b|\downarrow\rangle, \quad |a|^2 + |b|^2 = 1, \tag{7.1.2}$$

which gives rise to the projection operator

$$\begin{aligned} P(\vec{e}) &= |\vec{e}\rangle\langle\vec{e}| = (a|\uparrow\rangle + b|\downarrow\rangle)(a^*\langle\uparrow| + b^*\langle\downarrow|) = \begin{pmatrix} |a|^2 & a^*b \\ b^*a & |b|^2 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+e_3 & e_1 - ie_2 \\ e_1 + ie_2 & 1-e_3 \end{pmatrix} = \frac{1}{2}(\mathbf{1} + \vec{e}\cdot\vec{\sigma}), \end{aligned} \tag{7.1.3}$$

where we have identified

$$\vec{e} = \langle\vec{e}|\vec{\sigma}|\vec{e}\rangle = (a^*b + b^*a, -ia^*b + ib^*a, |a|^2 - |b|^2). \tag{7.1.4}$$

The vector $\vec{e} \in S^2$ associates a spin state with a point on the so-called Bloch sphere. Thereby we do not distinguish the state $|\vec{e}\rangle$ from states $\exp(i\alpha)|\vec{e}\rangle$ which differ only by an irrelevant phase, which cancels in the physical projection operator $P(\vec{e}) = |\vec{e}\rangle\langle\vec{e}|$.

Identifying $\vec{e}(t) = \pm\vec{B}(t)/|\vec{B}(t)|$ we obtain

$$\begin{aligned} H(t)|\vec{e}(t)\rangle\langle\vec{e}(t)| &= H(t)P(t) = \mu\vec{B}(t)\cdot\vec{\sigma}\frac{1}{2}(\mathbf{1} + \vec{e}\cdot\vec{\sigma}) \\ &= \frac{\mu}{2}\vec{B}(t)\cdot\vec{\sigma} + \frac{\mu}{2}\vec{B}(t)\cdot\vec{e}(t) = \frac{\mu}{2}\vec{B}(t)\cdot\vec{\sigma} \pm \frac{\mu}{2}|\vec{B}(t)| \\ &= \pm\mu|\vec{B}(t)||\vec{e}(t)\rangle\langle\vec{e}(t)|. \end{aligned} \tag{7.1.5}$$

Hence, the state $|\vec{e}(t)\rangle = |\pm\vec{B}(t)/|\vec{B}(t)|\rangle$ is a momentary eigenstate of the Hamiltonian with eigenvalue $\pm\mu|\vec{B}(t)|$.

Let us now consider the time-dependent Schrödinger equation

$$i\partial_t|\Psi(t)\rangle = H(t)|\Psi(t)\rangle. \tag{7.1.6}$$

For a slowly varying external magnetic field, the adiabatic theorem then suggests the ansatz

$$|\Psi(t)\rangle = \exp\left(\mp i\int_0^t dt'\ \mu|\vec{B}(t')|\right)\exp(i\gamma_\pm)|\vec{e}(t)\rangle, \tag{7.1.7}$$

i.e. the system always remains in a momentary eigenstate $|\vec{e}(t)\rangle$, but it also accumulates a phase. Inserting this ansatz in the time-dependent Schrödinger equation, we obtain

$$\begin{aligned} i\partial_t|\Psi(t)\rangle &= \left[\pm\mu|\vec{B}(t)| - \partial_t\gamma_\pm + i\partial_t\right]\exp\left(\mp i\int_0^t dt'\ \mu|\vec{B}(t')|\right)\exp(i\gamma_\pm)|\vec{e}(t)\rangle, \\ H(t)|\Psi(t)\rangle &= \pm\mu|\vec{B}(t)|\exp\left(\mp i\int_0^t dt'\ \mu|\vec{B}(t')|\right)\exp(i\gamma_\pm(t))|\vec{e}(t)\rangle, \end{aligned} \tag{7.1.8}$$

such that

$$|\vec{e}(t)\rangle\partial_t\gamma_\pm(t) = i\partial_t|\vec{e}(t)\rangle. \tag{7.1.9}$$

This equation can be satisfied only when the time-evolution is indeed adiabatic. In that case, one obtains

$$\partial_t \gamma_\pm(t) = i\langle \vec{e}(t)|\partial_t|\vec{e}(t)\rangle. \tag{7.1.10}$$

The Berry phase is defined for a cyclic variation of the external magnetic field with the period T, for which $\vec{B}(t+T) = \vec{B}(t)$,

$$\gamma_\pm(T) = \int_0^T dt\ \partial_t \gamma_\pm(t) = i\int_0^T dt\ \langle \vec{e}(t)|\partial_t|\vec{e}(t)\rangle. \tag{7.1.11}$$

Since $|\vec{e}(t)\rangle$ depends on t only through $\vec{B}(t)$, we can write

$$|\vec{e}(t)\rangle = |\vec{e}(\vec{B}(t))\rangle \ \Rightarrow\ \partial_t|\vec{e}(t)\rangle = \partial_t\vec{B}(t)\cdot\vec{\nabla}_B|\vec{e}(\vec{B}(t))\rangle, \tag{7.1.12}$$

such that

$$\gamma_\pm(T) = i\int_0^T dt\ \partial_t\vec{B}(t)\cdot\langle\vec{e}(\vec{B}(t))|\vec{\nabla}_B|\vec{e}(\vec{B}(t))\rangle = i\int_{\mathcal{C}} d\vec{B}\cdot\langle\vec{e}(\vec{B})|\vec{\nabla}_B|\vec{e}(\vec{B})\rangle. \tag{7.1.13}$$

This shows that the Berry phase depends only on the curve $\mathcal{C}$ along which $\vec{B}(t)$ varies with time, but not on the velocity of the variation, at least as long as it remains adiabatic. Consequently, the Berry phase is a purely geometric and not a dynamical object. This suggests to introduce an abstract vector potential, also known as the Berry connection,

$$\vec{A}(\vec{B}) = i\langle\vec{e}(\vec{B})|\vec{\nabla}_B|\vec{e}(\vec{B})\rangle, \tag{7.1.14}$$

which is real-valued despite the factor i. The Berry phase is then identified as a Wilson loop of the abstract Abelian Berry gauge field. It should be pointed out that this gauge field does not exist in coordinate space, but rather in the space of external parameters of a quantum system, in this case, in the space of all possible external magnetic fields $\vec{B}$.

Let us now make a different choice for the arbitrary phase of a momentary eigenstate

$$|\vec{e}(\vec{B})\rangle' = \exp(i\alpha(\vec{B}))|\vec{e}(\vec{B})\rangle. \tag{7.1.15}$$

This implies an Abelian gauge transformation of the vector potential

$$\vec{A}(\vec{B})' = \vec{A}(\vec{B}) - \vec{\nabla}_B\alpha(\vec{B}). \tag{7.1.16}$$

The Berry phase is gauge invariant, i.e.

$$\gamma_\pm(T)' = \int_{\mathcal{C}} d\vec{B}\cdot\vec{A}(\vec{B})' = \int_{\mathcal{C}} d\vec{B}\cdot[\vec{A}(\vec{B}) - \vec{\nabla}_B\alpha(\vec{B})] = \gamma_\pm(T). \tag{7.1.17}$$

Let us now associate an abstract Berry field strength with the Abelian vector potential

$$\vec{F}(\vec{B}) = \vec{\nabla}_B \times \vec{A}(\vec{B}), \tag{7.1.18}$$

which obviously is gauge invariant. Using Stoke's theorem, the Berry phase can then be expressed as

$$\gamma_{\pm}(T) = \int_{\mathcal{C}} d\vec{B} \cdot \vec{A}(\vec{B}) = \int_S d\vec{s} \cdot \vec{\nabla}_B \times \vec{A}(\vec{B}) = \int_S d\vec{s} \cdot \vec{F}(\vec{B}), \tag{7.1.19}$$

i.e. it represents the flux of the Berry field strength through a surface S bounded by the closed curve $\mathcal{C}$.

Let us evaluate the Berry gauge field for the spin $\frac{1}{2}$ in an external magnetic field

$$\vec{B} = |\vec{B}|(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \tag{7.1.20}$$

For $\vec{B} = (0, 0, |\vec{B}|)$ the ground state is given by $|\vec{e}(\vec{B})\rangle = |\uparrow\rangle$. The ground state for a general orientation of $\vec{B}$ is obtained by a rotation

$$|\vec{e}(\vec{B})\rangle = \exp(i\varphi S_z)\exp(-i\theta S_x)|\uparrow\rangle. \tag{7.1.21}$$

Based on this, it is straightforward to work out

$$\langle\vec{e}(\vec{B})|\partial_\varphi|\vec{e}(\vec{B})\rangle = \frac{i}{2}\cos\theta, \quad \langle\vec{e}(\vec{B})|\partial_\theta|\vec{e}(\vec{B})\rangle = \langle\vec{e}(\vec{B})|\partial_{|\vec{B}|}|\vec{e}(\vec{B})\rangle = 0, \tag{7.1.22}$$

such that one obtains

$$\vec{A}(\vec{B}) = -\frac{\cos\theta}{2|\vec{B}|\sin\theta}\vec{e}_\varphi. \tag{7.1.23}$$

The corresponding field strength is the one of a "magnetic monopole" in parameter space

$$\vec{F}(\vec{B}) = \vec{\nabla}_B \times \vec{A}(\vec{B}) = -\frac{1}{|\vec{B}|\sin\theta}\partial_\theta\left(\frac{\cos\theta}{2|\vec{B}|}\right)\vec{e}_B = \frac{1}{2}\frac{\vec{B}}{|\vec{B}|^3}. \tag{7.1.24}$$

7.2 Non-Abelian Berry Phase

Non-Abelian Berry phases arise when one considers the adiabatic evolution of a set of n degenerate states, which remain degenerate while some external parameters are varied. After a slow periodic variation of the external parameters, the initial state may then not turn back to itself, but may turn into another member of the set of degenerate states. The Berry phase, which rotates the initial into the final state, then takes the form of a non-Abelian $U(n)$ matrix.

As a concrete example, we consider a nuclear spin resonance experiment in which a probe rotates in a magnetic field. In this case, the interaction of the spin $\vec{S}$ proceeds via the nuclear quadrupole moment and is given by

$$H = \omega(\vec{B}(t) \cdot \vec{S})^2. \tag{7.2.1}$$

When $\vec{B} = (0, 0, |\vec{B}|)$, the energy eigenstates are $|m\rangle$ with $m = -S, -S+1, ..., S$ with eigenvalues $E_m = \omega|\vec{B}(t)|^2 m^2$. In particular, the states $|m\rangle$ and $|-m\rangle$ are degenerate. Again, we obtain the momentary eigenstates by a rotation

$$|m(\vec{B})\rangle = \exp(i\varphi S_z)\exp(-i\theta S_x)|m(\vec{B})\rangle. \tag{7.2.2}$$

In this case, the Berry gauge field is non-Abelian

$$\begin{aligned}
\vec{A}_{++}(\vec{B}) &= i\langle m(\vec{B})|\vec{\nabla}_B|m(\vec{B})\rangle = -\frac{m}{2}\frac{\cos\theta}{|\vec{B}|\sin\theta}\vec{e}_\varphi, \\
\vec{A}_{--}(\vec{B}) &= i\langle -m(\vec{B})|\vec{\nabla}_B|-m(\vec{B})\rangle = \frac{m}{2}\frac{\cos\theta}{|\vec{B}|\sin\theta}\vec{e}_\varphi, \\
\vec{A}_{+-}(\vec{B}) &= i\langle m(\vec{B})|\vec{\nabla}_B|-m(\vec{B})\rangle,\ \vec{A}_{-+}(\vec{B}) = i\langle -m(\vec{B})|\vec{\nabla}_B|m(\vec{B})\rangle.
\end{aligned} \tag{7.2.3}$$

A straightforward calculation reveals that $\vec{A}_{+-}(\vec{B})$ and $\vec{A}_{-+}(\vec{B})$ vanish, unless $m = \pm\frac{1}{2}$. In that case, one obtains

$$\vec{A}_{\pm\mp}(\vec{B}) = \frac{1}{2|\vec{B}|}\sqrt{S(S+1)+\frac{1}{4}}\left(\pm i\vec{e}_\varphi + \vec{e}_\theta\right), \tag{7.2.4}$$

such that the non-Abelian $SU(2)$ Berry vector potential takes the form

$$\vec{A}(\vec{B}) = \frac{1}{2|\vec{B}|}\left(\sqrt{S(S+1)+\frac{1}{4}}\vec{e}_\theta\sigma_1 + \sqrt{S(S+1)+\frac{1}{4}}\vec{e}_\varphi\sigma_2 - \frac{\cos\theta}{\sin\theta}\vec{e}_\varphi\sigma_3\right). \tag{7.2.5}$$

Under a unitary change $U \in SU(2)$ of the two basis states

$$|m(\vec{B})\rangle' = |n(\vec{B})\rangle U(\vec{B})^\dagger_{nm}, \tag{7.2.6}$$

the Berry gauge field transforms as one would expect for a non-Abelian gauge field

$$\begin{aligned}
\vec{A}(\vec{B})' &= i'\langle n(\vec{B})|\vec{\nabla}_B|m(\vec{B})\rangle' = iU(\vec{B})\langle n(\vec{B})|\vec{\nabla}_B|m(\vec{B})\rangle U(\vec{B})^\dagger \\
&= U(\vec{B})\vec{A}(\vec{B})U(\vec{B})^\dagger + iU(\vec{B})\vec{\nabla}_B U(\vec{B})^\dagger.
\end{aligned} \tag{7.2.7}$$

The corresponding non-Abelian field strength is then given by

$$\begin{aligned}
\vec{F}(\vec{B}) &= \vec{\nabla}_B \times \vec{A}(\vec{B}) + i\vec{A}(\vec{B}) \times \vec{A}(\vec{B}) \\
&= \frac{\vec{B}}{2|\vec{B}|^3}\sigma_3 + \frac{1}{2|\vec{B}|^2}\left[\frac{\cos\theta}{\sin\theta}\sqrt{S(S+1)+\frac{1}{4}}\vec{e}_z\sigma_2 - \left(S(S+1)+\frac{1}{4}\right)\vec{e}_z\sigma_3\right].
\end{aligned} \tag{7.2.8}$$

7.3 $SO(3)$ Gauge Fields in Falling Cats

Cats have the ability to land on their feet, even if they are dropped head down from some height. They twist their body and use their tail, such that their body undergoes a net rotation. Let us try to understand this phenomenon in mathematical terms. Somewhat surprisingly, we will encounter a non-Abelian gauge field in the space of all shapes of the cat's body, whose non-Abelian Berry phase gives the net rotation angle of the cat.

Let us discretize the cat by a set of point masses m_i at positions $\vec{x}_i$, and let us imagine that the cat has control over the shape of its body, by influencing the relative orientation of the point masses. While the cat is in free fall, it does not feel gravity, at least until it hits the ground. During the fall, we can simply go to the accelerated center of mass frame of the cat, and then describe the time-dependent shape of the cat in the absence of gravity. The key to the understanding of this problem is angular momentum conservation. The angular momentum of the cat is simply given by

$$\vec{L} = \sum_i m_i \vec{x}_i \times \frac{d\vec{x}_i}{dt}. \tag{7.3.1}$$

When the cat is originally released at rest, the angular momentum vanishes and will remain $\vec{L} = 0$ until the cat hits the ground, as a consequence of angular momentum conservation.

Let us now define a possible shape of the cat as a particular configuration of the points $\vec{x}_i$, with configurations being identified if they differ just by an $SO(3)$ spatial rotation. Any possible shape can be characterized by a reference configuration $\vec{y}_i$, which can then be realized in all possible orientations

$$\vec{x}_i = O\vec{y}_i, \quad O^T O = OO^T = \mathbf{1}, \quad \det O = 1. \tag{7.3.2}$$

Here $O \in SO(3)$ is an orthogonal rotation matrix that rotates the reference configuration $\vec{y}_i$ into the general orientation $\vec{x}_i$, keeping the shape fixed. Let us now assume that, by controlling its body, the cat can send the point masses inside its body through any time-dependent sequence of shapes, defined by time-dependent reference configurations $\vec{y}_i(t)$ that the cat can choose at will. The question then is how the reference configuration is rotated by a time-dependent orthogonal rotation $O(t)$ into the actual position of the cat $\vec{x}_i(t)$. This simply follows from angular momentum conservation

$$\begin{aligned} \vec{L} &= \sum_i m_i \vec{x}_i(t) \times \frac{d\vec{x}_i(t)}{dt} \\ &= \sum_i m_i O(t)\vec{y}_i(t) \times \frac{d}{dt}\left[O(t)\vec{y}_i(t)\right] \end{aligned}$$

$$= \sum_i m_i O(t)\vec{y}_i(t) \times \left[\frac{dO(t)}{dt}\vec{y}_i(t) + O(t)\frac{d\vec{y}_i(t)}{dt}\right] = 0, \tag{7.3.3}$$

which thus implies

$$O(t)\sum_i m_i\vec{y}_i(t) \times O(t)^T\frac{dO(t)}{dt}\vec{y}_i(t) = -O(t)\sum_i m_i\vec{y}_i(t) \times \frac{d\vec{y}_i(t)}{dt} \;\Rightarrow$$
$$\sum_i m_i\vec{y}_i(t) \times O(t)^T\frac{dO(t)}{dt}\vec{y}_i(t) = -\sum_i m_i\vec{y}_i(t) \times \frac{d\vec{y}_i(t)}{dt} = -\vec{M}(t), \tag{7.3.4}$$

where $\vec{M}(t)$ is the non-conserved "angular momentum" of the reference configuration. We now introduce the anti-Hermitean non-Abelian $SO(3)$ vector potential

$$A(t) = O(t)^T\frac{dO(t)}{dt} = iA^a(t)T^a, \quad T^a_{bc} = -i\varepsilon_{abc}. \tag{7.3.5}$$

Expressed in components, eq.(7.3.4) takes the form

$$\sum_i m_i\varepsilon_{abc}y^b_i(t)A^d(t)\varepsilon_{dce}y^e_i(t) = -M^a(t) \;\Rightarrow$$
$$\sum_i m_i[y^e_i(t)y^e_i(t)\delta_{ab} - y^a_i(t)y^b_i(t)]A^b(t) = M^a(t). \tag{7.3.6}$$

Introducing the moment of inertia tensor

$$I_{ab}(t) = \sum_i m_i[y^e_i(t)y^e_i(t)\delta_{ab} - y^a_i(t)y^b_i(t)], \tag{7.3.7}$$

we finally obtain

$$I_{ab}(t)A^b(t) = M^a(t) \;\Rightarrow\; \vec{A}(t) = I(t)^{-1}\vec{M}(t). \tag{7.3.8}$$

Of course, the reference configuration $\vec{y}_i$ for a given shape of the cat can be chosen arbitrarily. Let us investigate how the vector potential $A(t)$ changes when the reference configuration is rotated to a new configuration

$$\vec{y}_i(t)' = \Omega(t)\vec{y}_i(t), \quad \Omega \in SO(3). \tag{7.3.9}$$

After such a rotation, the new transformation $O'(t)$ that rotates $\vec{y}_i(t)'$ into $\vec{x}_i(t)$, is given by

$$\vec{x}_i(t) = O(t)'\vec{y}_i(t)' = O(t)'\Omega(t)\vec{y}_i(t) \;\rightarrow\; O(t)' = O(t)\Omega(t)^T. \tag{7.3.10}$$

The corresponding vector potential then takes the form

$$\begin{aligned} A(t)' &= O(t)'^T \frac{dO(t)'}{dt} = \Omega(t) O(t)^T \left(\frac{dO(t)}{dt} \Omega(t)^T + O(t) \frac{d\Omega(t)^T}{dt} \right) \\ &= \Omega(t) \left(A(t) + \frac{d}{dt} \right) \Omega(t)^T. \end{aligned} \tag{7.3.11}$$

Hence, the change of reference configuration for a given shape amounts to an $SO(3)$ gauge transformation. This is not surprising because the different shapes play the role of gauge equivalence classes.

Finally, let us calculate the total rotation of the cat during a sequence of shape changes after which the cat returns to its initial shape. Using

$$O(t) A(t) = \frac{dO(t)}{dt}, \tag{7.3.12}$$

we obtain

$$O(T) = \mathcal{P} \exp \left(\int_0^T dt \; A(t) \right), \tag{7.3.13}$$

where $\mathcal{P}$ denotes path ordering. Hence, when the cat returns from its initial shape to the same final shape after a time T, its net rotation can be computed as a closed Wilson loop in an $SO(3)$ gauge field, very much like a non-Abelian Berry phase in quantum mechanics. Of course, all this does not explain why the cat is actually able to perform the difficult task of landing on her feet. While it is unlikely that it has an $SO(3)$ Wilson loop computer hard-wired in its brain, at least our brain is capable of describing the cat's motion using the abstract mathematical concept of non-Abelian gauge fields.

7.4 Final Remarks

One message of all this is that there are gauge fields in many places. They arise naturally when we use redundant variables to describe Nature, be they fundamental quantum fields in the Standard Model of particle physics, ambiguous phases of quantum mechanical wave functions, or standard orientations of falling cats. From this point of view, gauge fields are clearly a human invention resulting from our choice of redundant variables. One may speculate whether Nature herself also uses redundancies in order to realize the phenomena that we describe with gauge theories. I personally like to think that this may not be the case. Trying to understand what Nature does (perhaps at the Planck scale) in order to generate effective gauge theories at low energies is interesting and may perhaps even reveal deep insights into

the emergence of space-time at short distances. While all this is highly speculative, there is now doubt that, endowed with great curiosity, theoretical physicists will continue to use their mathematical capabilities to push the boundaries of current knowledge further into the unknown. This course may be viewed as an invitation to participate in this most exciting and potentially quite satisfying enterprise.

Appendix A

Angular Momentum in Quantum Mechanics

Angular momentum is a fundamental quantity that is conserved in any known physical process. Angular momentum conservation is a consequence of the isotropy of space — the laws of Nature are invariant against spatial rotations. Of course, rotation invariance may be broken explicitly under certain conditions, for example, in the presence of external electric or magnetic fields. In that case, the subsystem without the fields is not rotation invariant. Still, the total system behaves in the same way when everything including the fields is spatially rotated. Angular momentum is a vector. In quantum mechanics its components cannot be measured simultaneously because the corresponding operators do not commute. This has interesting consequences for the physical behavior of quantum mechanical particles under spatial rotations.

A.1 Angular Momentum Commutation Relations

Besides the orbital angular momentum $\vec{L} = \vec{r} \times \vec{p}$ familiar from classical mechanics, quantum mechanical particles can carry an internal angular momentum known as spin. While orbital angular momentum is quantized in integer units, spin may be quantized in integer or half-integer units. Interestingly, there is an intimate connection between spin and statistics: particles with half-integer spin obey Fermi-Dirac statistics and are thus fermions, while particles with integer spin obey Bose-Einstein statistics and are hence bosons. This connection between spin and statistics can be understood in the framework of relativistic quantum field theory — but not from quantum mechanics alone.

When particles carry both orbital angular momentum $\vec{L}$ and spin $\vec{S}$, in general only their total angular momentum $\vec{J} = \vec{L} + \vec{S}$ is conserved. In that case, one is confronted with the problem of coupling two angular momenta together. The same problem arises when several particles add their angular momenta together to the conserved angular momentum of the total system. Performing the corresponding angular momentum "gymnastics" is an important tool of the quantum mechanic. Although the subject is a bit formal, its understanding is vital in atomic, molecular and particle physics, as well as in other branches of our field.

Let us consider the commutation relations of an arbitrary angular momentum $\vec{J}$ in quantum mechanics. Here $\vec{J}$ may be an orbital angular momentum, a spin, or any combination of these. For an orbital angular momentum $\vec{L}$ we have already derived the commutation relation

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k, \tag{A.1.1}$$

from the definition $\vec{L} = \vec{r} \times \vec{p}$ and from the fundamental commutation relation $[x_i, p_j] = i\hbar\delta_{ij}$. Now we postulate

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k, \tag{A.1.2}$$

for any angular momentum in quantum mechanics. In particular, different components of the angular momentum vector do not commute with one another. However, all components commute with the magnitude $\vec{J}^2 = J_x^2 + J_y^2 + J_z^2$, i.e. $[J_i, \vec{J}^2] = 0$. As a consequence, one can construct simultaneous eigenstates of the operator $\vec{J}^2$ and one component J_i. Usually one chooses J_z, i.e. the arbitrary quantization direction is then the z-direction. In general, this choice does not imply a physical violation of rotation invariance. One could have chosen any other quantization axis without any effect on the physics. When rotation invariance is already broken, for example, by the direction of an external electric or magnetic field, it is very convenient to choose the quantization axis along the same direction. As usual, we choose the z-direction as our quantization axis and we thus construct simultaneous eigenstates $|j, m\rangle$ of both $\vec{J}^2$ and J_z,

$$\vec{J}^2|j, m\rangle = \hbar^2 j(j+1)|j, m\rangle, \;\; J_z|j, m\rangle = \hbar m|j, m\rangle. \tag{A.1.3}$$

It will turn out that both j and m are either integer or half-integer.

For convenience, we introduce the angular momentum raising and lowering operators

$$J_\pm = J_x \pm iJ_y. \tag{A.1.4}$$

They obey the commutation relations

$$[J_+, J_-] = 2\hbar J_z, \;\; [\vec{J}^2, J_\pm] = 0, \;\; [J_z, J_\pm] = \pm\hbar J_\pm. \tag{A.1.5}$$

We have

$$J_z J_\pm |j,m\rangle = (J_\pm J_z + [J_z, J_\pm])|j,m\rangle = \hbar(m \pm 1) J_\pm |j,m\rangle, \tag{A.1.6}$$

i.e. the state $J_\pm|j,m\rangle$ is also an eigenstate of J_z and has the quantum number $m \pm 1$. Similarly

$$\vec{J}^2 J_\pm |j,m\rangle = J_\pm \vec{J}^2 |j,m\rangle = \hbar^2 j(j+1) J_\pm |j,m\rangle, \tag{A.1.7}$$

i.e. $J_\pm|j,m\rangle$ is still also an eigenstate of $\vec{J}^2$ with the unchanged quantum number j. Since j determines the magnitude of the angular momentum vector, one expects that for fixed j the m quantum number that measures the z-component of the angular momentum vector should be bounded from above and from below. On the other hand, for any state $|j,m\rangle$ with quantum number m one can construct the states

$$J_\pm |j,m\rangle = C_{j,m} |j, m \pm 1\rangle, \tag{A.1.8}$$

with quantum number $m \pm 1$. The apparent contradiction is resolved only if the constant $C_{j,m}$ vanishes for a given $m = m_{max}$ or $m = m_{min}$. Let us compute $C_{j,m}$ from the normalization condition $\langle j, m \pm 1 | j, m \pm 1\rangle = 1$. Using the relation

$$J_\pm^\dagger J_\pm = J_\mp J_\pm = \vec{J}^2 - J_z^2 \mp \hbar J_z, \tag{A.1.9}$$

we obtain

$$\begin{aligned} |C_{j,m}|^2 &= |C_{j,m}|^2 \langle j, m \pm 1 | j, m \pm 1\rangle = \langle j, m | J_\pm^\dagger J_\pm | j, m\rangle \\ &= \langle j, m | \vec{J}^2 - J_z^2 \mp \hbar J_z | j, m\rangle \\ &= \hbar^2 [j(j+1) - m(m \pm 1)]. \end{aligned} \tag{A.1.10}$$

For fixed j, the maximal value m_{max} of the quantum number m is determined by $C_{j,m_{max}} = 0$, which implies $m_{max} = j$. Similarly, the minimal value is determined by $C_{j,m_{min}} = 0$, which implies $m_{min} = -j$. Hence, for fixed j there are $2j + 1$ possible m values

$$m \in \{m_{min}, m_{min} + 1, ..., m_{max} - 1, m_{max}\} = \{-j, -j + 1, ..., j - 1, j\}. \tag{A.1.11}$$

Since the difference $m_{max} - m_{min} = 2j$ is an integer, j can be an integer or a half-integer. Both possibilities are realized in Nature.

A.2 Coupling of Angular Momenta

Let us now couple two angular momenta $\vec{J}_1$ and $\vec{J}_2$ together to a total angular momentum

$$\vec{J} = \vec{J}_1 + \vec{J}_2. \tag{A.2.1}$$

The angular momenta $\vec{J}_1$ and $\vec{J}_2$ could, for example, be orbital angular momentum and spin of the same particle, or angular momenta of two different particles. In any case, since they act in different Hilbert spaces the two angular momentum operators commute with one another, i.e.

$$[J_{1i}, J_{2j}] = 0. \tag{A.2.2}$$

As a consequence, the total angular momentum operator $\vec{J} = \vec{J}_1 + \vec{J}_2$ indeed obeys the usual commutation relations

$$[J_i, J_j] = [J_{1i}, J_{1j}] + [J_{2i}, J_{2j}] = i\hbar\varepsilon_{ijk}(J_{1k} + J_{2k}) = i\hbar\varepsilon_{ijk}J_k. \tag{A.2.3}$$

Let us assume that the states of subsystem 1 have a fixed quantum number j_1 and are given by $|j_1, m_1\rangle$. Similarly, the states of subsystem 2 have quantum number j_2 and are given by $|j_2, m_2\rangle$. Hence, the combined system has $(2j_1+1)(2j_2+1)$ product states $|j_1, m_1\rangle|j_2, m_2\rangle$ which span the Hilbert space of the total system. How does this space decompose into sectors of definite total angular momentum? It will turn out that the possible values for j are restricted by

$$j \in \{|j_1 - j_2|, |j_1 - j_2| + 1, ..., j_1 + j_2\}. \tag{A.2.4}$$

Indeed, the total number of states then is

$$(2|j_1 - j_2| + 1) + (2|j_1 - j_2| + 3) + ... + (2(j_1 + j_2) + 1) = (2j_1 + 1)(2j_2 + 1). \tag{A.2.5}$$

Also the question arises how one can construct states

$$|(j_1, j_2)j, m\rangle = \sum_{m_1,m_2} C_{m_1,m_2}|j_1, m_1\rangle|j_2, m_2\rangle, \tag{A.2.6}$$

as linear combinations of the product states? This is the "gymnastics" problem of coupling together two angular momenta. The factors C_{m_1,m_2} are known as Clebsch-Gordan coefficients.

A.3 Coupling of Two Spins $1/2$

Let us consider the spins of two spin $1/2$ particles, for example, a proton and a neutron forming the atomic nucleus of heavy hydrogen (deuterium). The corresponding bound state of proton and neutron is known as a deuteron. What are the possible total spins of the coupled system? In this case $j_1 = j_2 = 1/2$ and thus

$$j \in \{|j_1 - j_2|, |j_1 - j_2| + 1, ..., j_1 + j_2\} = \{0, 1\}, \tag{A.3.1}$$

i.e. the total spin j is either 0 (a singlet) or 1 (a triplet). Altogether, there are four states. This is consistent because there are also four product states $|j_1, m_1\rangle|j_2, m_2\rangle$ with $m_1 = \pm 1/2$ and $m_2 = \pm 1/2$. For these four states we introduce the short-hand notation

$$\begin{aligned} &|\frac{1}{2}, \frac{1}{2}\rangle|\frac{1}{2}, \frac{1}{2}\rangle = |\uparrow\uparrow\rangle,\ |\frac{1}{2}, \frac{1}{2}\rangle|\frac{1}{2}, -\frac{1}{2}\rangle = |\uparrow\downarrow\rangle, \\ &|\frac{1}{2}, -\frac{1}{2}\rangle|\frac{1}{2}, \frac{1}{2}\rangle = |\downarrow\uparrow\rangle,\ |\frac{1}{2}, -\frac{1}{2}\rangle|\frac{1}{2}, -\frac{1}{2}\rangle = |\downarrow\downarrow\rangle. \end{aligned} \tag{A.3.2}$$

These product states are eigenstates of $J_z = J_{1z} + J_{2z}$,

$$\begin{aligned} &J_z|\uparrow\uparrow\rangle = \hbar(\frac{1}{2} + \frac{1}{2})|\uparrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle, \\ &J_z|\uparrow\downarrow\rangle = \hbar(\frac{1}{2} - \frac{1}{2})|\uparrow\downarrow\rangle = 0,\ J_z|\downarrow\uparrow\rangle = \hbar(-\frac{1}{2} + \frac{1}{2})|\downarrow\uparrow\rangle = 0, \\ &J_z|\downarrow\downarrow\rangle = \hbar(-\frac{1}{2} - \frac{1}{2})|\uparrow\downarrow\rangle = -\hbar|\downarrow\downarrow\rangle. \end{aligned} \tag{A.3.3}$$

The first and the last of the four states must belong to $j = 1$ because they have $m = \pm 1$, i.e.

$$|(\frac{1}{2}, \frac{1}{2})1, 1\rangle = |\uparrow\uparrow\rangle,\ |(\frac{1}{2}, \frac{1}{2})1, -1\rangle = |\downarrow\downarrow\rangle. \tag{A.3.4}$$

The two remaining states have $m = 0$. One linear combination of them is the $m = 0$ state of the triplet, and the orthogonal combination is the $m = 0$ state with $j = 0$. In order to identify the $m = 0$ state with $j = 1$ we act with the lowering operator

$$|(\frac{1}{2}, \frac{1}{2})1, 0\rangle = \frac{1}{\sqrt{2}} J_-|(\frac{1}{2}, \frac{1}{2})1, 1\rangle = \frac{1}{\sqrt{2}}(J_{1-} + J_{2-})|\uparrow\uparrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle). \tag{A.3.5}$$

The orthogonal combination

$$|(\frac{1}{2}, \frac{1}{2})0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \tag{A.3.6}$$

should hence be the state with $j = 0$. This can be checked explicitly, for example, by acting with the operator $\vec{J}_+ = J_{1+} + J_{2+}$, i.e.

$$\vec{J}_+|(\frac{1}{2}, \frac{1}{2})0, 0\rangle = (J_{1+} + J_{2+})\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle - |\uparrow\uparrow\rangle) = 0. \tag{A.3.7}$$

Hence, according to the previous discussion the spin of the deuteron could be 0 or 1. The proton and neutron that form the deuteron nucleus attract each other through the so-called strong interaction. This interaction is spin-dependent. There is a term proportional to $-\vec{J}_1 \cdot \vec{J}_2$ (where $\vec{J}_1$ and $\vec{J}_2$ are the spin operators of proton and neutron) in the proton-neutron potential. One can write

$$\vec{J}^2 = (\vec{J}_1 + \vec{J}_2)^2 = \vec{J}_1^2 + \vec{J}_2^2 + 2\vec{J}_1 \cdot \vec{J}_2, \tag{A.3.8}$$

and hence

$$-\vec{J}_1 \cdot \vec{J}_2 = \frac{1}{2}(\vec{J}_1^2 + \vec{J}_2^2 - \vec{J}^2). \tag{A.3.9}$$

Since both the proton and the neutron have spin 1/2 we have

$$\vec{J}_1^2 = \vec{J}_2^2 = \hbar^2 \frac{1}{2}(\frac{1}{2} + 1) = \frac{3}{4}\hbar^2. \tag{A.3.10}$$

In the spin singlet state we have $\vec{J}^2 = 0$ and hence

$$-\vec{J}_1 \cdot \vec{J}_2 |(\frac{1}{2}, \frac{1}{2})0, 0\rangle = \frac{3}{4}\hbar^2 |(\frac{1}{2}, \frac{1}{2})0, 0\rangle. \tag{A.3.11}$$

In the spin triplet state, on the other hand, $\vec{J}^2 = 2\hbar^2$ such that

$$-\vec{J}_1 \cdot \vec{J}_2 |(\frac{1}{2}, \frac{1}{2})1, m\rangle = -\frac{1}{4}\hbar^2 |(\frac{1}{2}, \frac{1}{2})1, m\rangle. \tag{A.3.12}$$

In the triplet channel the proton-neutron interaction is attractive while in the singlet channel it is repulsive. As a consequence, the deuteron has spin 1, while in the spin 0 channel there is no bound state.

The strong interaction is mediated by gluons, just as electromagnetic interactions are mediated by photons. The strong interaction analogs of electrons and positrons are quarks and anti-quarks — the basic building blocks of protons and neutrons. There are two u-quarks and one d-quark in each proton. A u-quark has electric charge 2/3 and a d-quark has $-1/3$. Hence, the charge of a proton is indeed $2 \times 2/3 - 1/3 = 1$. Similarly, a neutron contains one u-quark and two d-quarks and thus has charge $2/3 - 2 \times 1/3 = 0$. Like electrons, quarks are fermions with spin 1/2. What is the possible total spin of a bound system of three quarks? As we learned before, two spin 1/2 particles can couple to a total spin $j = 0$ or $j = 1$. When a third spin 1/2 particle is added to the $j = 0$ state, the total spin is 1/2. If it is added to the $j = 1$ state the total spin can be either $j - 1/2 = 1/2$ or $j + 1/2 = 3/2$. Again, the strong interactions between quarks are spin-dependent and they favor the total spin 1/2 states corresponding to proton and neutron. The spin 3/2 state also exists but is unstable. It is known in particle physics as the Δ-isobar.

A.4 Coupling of Orbital Angular Momentum and Spin

Just as there are spin-dependent strong interactions between protons and neutrons or between quarks, there are also spin-dependent electromagnetic interactions between

electrons and protons. In particular, there are spin-orbit coupling terms proportional to

$$\vec{L}\cdot\vec{S} = \frac{1}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2) = \frac{\hbar^2}{2}(j(j+1) - l(l+1) - \frac{3}{4}). \tag{A.4.1}$$

Let us consider the coupling of an orbital angular momentum l and a spin $s = 1/2$ of an electron in more detail. In this case, there are $2(2l+1)$ product states $|l, m_l\rangle|s, m_s\rangle$. The possible values of the total angular momentum are

$$j \in \{|l-s|, l+s\} = \{l - \frac{1}{2}, l + \frac{1}{2}\}. \tag{A.4.2}$$

Indeed, there are again

$$2(l - \frac{1}{2}) + 1 + 2(l + \frac{1}{2}) + 1 = 2(2l+1) \tag{A.4.3}$$

states. The direct product state

$$|(l, \frac{1}{2})\ l + \frac{1}{2}, l + \frac{1}{2}\rangle = |l, l\rangle|\uparrow\rangle \tag{A.4.4}$$

has $m = l + 1/2$ and must thus have $j = l + 1/2$. We can construct the other states with $j = l + 1/2$ and with lower m-values by acting with $J_- = L_- + S_-$. For example,

$$\begin{aligned} |(l, \frac{1}{2})\ l + \frac{1}{2}, l - \frac{1}{2}\rangle &= \frac{1}{\sqrt{2l+1}} J_- |l, l\rangle|\uparrow\rangle \\ &= \frac{1}{\sqrt{2l+1}}(L_- + S_-)|l, l\rangle|\uparrow\rangle \\ &= \frac{1}{\sqrt{2l+1}}(|l, l\rangle|\downarrow\rangle + \sqrt{2l}|l, l-1\rangle|\uparrow\rangle). \end{aligned} \tag{A.4.5}$$

The orthogonal combination

$$|(l, \frac{1}{2})\ l - \frac{1}{2}, l - \frac{1}{2}\rangle = \frac{1}{\sqrt{2l+1}}(\sqrt{2l}|l, l\rangle|\downarrow\rangle - |l, l-1\rangle|\uparrow\rangle), \tag{A.4.6}$$

has $j = l - 1/2$ and $m = l - 1/2$. Again, by acting with J_- one can generate all other states with $j = l - 1/2$ and smaller m-values.

Appendix B

The Constituent Quark Model

The dynamics of quarks and gluons, which are permanently confined inside hadrons, i.e. mesons and baryons, is described by Quantum Chromodynamics (QCD). QCD is a strngly coupled relativistic quantum field theory, whose dynamics can be investigated from first principles using the regularization on a space-time lattice. The constituent quark model, provides a very crude approximation of these dynamics. It describes baryons as 3-quark states and mesons as quark-anti-quark states in a rather naive manner. Still, it accounts for the most prominent particle states that exist in the QCD spectrum. Here we are interested in the constituent quark model as an application of the group theory of $SU(2)$ and $SU(3)$.

B.1 Isospin Symmetry

Proton and neutron have almost the same masses

$$M_p = 0.938 \text{ GeV}, \; M_n = 0.940 \text{ GeV}. \tag{B.1.1}$$

While the proton seems to be absolutely stable, a free neutron decays radioactively into a proton, an electron and an electron-anti-neutrino $n \to p + e + \bar{\nu}_e$. Protons and neutrons (the nucleons) are the constituents of atomic nuclei. Originally, Yukawa postulated a light particle mediating the interaction between protons and neutrons. This π-meson or pion is a boson with spin 0, which exists in three charge states π^+, π^0 and π^-. The corresponding masses are

$$M_{\pi^+} = M_{\pi^-} = 0.140 \text{ GeV}, \; M_{\pi^0} = 0.135 \text{ GeV}. \tag{B.1.2}$$

In pion-nucleon scattering a resonance occurs in the total cross section as a function of the pion-nucleon center of mass energy. The resonance energy is interpreted as

Hadron	Representation	I	I_3	Q	S
p, n	$\{2\}$	$\frac{1}{2}$	$\frac{1}{2}$, $-\frac{1}{2}$	1, 0	$\frac{1}{2}$
Δ^{++}, Δ^{+}, Δ^{0}, Δ^{-}	$\{4\}$	$\frac{3}{2}$	$\frac{3}{2}$, $\frac{1}{2}$, $-\frac{1}{2}$, $-\frac{3}{2}$	2, 1, 0, -1	$\frac{3}{2}$
π^{+}, π^{0}, π^{-}	$\{3\}$	1	1, 0, -1	1, 0, -1	0
ρ^{+}, ρ^{0}, ρ^{-}	$\{3\}$	1	1, 0, -1	1, 0, -1	1

Table B.1: *The isospin classification of hadrons.*

the mass of an unstable particle — the so-called Δ-isobar. One may view the Δ-particle as an excited state of the nucleon. It exists in four charge states Δ^{++}, Δ^{+}, Δ^{0} and Δ^{-} with masses

$$M_{\Delta^{++}} \approx M_{\Delta^{+}} \approx M_{\Delta^{0}} \approx M_{\Delta^{-}} \approx 1.232 \text{ GeV} \tag{B.1.3}$$

Similar to pion-nucleon scattering there is also a resonance in pion-pion scattering. This so-called ρ-meson comes in three charge states ρ^{+}, ρ^{0} and ρ^{-} with masses

$$M_{\rho^{+}} \approx M_{\rho^{0}} \approx M_{\rho^{-}} \approx 0.768 \text{ GeV}. \tag{B.1.4}$$

Particles with different electric charges have (almost) degenerate masses, and it is natural to associate this with an (approximate) symmetry. This so-called isospin symmetry is similar to the ordinary spin $SU(2)$ rotational symmetry. Isospin is, however, not related to space-time transformations, it is an intrinsic symmetry. As we know each total spin $S = 0, 1/2, 1, 3/2, ...$ is associated with an irreducible representation of the $SU(2)_S$ rotation group containing $2S + 1$ states distinguished by their spin projection

$$S_z = -S, -S + 1, ..., S - 1, S. \tag{B.1.5}$$

In complete analogy the representations of the $SU(2)_I$ isospin symmetry group are characterized by their total isospin $I = 0, 1/2, 1, 3/2, ...$ The states of an isospin representation are distinguished by their isospin projection

$$I_3 = -I, -I + 1, ..., I - 1, I. \tag{B.1.6}$$

A representation with isospin I contains $2I+1$ states and is denoted by $\{2I+1\}$. We can classify the hadrons by their isospin. This is done in table B.1. For the baryons (nucleon and Δ) isospin projection and electric charge are related by $Q = I_3 + \frac{1}{2}$, and for the mesons (π and ρ) $Q = I_3$.

Isospin is an (approximate) symmetry of the strong interactions. For example, the proton-pion scattering reaction $p + \pi \to \Delta$ is consistent with isospin symmetry because the coupling of the isospin representations of nucleon and pion

$$\{2\} \otimes \{3\} = \{2\} \oplus \{4\} \tag{B.1.7}$$

does indeed contain the quadruplet isospin 3/2 representation of the Δ-isobar. The isospin symmetry of the hadron spectrum indicates that the strong interactions are charge independent. This is no surprise because the charge Q is responsible for the electromagnetic but not for the strong interactions.

B.2 Nucleon and Δ-Isobar in the Quark Model

We want to approach the question of the hadronic constituents by investigating various symmetries. First we consider isospin. Since the hadrons form isospin multiplets the same should be true for their constituents. The only $SU(2)$ representation from which we can generate all others is the fundamental representation — the isospin doublet $\{2\}$ with $I = 1/2$ and $I_3 = \pm 1/2$. We identify the two states of this multiplet with the constituent quarks up ($I_3 = 1/2$) and down ($I_3 = -1/2$). A constituent quark is a quasiparticle carrying the same quantum numbers as a fundamental (current) quark, but also containing numerous gluons. After all, a constituent quark is not a very well defined object. We can view it as a basic building block for hadrons that plays a role in some simple phenomenological models for the strong interactions. Still, the concept of constituent quarks leads to a rather successful group theoretical classification scheme for hadrons.

Since the Δ-isobar has isospin 3/2 it contains at least three constituent quarks. We couple

$$\{2\} \otimes \{2\} \otimes \{2\} = (\{1\} \oplus \{3\}) \otimes \{2\} = \{2\} \oplus \{2\} \oplus \{4\}, \tag{B.2.1}$$

and we do indeed find a quadruplet. For the charges of the baryons we have

$$Q = I_3 + \frac{1}{2} = \sum_{q=1}^{3} (I_{3q} + \frac{1}{6}) = \sum_{q=1}^{3} Q_q, \tag{B.2.2}$$

and hence we obtain for the charges of the quarks

$$Q_q = I_{3q} + \frac{1}{6},\ Q_u = \frac{1}{2} + \frac{1}{6} = \frac{2}{3},\ Q_d = -\frac{1}{2} + \frac{1}{6} = -\frac{1}{3}. \tag{B.2.3}$$

The quarks have fractional electric charges. Using Clebsch-Gordon coefficients of $SU(2)$ one finds

$$\begin{aligned}
&\boxed{1}\boxed{2}\boxed{3}_{3/2} = uuu \equiv \Delta^{++},\\
&\boxed{1}\boxed{2}\boxed{3}_{1/2} = \frac{1}{\sqrt{3}}(uud + udu + duu) \equiv \Delta^{+},\\
&\boxed{1}\boxed{2}\boxed{3}_{-1/2} = \frac{1}{\sqrt{3}}(udd + dud + ddu) \equiv \Delta^{0},\\
&\boxed{1}\boxed{2}\boxed{3}_{-3/2} = ddd \equiv \Delta^{-}.
\end{aligned} \tag{B.2.4}$$

These isospin states are completely symmetric against permutations of the constituent quarks.

We write the general coupling of the three quarks as

$$\boxed{1} \otimes \boxed{2} \otimes \boxed{3} = \boxed{1\,2\,3} \oplus \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \oplus \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \oplus \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \cline{1-1}\end{array}. \tag{B.2.5}$$

Translated into $SU(2)$ language this equation reads

$$\{2\} \otimes \{2\} \otimes \{2\} = \{4\} \oplus \{2\} \oplus \{2\} \oplus \{0\}. \tag{B.2.6}$$

Here $\{0\}$ denotes an empty representation — one that cannot be realized in $SU(2)$ because the corresponding Young tableau has more than two rows. We identify the totally symmetric representation as the four charge states of the Δ-isobar, and we write as before $\boxed{1\,2\,3}_{I_3}$.

Before we can characterize the state of the Δ-isobar in more detail we must consider the other symmetries of the problem. The Δ-isobar is a resonance in the scattering of spin 1/2 nucleons and spin 0 pions. The experimentally observed spin of the resonance is 3/2. To account for this we associate a spin 1/2 with the constituent quarks. Then, in complete analogy to isospin, we can construct a totally symmetric spin representation for the Δ-particle

$$\begin{aligned}
\boxed{1\,2\,3}_{3/2} &= \uparrow\uparrow\uparrow, \\
\boxed{1\,2\,3}_{1/2} &= \frac{1}{\sqrt{3}}(\uparrow\uparrow\downarrow \;+\; \uparrow\downarrow\uparrow \;+\; \downarrow\uparrow\uparrow), \\
\boxed{1\,2\,3}_{-1/2} &= \frac{1}{\sqrt{3}}(\uparrow\downarrow\downarrow \;+\; \downarrow\uparrow\downarrow \;+\; \downarrow\downarrow\uparrow), \\
\boxed{1\,2\,3}_{-3/2} &= \downarrow\downarrow\downarrow\,.
\end{aligned} \tag{B.2.7}$$

The isospin-spin part of the Δ-isobar state hence takes the form

$$|\Delta I_3 S_z\rangle = \boxed{1\,2\,3}_{I_3} \boxed{1\,2\,3}_{S_z}. \tag{B.2.8}$$

This state is symmetric with respect to both isospin and spin. Consequently, it is symmetric under simultaneous isospin-spin permutations. For illustrative purposes we write down the state for a Δ^+ particle with spin projection $S_z = 1/2$

$$\begin{aligned}
|\Delta\tfrac{1}{2}\tfrac{1}{2}\rangle \;=\; & \frac{1}{3}(u\uparrow u\uparrow d\downarrow + u\uparrow u\downarrow d\uparrow + u\downarrow u\uparrow d\uparrow \\
+\; & u\uparrow d\uparrow u\downarrow + u\uparrow d\downarrow u\uparrow + u\downarrow d\uparrow u\uparrow \\
+\; & d\uparrow u\uparrow u\downarrow + d\uparrow u\downarrow u\uparrow + d\downarrow u\uparrow u\uparrow).
\end{aligned} \tag{B.2.9}$$

One sees explicitly that this state is totally symmetric.

As we have seen, the Young tableau $\begin{array}{|c|c|}\hline&\\\hline\\\cline{1-1}\end{array}$ is associated with the isodoublet $\{2\}$. Hence, it is natural to expect that the nucleon state can be constructed from it. Now we have two possibilities $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{I_3}$ and $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{I_3}$ corresponding to symmetric or antisymmetric couplings of the quarks 1 and 2. Using Clebsch-Gordon coefficients one finds

$$\begin{aligned}
\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{1/2} &= \frac{1}{\sqrt{6}}(2uud - udu - duu),\\
\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{-1/2} &= \frac{1}{\sqrt{6}}(udd + dud - 2ddu),\\
\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{1/2} &= \frac{1}{\sqrt{2}}(udu - duu),\\
\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{-1/2} &= \frac{1}{\sqrt{2}}(udd - dud). \qquad \text{(B.2.10)}
\end{aligned}$$

Proton and neutron have spin 1/2. Hence, we have two possible coupling schemes $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{S_z}$ and $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{S_z}$. We now want to combine the mixed isospin and spin permutation symmetries to an isospin-spin representation of definite permutation symmetry. This requires to reduce the inner product

$$\begin{array}{|c|c|}\hline&\\\hline\\\cline{1-1}\end{array}_{I_3} \times \begin{array}{|c|c|}\hline&\\\hline\\\cline{1-1}\end{array}_{S_z} = \begin{array}{|c|c|c|}\hline&&\\\hline\end{array}_{I_3S_z} \oplus \begin{array}{|c|c|}\hline&\\\hline\\\cline{1-1}\end{array}_{I_3S_z} \oplus \begin{array}{|c|}\hline\\\hline\\\hline\\\hline\end{array}_{I_3S_z} \qquad \text{(B.2.11)}$$

in S_3. The two isospin and spin representations can be coupled to a symmetric, mixed or antisymmetric isospin-spin representation. As for the Δ-isobar we want to couple isospin and spin symmetrically. To do this explicitly, we need the Clebsch-Gordon coefficients of the group S_3. One finds

$$|NI_3S_z\rangle = \frac{1}{\sqrt{2}}\Big(\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{I_3}\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}_{S_z} + \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{I_3}\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}_{S_z}\Big). \qquad \text{(B.2.12)}$$

In our construction we have implicitly assumed that the orbital angular momentum of the constituent quarks inside a hadron vanishes. Then the orbital state is

completely symmetric in the coordinates of the quarks. The orbital part of the baryon wave function therefore is described by the Young tableau $\square\square\square$. Since also the isospin-spin part is totally symmetric, the baryon wave function is completely symmetric under permutations of the quarks. Since we have treated constituent quarks as spin 1/2 fermions, this contradicts the Pauli principle which requires a totally antisymmetric fermion wave function, and hence the Young tableau $\begin{array}{|c|}\hline\ \\\hline\ \\\hline\ \\\hline\end{array}$. To satisfy the Pauli principle the color symmetry comes to our rescue. In $SU(3)_c$, $\begin{array}{|c|}\hline\ \\\hline\ \\\hline\ \\\hline\end{array}$ corresponds to a singlet representation, which means that baryons are color-neutral. Since we have three colors we can now completely antisymmetrize three quarks

$$\begin{array}{|c|}\hline\ \\\hline\ \\\hline\ \\\hline\end{array} = \frac{1}{\sqrt{6}}(rgb - rbg + gbr - grb + brg - bgr). \tag{B.2.13}$$

The color symmetry is the key to the fundamental understanding of the strong interactions. As opposed to isospin, color is an exact and even local symmetry.

B.3 Anti-Quarks and Mesons

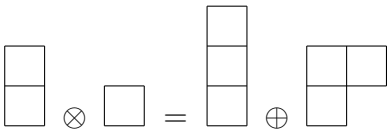
We have seen that the baryons (nucleon and Δ) consist of three constituent quarks (isospin doublets, spin doublets, color triplets). Now we want to construct the mesons (pion and ρ) in a similar manner. Since these particles have spin 0 and 1 respectively, they must contain an even number of constituent quarks. When we use two quarks, i.e. when we construct states like uu, ud, or dd, the resulting electric charges are 4/3, 1/3, and $-2/3$ in contradiction to experiment. Also the coupling of two color triplets

$$\begin{array}{|c|}\hline\ \\\hline\end{array} \otimes \begin{array}{|c|}\hline\ \\\hline\end{array} = \begin{array}{|c|c|}\hline\ & \ \\\hline\end{array} \oplus \begin{array}{|c|}\hline\ \\\hline\ \\\hline\end{array}$$
$$\{3\} \otimes \{3\} = \{6\} \oplus \{\bar{3}\}, \tag{B.3.1}$$

does not contain a singlet as desired by the confinement hypothesis.

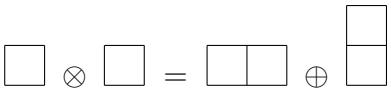
We have seen already that a representation together with its conjugate represen-

tation can always be coupled to a singlet. In $SU(3)$ this corresponds to



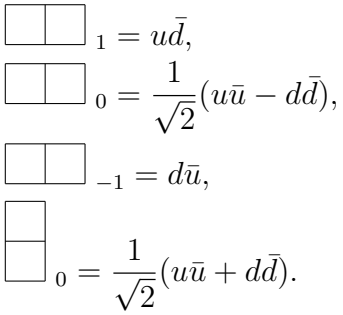
$$\{\bar{3}\} \otimes \{3\} = \{1\} \oplus \{8\}, \tag{B.3.2}$$

Hence it is natural to work with anti-quarks. Anti-quarks are isospin doublets, spin doublets, and color anti-triplets. We have quarks $\bar{u}$ and $\bar{d}$ with electric charges $Q_{\bar{u}} = -2/3$ and $Q_{\bar{d}} = 1/3$. Now we consider combinations of quark and anti-quark $u\bar{d}$, $u\bar{u}$, $d\bar{d}$, and d$\bar{u}$, which have charges 1, 0, and -1 as we need them for the mesons. First we couple the isospin wave function



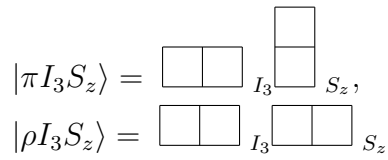
$$\{2\} \otimes \{2\} = \{3\} \oplus \{1\}, \tag{B.3.3}$$

and we obtain



$$\begin{aligned}
&= u\bar{d}, \\
&= \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}), \\
&= d\bar{u}, \\
&= \frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d}).
\end{aligned} \tag{B.3.4}$$

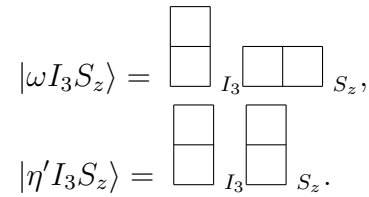
We proceed analogously for the spin and we obtain



$$\begin{aligned}
|\pi I_3 S_z\rangle &= \quad_{I_3} \quad_{S_z}, \\
|\rho I_3 S_z\rangle &= \quad_{I_3} \quad_{S_z}.
\end{aligned} \tag{B.3.5}$$

Since quarks and anti-quarks are distinguishable particles (for example they have different charges) we don't have to respect the Pauli principle in this case. As opposed to the baryons here the coupling to color singlets follows only from the confinement hypothesis.

Of course, we can combine isospin and spin wave functions also in a different way



$$\begin{aligned}
|\omega I_3 S_z\rangle &= \quad_{I_3} \quad_{S_z}, \\
|\eta' I_3 S_z\rangle &= \quad_{I_3} \quad_{S_z}.
\end{aligned} \tag{B.3.6}$$

Indeed one observes mesons with these quantum numbers with masses $M_\omega = 0.782\text{GeV}$ and $M_{\eta'} = 0.958\text{GeV}$.

B.4 Strange Hadrons

Up to now we have considered hadrons that consist of up and down quarks and their anti-particles. However, one also observes hadrons containing strange quarks. The masses of the scalar ($S = 0$) mesons are given by

$$M_\pi = 0.138\text{GeV}, M_K = 0.496\text{GeV}, M_\eta = 0.549\text{GeV}, M_{\eta'} = 0.958\text{GeV}, \quad \text{(B.4.1)}$$

while the vector ($S = 1$) meson masses are

$$M_\rho = 0.770\text{GeV}, M_\omega = 0.783\text{GeV}, M_{K^*} = 0.892\text{GeV}, M_\varphi = 1.020\text{GeV}. \quad \text{(B.4.2)}$$

Altogether we have nine scalar and nine vector mesons. In each group we have so far classified four (π^+, π^0, π^-, η' and ρ^+, ρ^0, ρ^-, ω). The number four resulted from the $SU(2)_I$ isospin relation

$$\{\bar{2}\} \otimes \{2\} = \{1\} \oplus \{3\}. \quad \text{(B.4.3)}$$

The number nine then suggests to consider the corresponding $SU(3)$ identity

$$\{\bar{3}\} \otimes \{3\} = \{1\} \oplus \{8\}. \quad \text{(B.4.4)}$$

Indeed we obtain nine mesons if we generalize isospin to a larger symmetry $SU(3)_F$. This so-called flavor group has nothing to with with the color symmetry $SU(3)_c$. It is only an approximate symmetry of QCD, with $SU(2)_I$ as a subgroup. In $SU(3)_F$ we have another quark flavor s — the strange quark.

The generators of $SU(3)$ can be chosen as follows

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \ \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \ \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \ \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \quad \text{(B.4.5)}$$

Two of the generators commute with each other $[\lambda_3, \lambda_8] = 0$. We say that the group $SU(3)$ has rank 2. One can now identify the generators of the isospin subgroup $SU(2)_I$

$$I_1 = \frac{1}{2}\lambda_1,\ I_2 = \frac{1}{2}\lambda_2,\ I_3 = \frac{1}{2}\lambda_3. \tag{B.4.6}$$

Also it is convenient to introduce the so-called strong hypercharge

$$Y = \frac{1}{\sqrt{3}}\lambda_8, \tag{B.4.7}$$

(not to be confused with the generator of $U(1)_Y$ gauge transformations in the standard model). Then I^2, I_3, and Y commute with each other, and we can characterize the states of an $SU(3)_F$ multiplet by their isospin quantum numbers and by their hypercharge. Starting with the $SU(3)_F$ triplet we have

$$\begin{aligned} I^2 u &= \frac{1}{2}(\frac{1}{2}+1)u = \frac{3}{4}u,\ I_3 u = \frac{1}{2}u,\ Yu = \frac{1}{3}u, \\ I^2 d &= \frac{1}{2}(\frac{1}{2}+1)d = \frac{3}{4}d,\ I_3 d = -\frac{1}{2}d,\ Yd = \frac{1}{3}d, \\ I^2 s &= 0,\ I_3 s = 0,\ Ys = -\frac{2}{3}s. \end{aligned} \tag{B.4.8}$$

The electric charge is now given by

$$Q = I_3 + \frac{1}{2}Y, \tag{B.4.9}$$

such that

$$Q_u = \frac{2}{3},\ Q_d = -\frac{1}{3},\ Q_s = -\frac{1}{3}, \tag{B.4.10}$$

i.e. the charge of the strange quark is the same as the one of the down quark. If $SU(3)_F$ would be a symmetry as good as $SU(2)_I$ the states in an $SU(3)_F$ multiplet should be almost degenerate. This is, however, not quite the case, and $SU(3)_F$ is only approximately a symmetry of QCD.

Of course, we can also include the s quark in baryons. Then we have

$$\{3\} \otimes \{3\} \otimes \{3\} = \{10\} \oplus 2\{8\} \oplus \{1\} \tag{B.4.11}$$

compared to the old $SU(2)_I$ result

$$\{2\} \otimes \{2\} \otimes \{2\} = \{4\} \oplus 2\{2\} \oplus \{0\}. \tag{B.4.12}$$

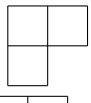
Indeed one observes more baryons than just nucleon and Δ-isobar.

The baryon masses for the spin 1/2 baryons are

$$M_N = 0.939\text{GeV},\ M_\Lambda = 1.116\text{GeV},\ M_\Sigma = 1.193\text{GeV},\ M_\Xi = 1.318\text{GeV}, \quad \text{(B.4.13)}$$

while the spin 3/2 baryon masses are

$$M_\Delta = 1.232\text{GeV},\ M_{\Sigma^*} = 1.385\text{GeV},\ M_{\Xi^*} = 1.530\text{GeV},\ M_\Omega = 1.672\text{GeV}. \quad \text{(B.4.14)}$$

Proton and neutron are part of an octet:  is $\{2\}$ in $SU(2)_I$ and $\{8\}$ in $SU(3)_F$. The Δ-isobar is part of a decouplet:  is $\{4\}$ in $SU(2)_I$ and $\{10\}$ in $SU(3)_F$. One does not find an $SU(3)_F$ singlet  . This is because a spatially symmetric color singlet wave function is totally antisymmetric. To obtain a totally antisymmetric wave function also the spin part should transform as  . Of course, in $SU(2)_S$ this is impossible.

We want to assume that the $SU(3)_F$ symmetry is explicitly broken because the s quark is heavier than the u and d quarks. Based on the quark content one would expect

$$M_{\Sigma^*} - M_\Delta = M_{\Xi^*} - M_{\Sigma^*} = M_\Omega - M_{\Xi^*} = M_s - M_q. \quad \text{(B.4.15)}$$

In fact one finds experimentally

$$M_{\Sigma^*} - M_\Delta = 0.153\text{GeV},\ M_{\Xi^*} - M_{\Sigma^*} = 0.145\text{GeV},\ M_\Omega - M_{\Xi^*} = 0.142\text{GeV}. \quad \text{(B.4.16)}$$

B.5 Gellman-Okubo Baryon Mass Formula

We have seen that the constituent quark model leads to a successful classification of hadron states in terms of flavor symmetry. The results about the hadron dynamics are, however, of more qualitative nature, and the assumption that a hadron is essentially a collection of a few constituent quarks is certainly too naive. The fundamental theory of the strong interactions is QCD. Here we want to use very basic QCD physics together with group theory to describe patterns in the hadron spectrum. The interaction between quarks and gluons is flavor-independent, and therefore $SU(3)_F$ symmetric. Also the gluon self-interaction is flavor symmetric because the gluons are flavor singlets. A violation of flavor symmetry results only from

the quark mass matrix

$$\mathcal{M} = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}. \tag{B.5.1}$$

We want to assume that u and d quark have the same mass m_q, while the s quark is heavier ($m_s > m_q$). The quark mass matrix can be written as

$$\begin{aligned} \mathcal{M} &= \frac{2m_q + m_s}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{m_q - m_s}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\ &= \frac{2m_q + m_s}{3} \mathbf{1} + \frac{m_q - m_s}{\sqrt{3}} \lambda_8. \end{aligned} \tag{B.5.2}$$

The mass matrix contains an $SU(3)_F$ singlet as well as an octet piece. Correspondingly, the QCD Hamilton operator can be written as

$$H_{QCD} = H_1 + H_8. \tag{B.5.3}$$

We want to assume that H_8 is small and can be treated as a perturbation. Then we first consider H_1 alone. This is justified if the mass difference $m_q - m_s$ is small. Since H_1 is $SU(3)_F$ symmetric we expect degenerate states in $SU(3)_F$ multiplets — the hadron octets and decouplets. Here we assume that the flavor symmetry is not spontaneously broken. This should indeed be correct for QCD.

Let us start with the baryons. The eigenstates of H_1 are denoted by $|B_1 Y I I_3\rangle$

$$H_1 |B_1 Y I I_3\rangle = M_{B_1} |B_1 Y I I_3\rangle. \tag{B.5.4}$$

We use degenerate perturbation theory to first order in H_8 and obtain

$$M_B = M_{B_1} + \langle B_1 Y I I_3 | H_8 | B_1 Y I I_3 \rangle. \tag{B.5.5}$$

A diagonalization in the space of degenerate states is not necessary, since H_8 transforms as the λ_8 component of an octet, and can therefore not change Y, I, and I_3. Next we will compute the required matrix elements using group theory. Starting with the baryon decouplet, we obtain a nonzero value only if $\{8\}$ and $\{10\}$ can couple to $\{10\}$. Indeed the decouplet appears in the reduction. Using the Wigner-Eckart theorem we obtain

$$\langle B_1 Y I I_3 | H_8 | B_1 Y I I_3 \rangle = \langle B_1 || H_8 || B_1 \rangle \langle \{10\} Y I I_3 | \{8\} 000 \{10\} Y I I_3 \rangle, \tag{B.5.6}$$

where $\langle B_1 || H_8 || B_1 \rangle$ is a reduced matrix element, and the second factor is an $SU(3)_F$ Clebsch-Gordon coefficient given by

$$\langle \{10\} Y I I_3 | \{8\} 000 \{10\} Y I I_3 \rangle = \frac{Y}{\sqrt{8}}. \tag{B.5.7}$$

Then we obtain for the baryon masses in the decouplet

$$M_B = M_{B_1} + \langle B_1||H_8||B_1\rangle \frac{Y}{\sqrt{8}}, \tag{B.5.8}$$

and hence

$$M_{\Sigma^*} - M_\Delta = M_{\Xi^*} - M_{\Sigma^*} = M_\Omega - M_{\Xi^*} = -\frac{1}{\sqrt{8}}\langle B_1||H_8||B_1\rangle. \tag{B.5.9}$$

Indeed, as we saw before, the three mass differences are almost identical. In view of the fact that we have just used first order perturbation theory, this is quite remarkable.

Next we consider the mass splittings in the baryon octet. Here we must ask if $\{8\}$ and $\{8\}$ can couple to $\{8\}$. One finds

$$\{8\} \otimes \{8\} = \{27\} \oplus \{10\} \oplus \{\bar{10}\} \oplus 2\{8\} \oplus \{1\}. \tag{B.5.10}$$

Hence there are even two ways to couple two octets to an octet. One is symmetric, the other is antisymmetric under the exchange of the two octets. We can write

$$\begin{aligned} \langle B_1 Y I I_3|H_8|B_1 Y I I_3\rangle &= \langle B_1||H_8||B_1\rangle_s \langle\{8\}Y I I_3|\{8\}000\{8\}Y I I_3\rangle_s \\ &+ \langle B_1||H_8||B_1\rangle_a \langle\{8\}Y I I_3|\{8\}000\{8\}Y I I_3\rangle_a. \end{aligned} \tag{B.5.11}$$

The Clebsch-Gordon coefficients are given by

$$\begin{aligned} \langle\{8\}Y I I_3|\{8\}000\{8\}Y I I_3\rangle_s &= \frac{1}{\sqrt{5}}(I(I+1) - \frac{1}{4}Y^2 - 1), \\ \langle\{8\}Y I I_3|\{8\}000\{8\}Y I I_3\rangle_a &= \sqrt{\frac{3}{4}}Y, \end{aligned} \tag{B.5.12}$$

and we obtain for the baryon octet

$$M_B = M_{B_1} + \langle B_1||H_8||B_1\rangle_s \frac{1}{\sqrt{5}}(I(I+1) - \frac{1}{4}Y^2 - 1) + \langle B_1||H_8||B_1\rangle_a \sqrt{\frac{3}{4}}Y. \tag{B.5.13}$$

These formulas for the baryon masses were first derived by Gellman and Okubo. From the octet formula one obtains

$$\begin{aligned} 2M_N + 2M_\Xi &= 4M_{B_1} + \langle B_1||H_8||B_1\rangle_s \frac{4}{\sqrt{5}}(\frac{3}{4} - \frac{1}{4} - 1), \\ M_\Sigma + 3M_\Lambda &= 4M_{B_1} + \langle B_1||H_8||B_1\rangle_s \frac{1}{\sqrt{5}}((2-1) + 3(-1)), \\ 2M_N + 2M_\Xi &= M_\Sigma + 3M_\Lambda. \end{aligned} \tag{B.5.14}$$

Experimentally the two sides of the last equation give 1.129 GeV and 1.135 GeV in excellent agreement with the theory.

Appendix C

Structure of Minkowski Space-Time

Relativistic theories such as Einstein's special relativity or Maxwell's electrodynamics are invariant not only under spatial translations and rotations but also under Lorentz transformations that rotate spatial and temporal coordinates into one another. The resulting enlarged invariance is known as Lorentz invariance (against space-time rotations) or as Poincaré invariance against space-time rotations and translations.

C.1 Lorentz Transformations

Hermann Minkowski was first to realize that in relativistic theories space and time (which are separate entities in Newtonian mechanics) are naturally united to space-time. A point in Minkowski space-time is described by four coordinates — one for time and three for space — which form a 4-vector

$$x^\mu = (x^0, x^1, x^2, x^3) \ , \quad x^0 = c\ t \ . \tag{C.1.1}$$

In particular, the time t (multiplied by the velocity of light c) plays the rôle of the zeroth component of the 4-vector. From now on, we choose natural units and simply put $c = 1$. Minkowski's space-time does not have Euclidean geometry. In particular, the length squared of the 4-vector x^μ is given by

$$s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \ , \tag{C.1.2}$$

and may thus be negative. Besides the contra-variant vector x^μ it is useful to introduce an equivalent form

$$x_\mu = (x_0, x_1, x_2, x_3) \ , \tag{C.1.3}$$

the co-variant 4-vector. Both the co- and the contra-variant 4-vectors contain the same physical information. Their components are simply related by

$$x_0 = x^0 \ , \quad x_1 = -x^1 \ , \quad x_2 = -x^2 \ , \quad x_3 = -x^3 \ . \tag{C.1.4}$$

The length squared of the 4-vector can then be written as

$$\sum_{\mu=0}^{3} x_\mu x^\mu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = s^2 \ . \tag{C.1.5}$$

Instead of always writing sums over space-time indices μ explicitly, Einstein has introduced a summation convention according to which repeated indices (one co- and one contra-variant index) will automatically be summed. Using Einstein's summation convention the above equation simply takes the form

$$x_\mu x^\mu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = s^2 \ . \tag{C.1.6}$$

The summation over μ is no longer written explicitly, but is still implicitly understood, because the index μ occurs twice (once as a co- and once as a contra-variant one).

The norm of a 4-vector induces a corresponding metric via

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} g_{\mu\nu} x^\mu x^\nu = g_{\mu\nu} x^\mu x^\nu \ . \tag{C.1.7}$$

In the last step, we have again used Einstein's summation convention and have dropped the explicit sums over the repeated indices μ and ν. The metric tensor g with the elements $g_{\mu\nu}$ is a 4×4 matrix given by

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \ . \tag{C.1.8}$$

The metric tensor can also be used to relate co- and contra-variant 4-vectors by lowering a contra-variant index, i.e.

$$x_\mu = g_{\mu\nu} x^\nu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu \ . \tag{C.1.9}$$

Again, the repeated index ν is summed over, while the unrepeated index μ is not summed. Let us also introduce the inverse metric g^{-1} with the components $g^{\mu\nu}$ which is given by

$$g^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \tag{C.1.10}$$

and which obeys

$$g g^{-1} = \mathbf{1} . \tag{C.1.11}$$

In components this relation takes the form

$$g_{\mu\nu} g^{\nu\rho} = \sum_{\nu=0}^{3} g_{\mu\nu} g^{\nu\rho} = \delta_{\mu}{}^{\rho} , \tag{C.1.12}$$

where $\delta_{\mu}{}^{\rho}$ is just the Kronecker symbol, i.e. it represents the matrix elements of the unit matrix $\mathbf{1}$. The inverse metric can now be used to raise co-variant indices, e.g.

$$x^{\mu} = g^{\mu\nu} x_{\nu} = \sum_{\nu=0}^{3} g^{\mu\nu} x_{\nu} . \tag{C.1.13}$$

Let us ask under what kind of rotations the length squared of a 4-vector is invariant. The rotated 4-vector can be written as

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} , \tag{C.1.14}$$

where Λ is a 4×4 space-time rotation matrix. Similarly, we obtain

$$x'_{\mu} = g_{\mu\nu} x'^{\nu} = g_{\mu\nu} \Lambda^{\nu}{}_{\rho} x^{\rho} . \tag{C.1.15}$$

The length squared of x'^{μ} is then given by

$$s'^2 = x'_{\mu} x'^{\mu} = g_{\mu\nu} \Lambda^{\nu}{}_{\rho} x^{\rho} \Lambda^{\mu}{}_{\sigma} x^{\sigma} . \tag{C.1.16}$$

It is invariant under space-time rotations, i.e. $s'^2 = s^2$, only if

$$g_{\mu\nu} \Lambda^{\nu}{}_{\rho} \Lambda^{\mu}{}_{\sigma} = g_{\rho\sigma} . \tag{C.1.17}$$

This can be rewritten as

$$\Lambda^{T}{}_{\sigma}^{\ \mu} g_{\mu\nu} \Lambda^{\nu}{}_{\rho} = g^{T}_{\sigma\rho} , \tag{C.1.18}$$

or equivalently as the matrix multiplication

$$\Lambda^T g \Lambda = g^T = g . \tag{C.1.19}$$

This condition is the Minkowski space-time analog of the Euclidean space condition $\Omega^T\Omega = \mathbf{1}$ for orthogonal spatial rotations. One now obtains

$$x'_\mu = g_{\mu\nu}\Lambda^\nu{}_\rho x^\rho = g_{\mu\nu}\Lambda^\nu{}_\rho g^{\rho\sigma}x_\sigma = [g\Lambda g^{-1}]_\mu{}^\sigma x_\sigma = x_\sigma [g\Lambda g^{-1}]^{T\sigma}{}_\mu = x_\sigma \Lambda^{-1\sigma}{}_\mu \ . \quad \text{(C.1.20)}$$

Here we have used eq.(C.1.19) which leads to

$$[g\Lambda g^{-1}]^T = g^{-1}\Lambda^T g = \Lambda^{-1} \ . \quad \text{(C.1.21)}$$

Finally, as a consequence of eq.(C.1.20) we obtain

$$x_\nu = x'_\mu \Lambda^\mu{}_\nu \ . \quad \text{(C.1.22)}$$

Space-time rotations which obey eq.(C.1.19) are known as Lorentz transformations.

The distance squared in space-time between two 4-vectors x^μ_a and x^μ_b is given by

$$(\Delta s)^2 = (x^0_a - x^0_b)^2 - (x^1_a - x^1_b)^2 - (x^2_a - x^2_b)^2 - (x^3_a - x^3_b)^2 \ , \quad \text{(C.1.23)}$$

and may again be negative. This distance is invariant under both Lorentz transformations Λ and space-time translations d^μ, i.e. $(\Delta s')^2 = (\Delta s)^2$, with

$$x'^\mu_a = \Lambda^\mu{}_\nu x^\nu_a + d^\mu \ , \quad x'^\mu_b = \Lambda^\mu{}_\nu x^\nu_b + d^\mu \ . \quad \text{(C.1.24)}$$

Lorentz transformations and space-time translations again form a group — the Poincaré group — which contains the Lorentz group as a subgroup.

C.2 Gradient as a 4-Vector and d'Alembert Operator

In order to do field theory in manifestly Lorentz-invariant form, we also need to combine temporal and spatial derivatives to a 4-vector. Let us introduce

$$\partial^\mu = (\partial^0, \partial^1, \partial^2, \partial^3) = \left(\frac{\partial}{\partial x_0}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right) \ . \quad \text{(C.2.1)}$$

How does this object transform under Lorentz transformations? Using eq.(C.1.22) one obtains

$$\partial^{\mu'} = \frac{\partial}{\partial x'_\mu} = \frac{\partial x_\nu}{\partial x'_\mu}\frac{\partial}{\partial x_\nu} = \Lambda^\mu{}_\nu \partial^\nu \ . \quad \text{(C.2.2)}$$

This shows that ∂^μ indeed transforms as a contra-variant 4-vector. Similarly, one can define the co-variant 4-vector

$$\partial_\mu = (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) \ . \quad \text{(C.2.3)}$$

One can form the scalar product of the co- and contra-variant derivative 4-vectors. In this way one obtains a second derivative operator which transforms as a space-time scalar, i.e. it is invariant under Lorentz transformations. This Minkowski space-time analog of the Laplace operator in Euclidean space is known as the d'Alembert operator and is given by

$$\Box = \partial_\mu \partial^\mu = \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2} \ . \tag{C.2.4}$$

Appendix D

Particle in an Electromagnetic Field

In this appendix we consider the quantum mechanics of a charged particle (e.g. an electron) in a general classical external electromagnetic field. In principle, the electromagnetic field itself should also be treated quantum mechanically. This is indeed possible and naturally leads to quantum electrodynamics (QED). QED is a relativistic quantum field theory — a subject beyond the scope of this course. Here we will limit ourselves to classical electrodynamics. Hence, we will only treat the charged particle moving in the external field (but not the field itself) quantum mechanically.

D.1 The Classical Electromagnetic Field

Let us investigate the quantum mechanical motion of a charged particle in a general classical external electromagnetic field. For this purpose, we remind ourselves of Maxwell's equations

$$\begin{aligned}
&\vec{\nabla} \cdot \vec{E}(\vec{x},t) = 4\pi\rho(\vec{x},t), \\
&\vec{\nabla} \times \vec{E}(\vec{x},t) + \frac{1}{c}\partial_t \vec{B}(\vec{x},t) = 0, \\
&\vec{\nabla} \cdot \vec{B}(\vec{x},t) = 0, \\
&\vec{\nabla} \times \vec{B}(\vec{x},t) - \frac{1}{c}\partial_t \vec{E}(\vec{x},t) = \frac{4\pi}{c}\vec{j}(\vec{x},t).
\end{aligned} \tag{D.1.1}$$

Adding the time-derivative of the first and c times the divergence of the last equation one obtains the continuity equation

$$\partial_t \rho(\vec{x},t) + \vec{\nabla}\cdot\vec{j}(\vec{x},t) = 0, \tag{D.1.2}$$

which guarantees charge conservation.

The electromagnetic fields $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ can be expressed in terms of scalar and vector potentials $\Phi(\vec{x},t)$ and $\vec{A}(\vec{x},t)$ as

$$\begin{aligned}\vec{E}(\vec{x},t) &= -\vec{\nabla}\Phi(\vec{x},t) - \frac{1}{c}\partial_t\vec{A}(\vec{x},t),\\ \vec{B}(\vec{x},t) &= \vec{\nabla}\times\vec{A}(\vec{x},t).\end{aligned} \tag{D.1.3}$$

Then the homogeneous Maxwell equations

$$\begin{aligned}&\vec{\nabla}\times\vec{E}(\vec{x},t) + \frac{1}{c}\partial_t\vec{B}(\vec{x},t)\\ &= -\vec{\nabla}\times\vec{\nabla}\cdot\Phi(\vec{x},t) - \frac{1}{c}\vec{\nabla}\times\partial_t\vec{A}(\vec{x},t) + \frac{1}{c}\partial_t\vec{\nabla}\times\vec{A}(\vec{x},t) = 0,\\ &\vec{\nabla}\cdot\vec{B}(\vec{x},t) = \vec{\nabla}\cdot\vec{\nabla}\times\vec{A}(\vec{x},t) = 0,\end{aligned} \tag{D.1.4}$$

are automatically satisfied. The inhomogeneous equations can be viewed as four equations for the four unknown functions $\Phi(\vec{x},t)$ and $\vec{A}(\vec{x},t)$.

All fundamental forces in Nature are described by gauge theories. This includes the electromagnetic, weak, and strong forces and even gravity. Gauge theories have a high degree of symmetry. In particular, their classical equations of motion (such as the Maxwell equations in the case of electrodynamics) are invariant against local space-time dependent gauge transformations. In electrodynamics a gauge transformation takes the form

$$\begin{aligned}\Phi(\vec{x},t)' &= \Phi(\vec{x},t) + \frac{1}{c}\partial_t\varphi(\vec{x},t),\\ \vec{A}(\vec{x},t)' &= \vec{A}(\vec{x},t) - \vec{\nabla}\varphi(\vec{x},t).\end{aligned} \tag{D.1.5}$$

Under this transformation the electromagnetic fields

$$\begin{aligned}\vec{E}(\vec{x},t)' &= -\vec{\nabla}\Phi(\vec{x},t)' - \frac{1}{c}\partial_t\vec{A}(\vec{x},t)' = -\vec{\nabla}\Phi(\vec{x},t) - \frac{1}{c}\partial_t\vec{A}(\vec{x},t)\\ &- \frac{1}{c}\vec{\nabla}\partial_t\varphi(\vec{x},t) + \frac{1}{c}\partial_t\vec{\nabla}\varphi(\vec{x},t) = \vec{E}(\vec{x},t),\\ \vec{B}(\vec{x},t)' &= \vec{\nabla}\times\vec{A}(\vec{x},t)' = \vec{\nabla}\times\vec{A}(\vec{x},t) - \vec{\nabla}\times\vec{\nabla}\varphi(\vec{x},t) = \vec{B}(\vec{x},t),\end{aligned} \tag{D.1.6}$$

remain unchanged — they are gauge invariant. As a consequence, Maxwell's equations themselves are gauge invariant as well. In fact, in a gauge theory only gauge invariant quantities have a physical meaning. The scalar and vector potentials $\Phi(\vec{x},t)$ and $\vec{A}(\vec{x},t)$ vary under gauge transformations and are not physically observable. Instead they are mathematical objects with an inherent unphysical gauge ambiguity. Instead, as gauge invariant quantities, the electromagnetic fields $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ are physically observable.

D.2 Classical Charged Particle in an External Electromagnetic Field

The motion of a point particle is governed by Newton's equation

$$m\vec{a}(t) = \vec{F}(t). \tag{D.2.1}$$

For a particle with charge $-e$ moving in an external electromagnetic field the force is given by

$$\vec{F}(t) = -e[\vec{E}(\vec{x}(t),t) + \frac{\vec{v}(t)}{c} \times \vec{B}(\vec{x}(t),t)]. \tag{D.2.2}$$

Newton's equation can be derived from the action

$$S[\vec{x}(t)] = \int dt\ \frac{m}{2}\vec{v}(t)^2 - \int dt d^3y\ [\rho(\vec{y},t)\Phi(\vec{y},t) - \vec{j}(\vec{y},t)\cdot\frac{1}{c}\vec{A}(\vec{y},t)], \tag{D.2.3}$$

where

$$\begin{aligned}\rho(\vec{y},t) &= -e\delta(\vec{y}-\vec{x}(t)),\\ \vec{j}(\vec{y},t) &= -e\vec{v}(t)\delta(\vec{y}-\vec{x}(t)),\end{aligned} \tag{D.2.4}$$

are the charge and current densities of the charged particle at position $\vec{x}(t)$. It is easy to show that charge is conserved, i.e.

$$\partial_t\rho(\vec{y},t) + \vec{\nabla}\cdot\vec{j}(\vec{y},t) = 0. \tag{D.2.5}$$

Inserting eq.(D.2.4) into eq.(D.2.3), for the action one obtains

$$S[\vec{x}(t)] = \int dt\ [\frac{m}{2}\vec{v}(t)^2 + e\Phi(\vec{x}(t),t) - e\frac{\vec{v}(t)}{c}\cdot\vec{A}(\vec{x}(t),t)]. \tag{D.2.6}$$

This action is indeed invariant under gauge transformations because

$$\begin{aligned}&\int dt\ [\Phi(\vec{x}(t),t)' - \frac{\vec{v}(t)}{c}\cdot\vec{A}(\vec{x}(t),t)'] =\\ &\int dt\ [\Phi(\vec{x}(t),t) + \frac{1}{c}\partial_t\varphi(\vec{x}(t),t) - \frac{\vec{v}(t)}{c}\cdot(\vec{A}(\vec{x}(t),t) - \vec{\nabla}\varphi(\vec{x}(t),t))] =\\ &\int dt\ [\Phi(\vec{x}(t),t) - \frac{\vec{v}(t)}{c}\cdot\vec{A}(\vec{x}(t),t) + \frac{1}{c}\frac{d}{dt}\varphi(\vec{x}(t),t)],\end{aligned} \tag{D.2.7}$$

and because the total derivative

$$\frac{d}{dt}\varphi(\vec{x}(t),t) = \partial_t\varphi(\vec{x}(t),t) + \vec{v}\cdot\vec{\nabla}\varphi(\vec{x}(t),t), \tag{D.2.8}$$

integrates to zero as long as $\varphi(\vec{x}(t),t)$ vanishes in the infinite past and future. Identifying the Lagrange function

$$L = \frac{m}{2}\vec{v}(t)^2 + e\Phi(\vec{x}(t),t) - e\frac{\vec{v}(t)}{c}\cdot\vec{A}(\vec{x}(t),t), \tag{D.2.9}$$

it is straightforward to derive Newton's equation as the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\delta L}{\delta v_i(t)} - \frac{\delta L}{\delta x_i} = 0. \tag{D.2.10}$$

The theory can also be formulated in terms of a classical Hamilton function

$$H = \vec{p}(t)\cdot\vec{v}(t) - L, \tag{D.2.11}$$

where $\vec{p}$ is the momentum canonically conjugate to the coordinate $\vec{x}$. One finds

$$m\vec{v}(t) = \vec{p}(t) + \frac{e}{c}\vec{A}(\vec{x}(t),t), \tag{D.2.12}$$

and thus one obtains

$$H = \frac{1}{2m}[\vec{p}(t) + \frac{e}{c}\vec{A}(\vec{x}(t),t)]^2 - e\Phi(\vec{x}(t),t). \tag{D.2.13}$$

This is indeed consistent because

$$v_i(t) = \frac{dx_i(t)}{dt} = \frac{\partial H}{\partial p_i(t)} = \frac{1}{m}[p_i(t) + \frac{e}{c}A_i(\vec{x}(t),t)]. \tag{D.2.14}$$

The other equation of motion is

$$\frac{dp_i(t)}{dt} = -\frac{\partial H}{\partial x_i(t)} = -\frac{e}{mc}[p_j(t)+\frac{e}{c}A_j(\vec{x}(t),t)]\partial_i A_j(\vec{x}(t),t)+e\partial_i\Phi(\vec{x}(t),t). \tag{D.2.15}$$

It is straightforward to show that these equations of motion are again equivalent to Newton's equation.

D.3 Gauge Invariant Form of the Schrödinger Equation

Remarkably, the gauge invariance of electrodynamics is intimately related to the phase ambiguity of the quantum mechanical wave function. As we have seen earlier, the Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{x},t) = -\frac{\hbar^2}{2m}\Delta\Psi(\vec{x},t) + V(\vec{x})\Psi(\vec{x},t), \tag{D.3.1}$$

determines the wave function only up to a global phase ambiguity

$$\Psi(\vec{x},t)' = \Psi(\vec{x},t)\exp(i\phi). \tag{D.3.2}$$

Here ϕ is a constant, independent of space and time.

We now apply the gauge principle to the Schrödinger equation, i.e. we demand that the physics is invariant even under local transformations

$$\Psi(\vec{x},t)' = \Psi(\vec{x},t)\exp(i\phi(\vec{x},t)), \tag{D.3.3}$$

with a space-time dependent phase $\phi(\vec{x},t)$. Of course, if the wave function $\Psi(\vec{x},t)$ solves the original Schrödinger equation (D.3.1), the wave function $\Psi(\vec{x},t)'$ of eq.(D.3.3) in general does not. This is easy to see because

$$\partial_t\Psi(\vec{x},t)' = [\partial_t\Psi(\vec{x},t) + i\Psi(\vec{x},t)\partial_t\phi(\vec{x},t)]\exp(i\phi(\vec{x},t)), \tag{D.3.4}$$

contains the second term on the right hand side that was not present in the original Schrödinger equation. However, if the potential energy $V(\vec{x})$ is replaced by a scalar potential $-e\Phi(\vec{x},t)$, the Schrödinger equation takes the form

$$i\hbar D_t\Psi(\vec{x},t) = -\frac{\hbar^2}{2m}\Delta\Psi(\vec{x},t), \tag{D.3.5}$$

with the covariant derivative

$$D_t\Psi(\vec{x},t) = \partial_t\Psi(\vec{x},t) - i\frac{e}{\hbar}\Phi(\vec{x},t)\Psi(\vec{x},t). \tag{D.3.6}$$

Using the gauge transformation property

$$\Phi(\vec{x},t)' = \Phi(\vec{x},t) + \frac{1}{c}\partial_t\varphi(\vec{x},t), \tag{D.3.7}$$

of the electromagnetic scalar potential, one obtains

$$\begin{aligned} D_t\Psi(\vec{x},t)' &= \partial_t\Psi(\vec{x},t)' - i\frac{e}{\hbar}\Phi(\vec{x},t)'\Psi(\vec{x},t)' \\ &= [\partial_t\Psi(\vec{x},t) + i\Psi(\vec{x},t)\partial_t\phi(\vec{x},t)]\exp(i\phi(\vec{x},t)) \\ &- i\frac{e}{\hbar}[\Phi(\vec{x},t) + \frac{1}{c}\partial_t\varphi(\vec{x},t)]\Psi(\vec{x},t)\exp(i\phi(\vec{x},t)) \\ &= D_t\Psi(\vec{x},t)\exp(i\phi(\vec{x},t)), \end{aligned} \tag{D.3.8}$$

provided that we identify

$$\phi(\vec{x},t) = \frac{e}{\hbar c}\varphi(\vec{x},t). \tag{D.3.9}$$

We also introduce a space-like covariant derivative

$$\vec{D}\Psi(\vec{x},t) = \vec{\nabla}\Psi(\vec{x},t) + i\frac{e}{\hbar c}\vec{A}(\vec{x},t)\Psi(\vec{x},t), \tag{D.3.10}$$

which also transforms as

$$\begin{aligned}\vec{D}\Psi(\vec{x},t)' &= \vec{\nabla}\Psi(\vec{x},t)' + i\frac{e}{\hbar c}\vec{A}(\vec{x},t)'\Psi(\vec{x},t)' \\ &= [\vec{\nabla}\Psi(\vec{x},t) + i\Psi(\vec{x},t)\vec{\nabla}\phi(\vec{x},t)]\exp(i\phi(\vec{x},t)) \\ &+ i\frac{e}{\hbar c}[\vec{A}(\vec{x},t) - \vec{\nabla}\varphi(\vec{x},t)]\Psi(\vec{x},t)\exp(i\phi(\vec{x},t)) \\ &= \vec{D}\Psi(\vec{x},t)\exp(i\phi(\vec{x},t)), \end{aligned} \tag{D.3.11}$$

under a gauge transformation. Using $\vec{p} = (\hbar/i)\vec{\nabla}$ one obtains

$$\frac{\hbar}{i}\vec{D} = \vec{p} + \frac{e}{c}\vec{A}(\vec{x},t), \tag{D.3.12}$$

which is the quantum version of $m\vec{v}$ from eq.(D.2.12) that we encountered in the classical theory.

Finally, in the Schrödinger equation we replace $\Delta\Psi(\vec{x},t) = \vec{\nabla}\cdot\vec{\nabla}\Psi(\vec{x},t)$ with $\vec{D}\cdot\vec{D}\Psi(\vec{x},t)$ and we obtain

$$i\hbar D_t\Psi(\vec{x},t) = -\frac{\hbar^2}{2m}\vec{D}\cdot\vec{D}\Psi(\vec{x},t). \tag{D.3.13}$$

Inserting the explicit form of the covariant derivatives, the Schrödinger equation for a charged particle in an arbitrary external electromagnetic field takes the form

$$i\hbar[\partial_t - i\frac{e}{\hbar}\Phi(\vec{x},t)]\Psi(\vec{x},t) = -\frac{\hbar^2}{2m}[\vec{\nabla} + i\frac{e}{\hbar c}\vec{A}(\vec{x},t)]\cdot[\vec{\nabla} + i\frac{e}{\hbar c}\vec{A}(\vec{x},t)]\Psi(\vec{x},t). \tag{D.3.14}$$

This equation is invariant under gauge transformations of the form

$$\begin{aligned}\Phi(\vec{x},t)' &= \Phi(\vec{x},t) + \frac{1}{c}\partial_t\varphi(\vec{x},t), \\ \vec{A}(\vec{x},t)' &= \vec{A}(\vec{x},t) - \vec{\nabla}\varphi(\vec{x},t), \\ \Psi(\vec{x},t)' &= \Psi(\vec{x},t)\exp(i\frac{e}{\hbar c}\varphi(\vec{x},t)). \end{aligned} \tag{D.3.15}$$

Under this transformation, both sides of the Schrödinger equation change by a factor $\exp(i(e/\hbar c)\varphi(\vec{x},t))$. Canceling this factor out, the equation remains invariant.

As usual, the wave function $\Psi(\vec{x},t)$ that solves the gauged Schrödinger equation (D.3.14) can be interpreted as the probability amplitude for finding the particle at position $\vec{x}$ at time t. In particular, the probability density

$$\rho(\vec{x},t) = |\Psi(\vec{x},t)|^2, \tag{D.3.16}$$

is gauge invariant and hence physically meaningful. Again, probability conservation follows from a continuity equation

$$\partial_t\rho(\vec{x},t) + \vec{\nabla}\cdot\vec{j}(\vec{x},t) = 0. \tag{D.3.17}$$

However, in the presence of electromagnetic fields the usual probability current must be modified by replacing ordinary with covariant derivatives such that now

$$\vec{j}(\vec{x},t) = \frac{\hbar}{2mi}[\Psi(\vec{x},t)^* \vec{D}\Psi(\vec{x},t) - (\vec{D}\Psi(\vec{x},t))^* \Psi(\vec{x},t)]. \tag{D.3.18}$$

Appendix E

Relativistic Formulation of Classical Electrodynamics

The relativistic nature of Maxwell's equations is not manifest in their original form. In this Appendix, we formulate electrodynamics such that its invariance under Lorentz transformations — i.e. under rotations in Minkowski space-time — becomes manifest. In order to better see the connection with the original Maxwell equations, in this appendix we do not put $c = 1$.

E.1 Current and Vector Potential

In order to express electrodynamics in manifestly Lorentz co-variant form, we proceed step by step and begin with the charge and current densities $\rho(\vec{x}, t)$ and $\vec{j}(\vec{x}, t)$. The corresponding continuity equation which expresses charge conservation takes the form

$$\partial_t \rho(\vec{x}, t) + \vec{\nabla} \cdot \vec{j}(\vec{x}, t) = 0 \ . \tag{E.1.1}$$

Since temporal and spatial derivatives are combined to a gradient 4-vector, it is natural to combine $\rho(\vec{x}, t)$ and $\vec{j}(\vec{x}, t)$ to the current 4-vector

$$j^\mu(x) = (c\rho(\vec{x}, t), j_x(\vec{x}, t), j_y(\vec{x}, t), j_z(\vec{x}, t)) \ . \tag{E.1.2}$$

Here we have introduced $x = (\vec{x}, t)$ as a short-hand notation for a point in space-time. It goes without saying that this should not be confused with the x-component of the spatial vector $\vec{x}$. In Lorentz-invariant form the continuity equation thus takes the form

$$\partial_\mu j^\mu(x) = \frac{1}{c}\partial_t c\rho(\vec{x}, t) + \partial_x j_x(\vec{x}, t) + \partial_y j_y(\vec{x}, t) + \partial_z j_z(\vec{x}, t) = 0 \ . \tag{E.1.3}$$

Here we have combined the co-variant 4-vector ∂_μ and the contra-variant 4-vector $j^\mu(x)$ to the Lorentz-scalar zero. The Lorentz invariance of the continuity equation implies that charge conservation is valid in any inertial frame, independent of the motion of an observer.

Of course, the charge and current densities themselves are dependent on the reference frame in which they are considered. If a general (non-uniform and non-static) charge and current density is transformed into another reference frame, one must also transform the space-time point x at which the density is evaluated, i.e.

$$j'^\mu(x') = \Lambda^\mu{}_\nu j^\nu(x) = \Lambda^\mu{}_\nu j^\nu(\Lambda^{-1}x') \ . \tag{E.1.4}$$

From the scalar potential $\phi(\vec{x},t)$ and the vector potential $\vec{A}(\vec{x},t)$ one can construct another 4-vector field

$$A^\mu(x) = (\phi(\vec{x},t), A_x(\vec{x},t), A_y(\vec{x},t), A_z(\vec{x},t)) \ . \tag{E.1.5}$$

Under Lorentz transformations it again transforms as

$$A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x') \ . \tag{E.1.6}$$

Scalar and vector potentials transform non-trivially under gauge transformations

$${}^\varphi\phi(\vec{x},t) = \phi(\vec{x},t) + \frac{1}{c}\partial_t\varphi(\vec{x},t) \ , \quad {}^\varphi\vec{A}(\vec{x},t) = \vec{A}(\vec{x},t) - \vec{\nabla}\varphi(\vec{x},t) \ . \tag{E.1.7}$$

In 4-vector notation this relation takes the form

$${}^\varphi A^\mu(x) = A^\mu(x) + \partial^\mu\varphi(x) \ . \tag{E.1.8}$$

Here $\varphi(x)$ is an arbitrary space-time dependent gauge transformation function. This function is a space-time scalar, i.e. under Lorentz transformations it transforms as

$$\varphi'(x') = \varphi(\Lambda^{-1}x') \ . \tag{E.1.9}$$

The Lorentz gauge fixing condition

$$\frac{1}{c}\partial_t\phi(\vec{x},t) + \vec{\nabla}\cdot\vec{A}(\vec{x},t) = 0 \tag{E.1.10}$$

can be re-expressed as

$$\partial_\mu A^\mu(x) = 0 \ . \tag{E.1.11}$$

In the Lorentz gauge the wave equations take the form

$$\frac{1}{c^2}\partial_t^2\phi(\vec{x},t) - \Delta\phi(\vec{x},t) = \rho(\vec{x},t) \ , \quad \frac{1}{c^2}\partial_t^2\vec{A}(\vec{x},t) - \Delta\vec{A}(\vec{x},t) = \frac{1}{c}\vec{j}(\vec{x},t) \ , \tag{E.1.12}$$

which can be combined to

$$\Box A^\mu(x) = \frac{1}{c}j^\mu(x) \ . \tag{E.1.13}$$

E.2 Field Strength Tensor

It may not be entirely obvious how to express the electric and magnetic fields $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ in relativistic form. We need to use

$$\vec{E}(\vec{x},t) = -\vec{\nabla}\phi(\vec{x},t) - \frac{1}{c}\partial_t \vec{A}(\vec{x},t) \ , \quad \vec{B}(\vec{x},t) = \vec{\nabla} \times \vec{A}(\vec{x},t) \ . \tag{E.2.1}$$

Obviously, $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ are constructed from the 4-vectors ∂^μ and $A^\mu(x)$. The scalar product of these two 4-vectors

$$\partial_\mu A^\mu(x) = \frac{1}{c}\partial_t \phi(\vec{x},t) + \vec{\nabla} \cdot \vec{A}(\vec{x},t) \tag{E.2.2}$$

appears in the Lorentz gauge fixing condition but does not yield the electric or magnetic field. The 4-vectors ∂^μ and $A^\mu(x)$ can also be combined to the symmetric tensor

$$D^{\mu\nu} = \partial^\mu A^\nu(x) + \partial^\nu A^\mu(x) \ , \tag{E.2.3}$$

as well as to the anti-symmetric tensor

$$F^{\mu\nu} = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x) \ . \tag{E.2.4}$$

Under gauge transformations these tensors transform as

$$\begin{aligned}
{}^{\varphi}D^{\mu\nu} &= \partial^\mu \ {}^{\varphi}A^\nu(x) + \partial^\nu \ {}^{\varphi}A^\mu(x) \\
&= \partial^\mu A^\nu(x) + \partial^\mu\partial^\nu\varphi(x) + \partial^\nu A^\mu(x) + \partial^\nu\partial^\mu\varphi(x) = D^{\mu\nu} + 2\partial^\mu\partial^\nu\varphi(x) \ , \\
{}^{\varphi}F^{\mu\nu} &= \partial^\mu \ {}^{\varphi}A^\nu(x) - \partial^\nu \ {}^{\varphi}A^\mu(x) \\
&= \partial^\mu A^\nu(x) + \partial^\mu\partial^\nu\varphi(x) - \partial^\nu A^\mu(x) - \partial^\nu\partial^\mu\varphi(x) = F^{\mu\nu} \ ,
\end{aligned} \tag{E.2.5}$$

i.e. the anti-symmetric tensor $F^{\mu\nu}$ is gauge invariant, while the symmetric tensor $D^{\mu\nu}$ is not. As a consequence, it does not play any particular rôle in electrodynamics. Since the electromagnetic fields are gauge invariant, we expect them to be related to $F^{\mu\nu}$. Let us consider the various components of this tensor

$$\begin{aligned}
F^{01}(x) &= \partial^0 A^1(x) - \partial^1 A^0(x) = \frac{1}{c}\partial_t A_x(\vec{x},t) + \partial_x\phi(\vec{x},t) = -E_x(\vec{x},t) \ , \\
F^{02}(x) &= \partial^0 A^2(x) - \partial^2 A^0(x) = \frac{1}{c}\partial_t A_y(\vec{x},t) + \partial_y\phi(\vec{x},t) = -E_y(\vec{x},t) \ , \\
F^{03}(x) &= \partial^0 A^3(x) - \partial^3 A^0(x) = \frac{1}{c}\partial_t A_z(\vec{x},t) + \partial_z\phi(\vec{x},t) = -E_z(\vec{x},t) \ , \\
F^{12}(x) &= \partial^1 A^2(x) - \partial^2 A^1(x) = -\partial_x A_y(\vec{x},t) + \partial_y A_x(\vec{x},t) = -B_z(\vec{x},t) \ , \\
F^{23}(x) &= \partial^2 A^3(x) - \partial^3 A^2(x) = -\partial_y A_z(\vec{x},t) + \partial_z A_y(\vec{x},t) = -B_x(\vec{x},t) \ , \\
F^{31}(x) &= \partial^3 A^1(x) - \partial^1 A^3(x) = -\partial_z A_x(\vec{x},t) + \partial_x A_z(\vec{x},t) = -B_y(\vec{x},t) \ .
\end{aligned} \tag{E.2.6}$$

Hence, the anti-symmetric tensor indeed contains the electric and magnetic fields as

$$F^{\mu\nu}(x) = \begin{pmatrix} 0 & -E_x(\vec{x},t) & -E_y(\vec{x},t) & -E_z(\vec{x},t) \\ E_x(\vec{x},t) & 0 & -B_z(\vec{x},t) & B_y(\vec{x},t) \\ E_y(\vec{x},t) & B_z(\vec{x},t) & 0 & -B_x(\vec{x},t) \\ E_z(\vec{x},t) & -B_y(\vec{x},t) & B_x(\vec{x},t) & 0 \end{pmatrix} . \tag{E.2.7}$$

The co-variant components of this tensor are given by

$$F_{\mu\nu}(x) = \begin{pmatrix} 0 & E_x(\vec{x},t) & E_y(\vec{x},t) & E_z(\vec{x},t) \\ -E_x(\vec{x},t) & 0 & -B_z(\vec{x},t) & B_y(\vec{x},t) \\ -E_y(\vec{x},t) & B_z(\vec{x},t) & 0 & -B_x(\vec{x},t) \\ -E_z(\vec{x},t) & -B_y(\vec{x},t) & B_x(\vec{x},t) & 0 \end{pmatrix} . \tag{E.2.8}$$

E.3 Maxwell Equations

Let us first consider the inhomogeneous Maxwell equations

$$\vec{\nabla} \cdot \vec{E}(\vec{x},t) = \rho(\vec{x},t) \ , \quad \vec{\nabla} \times \vec{B}(\vec{x},t) - \frac{1}{c}\partial_t \vec{E}(\vec{x},t) = \frac{1}{c}\vec{j}(\vec{x},t) \ . \tag{E.3.1}$$

These are four equations with the components of the 4-vector current $j^\mu(x)$ on the right-hand side. Hence, on the left-hand side there must also be a 4-vector. The left-hand side consists of derivatives, i.e. of components of the gradient 4-vectors ∂_μ, and of the electromagnetic fields, i.e. of the components of the field strength tensor $F^{\mu\nu}$. Hence, the 4-vector ∂_μ and the tensor $F^{\mu\nu}$ on the left-hand side must be combined to another 4-vector. This can be achieved by forming $\partial_\mu F^{\mu\nu}(x)$ and thus by contracting (i.e. by summing) one co- and one contra-variant index. The various components of this object take the form

$$\begin{aligned}
\partial_\mu F^{\mu 0}(x) &= \partial_x E_x(\vec{x},t) + \partial_y E_y(\vec{x},t) + \partial_z E_z(\vec{x},t) \\
&= \vec{\nabla} \cdot \vec{E}(\vec{x},t) = \rho(\vec{x},t) \ , \\
\partial_\mu F^{\mu 1}(x) &= -\frac{1}{c}\partial_t E_x(\vec{x},t) + \partial_y B_z(\vec{x},t) - \partial_z B_y(\vec{x},t) \\
&= [\vec{\nabla} \times \vec{B}]_x(\vec{x},t) - \frac{1}{c}\partial_t E_x(\vec{x},t) = \frac{1}{c} j_x(\vec{x},t) \ , \\
\partial_\mu F^{\mu 2}(x) &= -\frac{1}{c}\partial_t E_y(\vec{x},t) - \partial_x B_z(\vec{x},t) + \partial_z B_x(\vec{x},t) \\
&= [\vec{\nabla} \times \vec{B}]_y(\vec{x},t) - \frac{1}{c}\partial_t E_y(\vec{x},t) = \frac{1}{c} j_y(\vec{x},t) \ , \\
\partial_\mu F^{\mu 3}(x) &= -\frac{1}{c}\partial_t E_z(\vec{x},t) + \partial_x B_y(\vec{x},t) - \partial_y B_x(\vec{x},t) \\
&= [\vec{\nabla} \times \vec{B}]_z(\vec{x},t) - \frac{1}{c}\partial_t E_z(\vec{x},t) = \frac{1}{c} j_z(\vec{x},t) \ .
\end{aligned} \tag{E.3.2}$$

Indeed, these equations can be summarized as

$$\partial_\mu F^{\mu\nu}(x) = \frac{1}{c} j^\nu(x) \ . \tag{E.3.3}$$

At this point the usefulness of the compact 4-dimensional notation should be obvious. Inserting eq.(E.2.4) into the inhomogeneous Maxwell equations, we obtain

$$\partial_\mu F^{\mu\nu}(x) = \partial_\mu(\partial^\mu A^\mu(x) - \partial^\nu A^\mu(x)) = \Box A^\nu(x) - \partial^\nu \partial_\mu A^\mu(x) = \frac{1}{c} j^\nu(x) \ . \tag{E.3.4}$$

If the 4-vector potential obeys the Lorentz gauge fixing condition $\partial_\mu A^\mu(x) = 0$, this is nothing but the wave equation eq.(E.1.13).

How can we express the homogeneous Maxwell equations

$$\vec{\nabla} \cdot \vec{B}(\vec{x},t) = 0 \ , \ \vec{\nabla} \times \vec{E}(\vec{x},t) + \frac{1}{c}\partial_t \vec{B}(\vec{x},t) = 0 \tag{E.3.5}$$

in 4-dimensional form? Except for the vanishing right-hand side, they look very similar to the inhomogeneous equations. All we need to do is to substitute $\vec{E}(\vec{x},t)$ by $-\vec{B}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ by $\vec{E}(\vec{x},t)$. Such a substitution is known as a duality transformation. Under this operation the field strength tensor turns into the dual tensor

$$\widetilde{F}^{\mu\nu}(x) = \begin{pmatrix} 0 & B_x(\vec{x},t) & B_y(\vec{x},t) & B_z(\vec{x},t) \\ -B_x(\vec{x},t) & 0 & -E_z(\vec{x},t) & E_y(\vec{x},t) \\ -B_y(\vec{x},t) & E_z(\vec{x},t) & 0 & -E_x(\vec{x},t) \\ -B_z(\vec{x},t) & -E_y(\vec{x},t) & E_x(\vec{x},t) & 0 \end{pmatrix} \ , \tag{E.3.6}$$

and the homogeneous Maxwell equations can thus be expressed as

$$\partial_\mu \widetilde{F}^{\mu\nu}(x) = 0 \ . \tag{E.3.7}$$

The co-variant components of the dual field strength tensor take the form

$$\widetilde{F}_{\mu\nu}(x) = \begin{pmatrix} 0 & -B_x(\vec{x},t) & -B_y(\vec{x},t) & -B_z(\vec{x},t) \\ B_x(\vec{x},t) & 0 & -E_z(\vec{x},t) & E_y(\vec{x},t) \\ B_y(\vec{x},t) & E_z(\vec{x},t) & 0 & -E_x(\vec{x},t) \\ B_z(\vec{x},t) & -E_y(\vec{x},t) & E_x(\vec{x},t) & 0 \end{pmatrix} \ . \tag{E.3.8}$$

It is obvious that the field strength tensor $F^{\mu\nu}(x)$ and its dual $\widetilde{F}_{\mu\nu}(x)$ consist of the same components, namely of the electric and magnetic fields $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$. Hence, there must be a relation between the two tensors. This relation takes the form

$$\widetilde{F}_{\mu\nu}(x) = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}(x) \ . \tag{E.3.9}$$

Here $\epsilon_{\mu\nu\rho\sigma}$ is a totally anti-symmetric tensor with components 0 or ± 1. If any of the indices μ, ν, ρ, and σ are equal, the value of $\epsilon_{\mu\nu\rho\sigma}$ is zero. Only if all indices are different, the value of $\epsilon_{\mu\nu\rho\sigma}$ is non-zero. The value is $\epsilon_{\mu\nu\rho\sigma} = 1$ if $\mu\nu\rho\sigma$ is an even permutation of 0123 (i.e. it requires an even number of index pair permutations to turn $\mu\nu\rho\sigma$ into 0123). Similarly, $\epsilon_{\mu\nu\rho\sigma} = -1$ if $\mu\nu\rho\sigma$ is an odd permutation of 0123. For example, we obtain

$$\widetilde{F}_{01}(x) = \frac{1}{2}\epsilon_{01\rho\sigma}F^{\rho\sigma}(x) = \frac{1}{2}\left(\epsilon_{0123}F^{23}(x) - \epsilon_{0132}F^{32}(x)\right) = F^{23}(x) = -B_x(\vec{x},t) \ , \tag{E.3.10}$$

as well as

$$\widetilde{F}_{12}(x) = \frac{1}{2}\epsilon_{12\rho\sigma}F^{\rho\sigma}(x) = \frac{1}{2}\left(\epsilon_{1203}F^{03}(x) - \epsilon_{1230}F^{30}(x)\right) = F^{03}(x) = -E_z(\vec{x},t) \ , \tag{E.3.11}$$

Inserting eq.(E.3.9) into the homogeneous Maxwell equations (E.3.7) one obtains

$$\partial^\mu \widetilde{F}_{\mu\nu}(x) = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\mu F^{\rho\sigma}(x) = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\mu\left(\partial^\rho A^\sigma(x) - \partial^\sigma A^\rho(x)\right) = 0 \ . \tag{E.3.12}$$

Due to the anti-symmetry of $\epsilon_{\mu\nu\rho\sigma}$ and the commutativity of the derivatives ∂^μ and ∂^ρ, this equation is automatically satisfied. This is no surprise, because the original Maxwell equations were also automatically satisfied by the introduction of the scalar and vector potentials $\phi(\vec{x},t)$ and $\vec{A}(\vec{x},t)$. The homogeneous Maxwell equations can alternatively be expressed as

$$\partial^\mu F^{\rho\sigma} + \partial^\rho F^{\sigma\mu} + \partial^\sigma F^{\mu\rho} = 0 \ . \tag{E.3.13}$$

Indeed, multiplying this relation with $\epsilon_{\mu\nu\rho\sigma}$ and applying cyclic permutations to the indices μ, ρ, and σ, one again arrives at eq.(E.3.12).

E.4 Space-Time Scalars from Field Strength Tensors

Which scalar quantities can be formed by combining the field strength tensors $F^{\mu\nu}(x)$ and $\widetilde{F}^{\mu\nu}(x)$? First, we can construct the combination

$$\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) = \frac{1}{2}\left(\vec{B}(\vec{x},t)^2 - \vec{E}(\vec{x},t)^2\right) \ , \tag{E.4.1}$$

which will later turn out to be the Lagrange density of electrodynamics. Then we can construct

$$\frac{1}{4}\widetilde{F}_{\mu\nu}(x)\widetilde{F}^{\mu\nu}(x) = \frac{1}{2}\left(\vec{E}(\vec{x},t)^2 - \vec{B}(\vec{x},t)^2\right) \ , \tag{E.4.2}$$

which is thus the same up to a minus-sign. While the electromagnetic fields themselves are obviously not Lorentz-invariant, the difference of their magnitudes squared is.

One can also mix the field strength tensor with its dual and one then obtains

$$\frac{1}{4}F_{\mu\nu}(x)\widetilde{F}^{\mu\nu}(x) = \vec{E}(\vec{x},t)\cdot\vec{B}(\vec{x},t)\ . \tag{E.4.3}$$

Similarly, one obtains

$$\frac{1}{4}\widetilde{F}_{\mu\nu}(x)F^{\mu\nu}(x) = \vec{E}(\vec{x},t)\cdot\vec{B}(\vec{x},t)\ , \tag{E.4.4}$$

which is thus equivalent. Interestingly, the projection of the electric on the magnetic field $\vec{E}(\vec{x},t)\cdot\vec{B}(\vec{x},t)$ is also Lorentz invariant, i.e. it has the same value in all inertial frames.

E.5 Transformation of Electromagnetic Fields

Since they form the components of the field strength tensors, it is obvious that the electromagnetic fields $\vec{E}(\vec{x},t)$ and $\vec{B}(\vec{x},t)$ are not Lorentz-invariant, i.e. they depend on the motion of the observer. Under a Lorentz transformation the field strength tensor transforms as

$$F^{\mu\nu\prime}(x') = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}F^{\rho\sigma}(x) = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}F^{\rho\sigma}(\Lambda^{-1}x')\ . \tag{E.5.1}$$

This can be rewritten in matrix form as

$$F^{\mu\nu\prime}(x') = \Lambda^{\mu}{}_{\rho}F^{\rho\sigma}(\Lambda^{-1}x')\Lambda^{T}{}_{\sigma}^{\ \nu} \ \Rightarrow\ F'(x') = \Lambda F(\Lambda^{-1}x')\Lambda^{T}\ . \tag{E.5.2}$$

E.6 Action and Euler-Lagrange Equation

The Lagrangian for the electromagnetic field interacting with a charge and current distribution j^{μ} is given by

$$\mathcal{L}(A_\nu,\partial_\mu A_\nu) = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}A_\nu j^\nu\ . \tag{E.6.1}$$

The corresponding action is obtained by integrating the Lagrangian over space-time, i.e.

$$S[A] = \int dt\ d^3x\ \mathcal{L} = \int d^4x\ \frac{1}{c}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}A_\nu j^\nu\right)\ . \tag{E.6.2}$$

The action can be viewed as a functional (i.e. a function of a function) of the electromagnetic 4-vector potential A^μ. The Euler-Lagrange equation of motion resulting from the principle of least action now takes the form

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta \partial_\mu A_\nu} - \frac{\delta \mathcal{L}}{\delta A_\nu} = \partial_\mu F^{\mu\nu} - \frac{1}{c} j^\nu = 0 \ . \tag{E.6.3}$$

Indeed, this yields just the inhomogeneous Maxwell equations

$$\partial_\mu F^{\mu\nu} = \frac{1}{c} j^\nu \ . \tag{E.6.4}$$

The homogeneous Maxwell equations are automatically satisfied as a consequence of $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

E.7 Energy-Momentum Tensor

Let us consider the energy-momentum tensor of the free electromagnetic field (i.e. in the absence of charges and currents)

$$T^{\mu\nu} = -F^{\mu\rho} F^{\nu}_{\ \rho} - \mathcal{L} g^{\mu\nu} \ , \tag{E.7.1}$$

which obeys the continuity equation

$$\partial_\mu T^{\mu\nu} = 0 \ . \tag{E.7.2}$$

The time-time component of the energy-momentum tensor is given by

$$T^{00} = -F^{0\rho} F^{0}_{\ \rho} + \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} g^{00} = \vec{E}^2 + \frac{1}{2}(\vec{B}^2 - \vec{E}^2) = \frac{1}{2}(\vec{E}^2 + \vec{B}^2) \ . \tag{E.7.3}$$

which is the energy density of the electromagnetic field. The space-time components of the energy-momentum tensor take the form

$$T^{i0} = -F^{i\rho} F^{0}_{\ \rho} + \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} g^{i0} = \epsilon_{ijk} E_j B_k = (\vec{E} \times \vec{B})_i \ . \tag{E.7.4}$$

Here ϵ_{ijk} is the completely anti-symmetric Levi-Civita tensor and $\vec{E} \times \vec{B}$ is the Poynting vector which is known to represent the momentum density of the electromagnetic field. The continuity equation that represents energy conservation takes the form

$$\partial_\mu T^{\mu 0} = \partial_0 T^{00} + \partial_i T^{i0} = 0 \ . \tag{E.7.5}$$

Appendix F

Why Physics Exists

In the following, I present a few thoughts that may shed some light on the question why physics exists. People seem not to talk about this very much. All I'll have to say is therefore wide open for criticism and far from being well established. While this appendix also addresses some questions at the interface of physics and philosophy, it does not comply with the conceptual rigor of the latter discipline. For example, in the text below the notion of "existence" is not defined or clarified further in a philosophical context. The reader is assumed to be familiar with some notions of modern physics, such as, for example, the concept of an effective field theory.

F.1 Structures and Physicists

Physics is the man-made observation and experimental investigation as well as the theoretical mathematical description of Nature. As such, the existence of physics requires the existence of physicists, i.e. the existence of curious minds and capable brains. Of course, it is equally important that there is something "out there" to be discovered, i.e. that Nature does contain structures that can be adequately described in mathematical terms. Both such structures as well as curious minds and capable brains obviously do exist in our Universe. Of course, we can easily imagine much simpler hypothetical "Universes" that follow some strict mathematical laws of Nature. For example, we can imagine a "Universe" consisting of just two Newtonian point particles, to be called Sun and Earth, orbiting their center of mass in elliptical orbits for the rest of time. While such a simple system can easily be understood from outside, it is obvious that it cannot understand itself from within. Remarkably, by evolving the human subspecies of the experimental and theoretical physicist, in some sense our Universe is "thinking" about itself. This requires an

enormous degree of complexity.

F.2 Space, Time, and Separable Entities

In order to understand something from within, one must still be able to distant oneself from the object of study, at least to some extent. It is the existence of space and time that allows us to do just that. While we currently do not understand the origin of space and time, the related puzzles lead to the deepest questions we can currently ask in physics. After all, space and time are fundamental ordering principles, deeply rooted in physical reality, that allow us to separate ourselves from other more or less independent entities. In this sense, space and time are necessary prerequisites for doing physics from within the system. It should be pointed out that we can easily imagine a mathematical "Universe" without invoking the concept of time. For example, equilibrium statistical mechanics follows strict mathematical rules, but knows nothing about time. Indeed, a "Universe" following just the rules of equilibrium statistical mechanics would have a hard "time" understanding even part of itself from within.

F.3 Space-Time Locality

Along with the concepts of space and time comes the concept of space-time locality. Indeed without locality, space and time would loose their meaning as the most basic ordering principles, and would therefore cease to be useful concepts. In our hypothetical "Universe" consisting of just Sun and Earth interacting via instantaneous Newtonian gravity, there is locality in time but not in space. As a consequence, space is not even a very useful concept in a hypothetical world like this. The coordinates of the two particles as well as their distance are completely sufficient. Thinking about the particles as existing in an otherwise empty space seems natural to us, who live in a much more dynamical space. However, in the very restricted two-particle "Universe" the whole concept of space is to a large extent an unnecessary luxury. The situation changes drastically when we endow the two point particles with Einstein gravity (i.e. general relativity). Then there is a metric attached to each point in space-time, which becomes a dynamical entity governed by local laws of Nature.

F.4 Hierarchies of Scales

The existence of curious minds and capable brains is likely to require the existence of vastly different length scales. If every separable entity would be of a similar size, for example, if everything including the potential physicist would exist at atomic scales, it is unlikely that physics would ever have taken off. This is because, in that case, one must identify the correct "Theory of Everything" (TOE) before physics can get started. In a hypothetical "Universe" consisting only of atomic scale entities governed by the rules of quantum mechanics, even if something like a "brain" would exist, it would almost certainly not be capable enough to discover quantum mechanics (which we may think of as the TOE in that world), in one gigantic strike of genius.

F.5 Anthropic Arguments

Although we don't know why they exist in our Universe, we have identified the existence of mathematical structures including space, time, and locality, as well as hierarchies of vastly different length scales as necessary prerequisites for the existence of physics. Today it is all too popular to invoke the anthropic principle. Still, the above mentioned prerequisites for the existence of physics may very well be necessary conditions for the existence of curious minds and capable brains as well. Hence, we can argue that potential physicists (i.e. curious minds and capable brains) can only exist in a world in which the basis for doing physics also automatically exists. Invoking the anthropic principle, i.e. using the fact of our own existence, we can then "explain" why the basis for physics exists as well.

F.6 Effective Field Theory

Once we have mathematical structures, space-time locality, hierarchies of different length scales, as well as curious minds and capable brains, we are suddenly in business for doing physics. In particular, thanks to space-time locality and the existence of hierarchies, we are free to do physics at some length scale, without the need to discover the "Theory of Everything" before we can even get started. Identifying the relevant degrees of freedom at some scale, and constructing the most general local dynamics while respecting the relevant symmetries is the technique of systematic effective field theory. Being able to understand aspects of the world step by step in scale, is a major reason why physics is so successful. An effective field theory

always contains some a priori unknown physical parameters, whose values can be fixed, for example, by comparison with experiments. Eventually, one may reach a deeper level of understanding by matching these parameters to a more fundamental effective theory valid at shorter length scales. In this way, by matching effective field theories, we may patch together a "quilt" of mathematical descriptions of Nature that may eventually cover the entire landscape of physical phenomena.

F.7 Renormalization Group, Universality, and the Benefit of 4-d

The success of the effective field theory method relies on Kenneth Wilson's renormalization group and the related concept of universality. Universality implies that the physics at long distances or low energies is insensitive to the details of what happens at much higher energy scales. In the process of renormalization, the dynamics is attracted to a renormalization group fixed point that is characteristic of a universality class. Theories which may differ substantially at high energies may still flow to the same fixed point at low energies. This implies that we need not know all about physics (i.e. the TOE) before we can start to explore the low-energy regime. As a flip-side of the same coin, it is practically impossible to infer the correct TOE from low-energy data. In view of universality, doing successful fundamental physics is reduced to identifying the correct renormalization group fixed point, which makes life a lot easier. This is also thanks to the dimension of space-time. In particular, in the 4-dimensional space-time we live in, there are relatively few non-trivial fixed points of the renormalization group, namely those related to non-Abelian gauge theories. Indeed, it is non-Abelian gauge fields that dominate the physics of the Standard Model of particle physics — our most fundamental theory today. In a 2-d space-time, for example, the situation would be very different. In 2-d, there is an enormous number of non-trivial fixed points, such that identifying the correct one would be much more difficult than in 4-d. As Martin Lüscher from CERN once said: in 2 space-time dimensions, physics would be as complicated as politics. Thanks to our 4-d space-time, as complicated as physics may seem, it does, in fact, exist because it is not too difficult (for the more capable brains).

F.8 The Theory of Everything

By moving towards ever decreasing length scales (and thus ever increasing energy scales), we may (or may not) eventually discover the true "Theory of Everything".

Already the Standard Model of particle physics is based on fundamental objects such as quarks or W-bosons, which are precise embodiments of abstract mathematical concepts. Even if physicists may eventually discover the ultimate TOE, which would obviously be a tremendous achievement, this would in no way obviate the need for effective field theories covering the low-energy domain. Just as the Standard Model is extremely powerful at particle physics energy scales, but pretty useless for understanding the complex dynamics of condensed matter, the TOE would be really useful only at the shortest distance scales. In this sense, the "Theory of Everything" may very well be a "Theory of Nothing" relevant at presently accessible energy scales. Furthermore, just as Bertrand Russell's Principia Mathematica, a "Theory of Everything" may suffer from Gödelian incompletenesses. We may have to be content with consistency, instead of urging for completeness.

F.9 Model Building

When doing physics, it is often useful to build models. Our model "Universe" consisting of just Sun and Earth is a good example. In this way, by mentally or even experimentally isolating a small part of the world from external influences, we have a chance to completely understand it. In some sense, Newton's classical mechanics is the theory of everything that is important to understand the dynamics of slowly moving macroscopic objects under the influence of gravity. Similarly, the Hubbard model describes the "world" inside doped antiferromagnets. Experts argue whether or not the Hubbard model is the theory of everything that is necessary to understand the origin of high-temperature superconductivity. Quantum Chromodynamics (QCD) is far more than just a model in this sense. In fact, it is an integral part of our most fundamental description of Nature — the Standard Model of particle physics. While (due to its "triviality" in the renormalization group sense) the Standard Model is necessarily "just" an effective theory, thanks to its asymptotic freedom, QCD could hold at arbitrarily high energy scales. In this sense, it may be considered as the "Theory of Everything" about the strong interaction. Isolated "worlds" like the Hubbard model or QCD are still sufficiently complicated that we cannot understand them completely analytically. In that case, effective field theory again plays an important role. The low-energy regimes of both the Hubbard model at low doping and QCD at low baryon density can be described by systematic low-energy effective field theories. Although they cannot be derived rigorously from the underlying microscopic theories, the physical consequences of emergent phenomena like antiferromagnetism or chiral symmetry breaking can then be addressed quantitatively in the low-energy effective theory. The matching between the underlying short-distance and the emergent long-distance theory works, because it connects two local theories which are mathematically formulated using the basic concepts of

space and time.

F.10 Numerical Simulation

To accurately investigate model "worlds" such as the Hubbard model or (lattice) QCD, numerical simulations play an important role. For example, the low-energy parameters of the corresponding effective theories can be derived numerically from the underlying microscopic models. Still, in several interesting cases accurate numerical simulations are prevented by severe sign or complex action problems. For example, the exploration of high-temperature superconductivity in the Hubbard model is hindered by a severe fermion sign problem, and the exploration of the "condensed matter physics of QCD" (as Krishna Rajagopal and Frank Wilczek from MIT have called it) is hindered by a severe complex action problem of lattice QCD at non-zero baryon chemical potential. Quantum simulation, i.e. the use of special purpose quantum analog computers, is currently arising as a promising tool that may eventually help us to overcome such problems.

F.11 Different Layers of Reality and Limits of Physics

As we have discussed, in physics we patch together effective theories valid at different length scales. Indeed there are different layers of physical reality, each with its own appropriate mathematical description. While the dynamics at the shortest presently accessible distances are described by the Standard Model of particle physics, atomic nuclei are described by their own effective field theory valid at larger distance scales. Stepping further up in length scale, we reach atoms, molecules, and then condensed matter, which are all described by their own effective theories. While the resulting "quilt" of effective theories may eventually cover all physical phenomena, it is very unlikely to cover even a small fraction of other layers of reality. After all, the success of physics relies on the identification of subsystems with a sufficiently structured dynamics in space and time that can be described mathematically. While mathematics is a universal language that even Nature uses to express herself in, it is not a useful tool for talking about other layers of reality, such as art and poetry, or mind, consciousness, and free will. Those are very unlikely to ever become subjects of physics, just because mathematics is not an appropriate language for communicating about these emergent concepts which cannot be appropriately described, for example, in terms of space-time coordinates. While the subjects of physics are thus somewhat

limited, the Universe is such a rich and diverse place that there is no reason to think that physics will be exhausted any time soon.

F.12 Understanding our own Brain

As suggested above, by means of physicists the Universe is "thinking" about itself, thus enabling an understanding of part of the system from within. This all happens thanks to the wonderful device we carry around in our head. Will physicists, biologists, or neuroscientists ever be able to understand the brain itself, essentially by using it to think hard enough about itself? While Gödelian self-referential situations are likely to prevent complete understanding, there may be no reason to be overly pessimistic. In particular, numerical simulations of certain brain functions should be possible. Except for that, however, the brain is probably the last system a typical physicist is well prepared to deal with. First of all, due to its large interconnectedness, the structure of the brain is to a large degree spatially non-local. Due to long-term memories stored in the brain, locality in time is disfigured as well. Furthermore, due to a lack of obvious hierarchies in length scales, it seems difficult to identify relevant degrees of freedom that can be separated from the rest of the system. In other words, the physicist's most powerful tool of effective field theory is not expected to work in an environment as complex as the brain itself.

F.13 Neuroscience as a "Theory of Everything" related to Neuronal Activity?

In recent years, some influential neuroscientists including Gerhard Roth and Wolf Singer have come to the conclusion that free will is an illusion, because conscious decisions are preceded by sub-conscious neuronal events in the brain. This has even led them to argue that law should be rewritten, in order to take into account that people cannot be held responsible for their actions. As a physicist working with effective field theories applied to different layers of physical reality, I find these arguments rather absurd. The neuronal scale "effective theory" of the neuroscientist has not been properly matched to the framework of mind and free will used in the humanities. Cross-communication between these different layers of reality is hence not possible in a truly meaningful way. A neuroscientist, who claims that free will is an illusion, uses his neuronal scale "effective theory" as a "Theory of Everything" related to neuronal activity. While emergent concepts like free will or consciousness should indeed not be inconsistent with fundamental theories underlying neuronal

activity, they can neither be derived nor disproved by these theories. Therefore, the "Theory of Everything" concerning neuronal activity is very likely at the same time a "Theory of Nothing" relevant to legislation. Interestingly, based on arguments of a similar nature, some philosophers including Peter Bieri, David Chalmers, Joseph Levine, and Thomas Nagel have reached the same conclusion.

If it exists, free will should indeed not be inconsistent with fundamental theories underlying the brain. Isn't this the case if neuronal activity precedes our conscious decisions? As pointed out before, the brain itself is a highly non-local structure both in space and in time. Indeed, it is a wonderful device that turns chemical energy, which can be described mathematically by an effective theory using space-time coordinates, into concepts of the mind, which are subjects of the humanities that exist beyond space and time. Since a conscious decision is an emergent phenomenon which is not associated with a unique time-coordinate, it does not really make sense to say that it was preceded by some specific neuronal activity. Since the fundamental theories of physics are based on quantum uncertainty rather than classical determinism, they are consistent with free will as well, but cannot derive it either. At the moment, we simply don't know how to match the two separate layers of reality in which we can talk about either mind or matter, but not about the relations between both of them.

F.14 Matching the Mind and Matter Layers of Reality

The mind-body problem, i.e. to understand how mind and matter are related, and, if possible, to explain how mind emerges from matter, has been around for a long time. Since effective field theory does not work for a structure as complex and non-local as the brain, physics is currently nowhere near contributing to the solution of the mind-body problem. As argued before, neuroscience or any other natural science isn't either. Perhaps simulation by very powerful future computers (either involving classical randomness or quantum indeterminacy) may eventually mimic something like a "brain". While this may contribute to bridging the gap between mind and matter, it will not necessarily lead to deep understanding. In physics, renormalization group theory leads to deep understanding and enables us to match the different layers of physical reality separated by different length scales. Similarly, one may speculate that a future "Theory of Mind and Matter" may be able to match the natural sciences' space-time description of matter with the humanities' description of mind beyond space and time. Developing such a theory, and thus eventually solving the mind-body problem, is a tremendous challenge for both natural science and the humanities. Most likely, time is not ripe yet for this endeavor, because the

individual disciplines still need to sharpen their tools before they can address the problem in a meaningful manner. However, it seems very well worth keeping the solution of the mind-body problem on the agenda, as a potential common long-term goal of natural science and the humanities. Negating the problem, by declaring neuroscience as the "Theory of Everything" related to neuronal activity, is much easier but completely misleading.

F.15 Evolution, Brain Chemistry, and Physics

After this exploration of the limits of physics and other sciences (including neuroscience), let us return to the question why physics exists. While all necessary conditions for physics are indeed fulfilled in our Universe, we should not forget that physics exists also because it is a lot of fun. Indeed it is the pleasure of figuring things out that is driving many new discoveries. The reward system in our brain, which evolved in order to help our ancestors outsmart predators as well as prey, also seems to encourage the creativity that one needs today as a physicist. However, in order not to paint a too naive picture of what drives physicists, we should also not forget that the power of physics to advance technology, unfortunately including the one used to fight other nations, has also motivated numerous research projects. In any case, mankind's well-being for the rest of the century is likely to benefit from physicists whose imagination is spurred by more honorable causes, including, for example, advancing climate research or medical applications of physics. Still, based on my own experience, I think that curiosity-driven basic research in physics benefits tremendously from our brain chemistry.

F.16 Summary

To summarize, as far as I understand, physics crucially depends on the existence of mathematical structures, space-time locality, as well as large hierarchies of length scales. These may all be necessary prerequisites for the existence of curious minds and capable brains of potential physicists as well. Invoking the anthropic principle, taking into account that physics in 4-d is not too difficult and can actually be quite useful, and knowing that it is a pleasure to figure things out, it is perhaps not surprising that physics does indeed exist. To a large extent, the success of physics relies on the powerful tool of effective theories. Model building and numerical simulation are very important tools as well, which are about to be enriched by quantum simulators. Of course, as usual when one invokes the anthropic principle, many questions remain unanswered. As curious minds, we still urge to understand why

there are mathematical structures, space and time, locality, as well as large hierarchies. There is no reason not to think about these deep questions. Understanding why there are curious minds seems even harder, and may require a future "Theory of Mind and Matter" that matches the mathematical language of natural science to the concepts of the humanities. While the subjects of physics cover only a small fraction of reality, as far as I'm concerned, there is plenty for a life-time of a curious mind. Finally, taking part in the process of the Universe "thinking" about itself, together with colleagues all around the world, is a fascinating and most rewarding experience (not only in the brain chemistry sense).