Nair's Lecture Notes on Algebraic & Geometric Methods in Physics

Contents

Part I

Algebraic Methods

Chapter 1

Harmonic oscillators and the $SU(N)$ algebra

1.1 Generalities on the algebraic approach to quantum theory

In this first part, we will start with some algebraically solvable problems. The key to this approach is the observation that the quantum theory of any physical system can be viewed as a unitary irreducible representation of the algebra of observables. We will explore and elucidate the meaning of unitarity and irreducibility as we get deeper into various examples we consider. Our approach will be more of a bottom-up approach, moving from the particulars to an appreciation of the general. However, a few general observations can be useful at this point. The operator algebra for the observables cannot be just any algebra. We need a way to connect the operators or elements of the algebra to the real numbers we can measure in the laboratory. Thus a certain notion of a norm on the algebra is necessary. A notion of conjugation is also needed to give meaning to hermiticity of an operator. The minimal requirements would then characterize the observables as a C^*-algebra. (Additional requirements such as Poincaré invariance would be needed for a relativistically invariant field theory context.)

Let us specify, for completeness, the definition of a C^*-algebra. Recall that an algebra $\mathcal{A}$ can be considered as a linear vector space in the sense that elements of the algebra can be added with coefficients in a field, say, complex numbers for us. Thus if $A, B \in \mathcal{A}$, then $\lambda_1 A + \lambda_2 B \in \mathcal{A}$, for $\lambda_1, \lambda_2 \in \mathbb{C}$. Further we need a norm $||A||$ for the elements of the algebra and we want to impose completeness with respect to this norm. If we also require that we have an associative multiplication law for elements of the algebra and that this obeys the Schwarz inequality in the sense $||A\,B|| \leq ||A||\,||B||$, then $\mathcal{A}$ is a Banach algebra. Finally, we need a $*$-operation which is an additional requirement

making $\mathcal{A}$ into a Banach $*$-algebra. The $*$-operation obeys the requirements

$$\begin{aligned}(A+B)^* &= A^*+B^* \\ (\lambda A)^* &= \bar{\lambda}\, A^* \\ (A\,B)^* &= B^*\, A^* \\ (A^*)^* &= A\end{aligned}$$

(The adjoint operation on matrices is an example of this.)

A C^*-algebra is a Banach $*$-algebra with the additional requirement that $||A^*\,A|| = ||A||^2$. By a theorem of Gel'fand and Naimark, a C^*-algebra is equivalent to the Banach algebra of operators on a Hilbert space that contains the adjoint along with each operator.

Having made these general statements, we will not use them very much. We will discuss examples and try to evolve into an appreciation of the general framework. In fact, we will start with something very elementary, the harmonic oscillator in quantum mechanics.

1.2 The one-dimensional oscillator

We start by considering a one-dimensional harmonic oscillator characterized by the angular frequency ω. The Hamiltonian of this oscillator is given by

$$H = \frac{1}{2}\left(\hat{p}^2 + \omega^2 \hat{q}^2\right) \tag{1.1}$$

where the operators corresponding to the position (or coordinate) and momentum, *i.e.*, $\hat{q}$ and $\hat{p}$, respectively, satisfy the well-known Heisenberg algebra

$$[\hat{q}, \hat{p}] = i\,. \tag{1.2}$$

(We use units in which $\hbar = 1$.) We can change variables for these operators, parametrizing them as

$$\hat{q} = \sqrt{\frac{1}{2\omega}}(a^\dagger + a)\,, \quad \hat{p} = i\sqrt{\frac{\omega}{2}}(a^\dagger - a) \tag{1.3}$$

The new operators $a^\dagger$ and a are known as the creation and annihilation operators, respectively. We can invert the relations given above and express $a^\dagger$ and a as

$$a = \sqrt{\frac{\omega}{2}}q + i\frac{p}{\sqrt{2\omega}}\,, \quad a^\dagger = \sqrt{\frac{\omega}{2}}q - i\frac{p}{\sqrt{2\omega}}\,. \tag{1.4}$$

In the above expressions, and in what follows, we omit the hats for operators. Using the commutation relation (1.2), we find

$$[a, a^\dagger] = 1\,. \tag{1.5}$$

This can be considered as an alternative presentation of the Heisenberg algebra. From (1.4), we can easily find

$$a^\dagger a = \frac{\omega}{2} q^2 + \frac{p^2}{2\omega} - \frac{1}{2}\,, \tag{1.6}$$

$$H = \omega \left(a^\dagger a + \frac{1}{2} \right) . \tag{1.7}$$

The positivity of the operator $a^\dagger a$ can be seen from the relation

$$\langle \alpha | a^\dagger a | \alpha \rangle = \sum_\lambda \langle \alpha | a^\dagger | \lambda \rangle \langle \lambda | a | \alpha \rangle = \sum_\lambda |C_{\lambda\alpha}|^2 \geq 0 \tag{1.8}$$

where $|\alpha\rangle$ and $|\lambda\rangle$ denote arbitrary states, and $C_{\lambda\alpha} = \langle \lambda | a | \alpha \rangle$ and $C^*_{\lambda\alpha} = \langle \alpha | a^\dagger | \lambda \rangle$. We have also used the completeness relation $\sum_\lambda |\lambda\rangle\langle\lambda| = \mathbb{1}$. The ground state may be obtained by minimizing the expectation value of $a^\dagger a$. From (1.8), the lowest possible value is zero and hence we need $C_{\lambda 0} = \langle \lambda | a | 0 \rangle = 0$, for any choice of $|\lambda\rangle$. Thus we can define the ground state $|0\rangle$ to be specified by

$$a|0\rangle = 0\,, \quad H|0\rangle = \frac{\omega}{2}|0\rangle\,. \tag{1.9}$$

We have to still show that such a state exists. This equation, in the x-diagonal basis, may be written as $\langle x | a | 0 \rangle = 0$. Using the expression of a in (1.4), we find that this equation can be written as

$$\left(\sqrt{\frac{\omega}{2}}\, x + \frac{1}{\sqrt{2\omega}} \frac{\partial}{\partial x} \right) \langle x | 0 \rangle = 0\,. \tag{1.10}$$

The solution to this equation is given by

$$\langle x | 0 \rangle = C \exp\left(-\omega\, \frac{x^2}{2} \right) \tag{1.11}$$

where C is a normalization constant. We have found a normalizable wave function for the ground state; this guarantees the existence of the ground state $|0\rangle$.

Since a acting on the ground state annihilates it, the only thing we can do to get another state is the application of $a^\dagger$. We can easily verify that

$$[H, a^\dagger] = \omega\, a^\dagger \tag{1.12}$$

which leads to

$$H\, a^\dagger\, |0\rangle = a^\dagger\, H\, |0\rangle + \omega\, a^\dagger |0\rangle = \omega\, (1 + \frac{1}{2})\, a^\dagger |0\rangle\,. \tag{1.13}$$

Thus $a^\dagger|0\rangle$ is an eigenstate of the Hamiltonian with eigenvalue $\frac{3}{2}\omega$. Higher excited states can be constructed from $|0\rangle$ by the multiple application of the $a^\dagger$'s. Accordingly, the energy spectrum of the oscillator becomes discrete. The list of

states and corresponding energies for the one-dimensional harmonic oscillator is given below.

State	$\|0\rangle$	$a^\dagger\|0\rangle$	$\frac{1}{\sqrt{2}}(a^\dagger)^2\|0\rangle$	$\cdots$	$\frac{1}{\sqrt{n!}}(a^\dagger)^n\|0\rangle$
Energy	$\frac{1}{2}\omega$	$\frac{3}{2}\omega$	$\frac{5}{2}\omega$	$\cdots$	$\left(n+\frac{1}{2}\right)\omega$

(1.14)

On the states $|n\rangle = (1/\sqrt{n!})(a^\dagger)^n|0\rangle$, we can work out the action of the operators a and $a^\dagger$ as

$$a\,|n\rangle = \sqrt{n}\;|n-1\rangle\,, \qquad a^\dagger\,|n\rangle = \sqrt{n+1}\;|n+1\rangle\,. \tag{1.15}$$

This constitutes an explicit representation of the Heisenberg algebra.

There are many other ways to represent the Heisenberg algebra. For example, we can take

$$a = \frac{\partial}{\partial z}, \qquad a^\dagger = z \tag{1.16}$$

with the understanding that these act on holomorphic functions f of the complex variable z. If we take the inner product to be

$$\langle\alpha|\beta\rangle = \int \frac{d^2z_1}{\pi} e^{-\bar{z}z}\;\bar{f}_\alpha\, f_\beta \tag{1.17}$$

then we can check that a and $a^\dagger$ are adjoints, as required. (This representation (1.16, 1.17) is the so-called coherent state representation.) So the question arises: Is this a brand new representation and, if so, which representation should one use for the quantum theory? The following theorem assures us that this is not quite a new representation, but is related to the one in (1.15) by a unitary transformation.

Theorem 1.1 *(Stone–von Neumann Theorem) For any finite number of operators a_i, $a_i^\dagger$ obeying the Heisenberg algebra, there is only one irreducible representation up to a unitary transformation, provided the underlying phase space (of x_i, p_i) is simply connected.*

Since unitarily equivalent descriptions give the same physics in quantum mechanics, there is no ambiguity; use of any representation will give the same physical results.

Another useful remark may be that the existence of the ground state for the oscillator arises from the fact that the Hamiltonian for the oscillator is bounded from below If we have a Hamiltonian which is not bounded below, then there is no ground state in general.

1.3 The two-dimensional oscillator

We now consider application of a similar analysis as in the previous section to two-dimensional oscillators. Operators p, q, a, $a^\dagger$ are now labeled by a directional index $i = 1, 2$, say, corresponding to x and y directions. We will consider an isotropic oscillator so that its angular frequency remains the same in both directions; $\omega \equiv \omega_1 = \omega_2$. The algebra of a_i and $a_j^\dagger$ can be expressed as

$$[a_i, a_j^\dagger] = \delta_{ij}, \qquad [a_i, a_j] = 0, \qquad [a_i^\dagger, a_j^\dagger] = 0\,. \tag{1.18}$$

This can be interpreted as two copies of the Heisenberg algebra (1.5). The Hamiltonian of the oscillator is given by

$$\begin{aligned} H &= \omega(a_1^\dagger a_1 + a_2^\dagger a_2 + 1) \\ &= \omega(n_1 + n_2 + 1) \end{aligned} \tag{1.19}$$

where n_i denotes the number operator $n_i = a_i^\dagger a_i$ $(i = 1, 2)$. The ground state $|0\rangle$ is defined by

$$a_1|0\rangle = a_2|0\rangle = 0\,. \tag{1.20}$$

The first excited state is then given either by $a_1^\dagger|0\rangle$ or $a_2^\dagger|0\rangle$. Using the fact that the operators in different directions commute with each other, we can easily find that the action of $n = n_1 + n_2$ on $a_i^\dagger|0\rangle$ becomes $1 \cdot |0\rangle$. This means that the first excited state is a degenerate state of energy ω with degeneracy equal to 2. We may denote the states as $|1, 0\rangle$ and $|0, 1\rangle$ in the notation $|n_1, n_2\rangle$ for eigenstates of the number operators n_1 and n_2. Higher excited states can analogously be constructed. For $n = n_1 + n_2 > 1$, a set of degenerate states can be written, up to normalization, as

$$(a_1^\dagger)^n|0\rangle\,,\ (a_1^\dagger)^{n-1}a_2^\dagger|0\rangle\,,\ \cdots,\ (a_2^\dagger)^n|0\rangle\,. \tag{1.21}$$

All these states have the same energy and thus the degeneracy is given by $n+1$. These states can generally be labeled as

$$|n_1, n_2\rangle = \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} |0\rangle \tag{1.22}$$

whose energy eigenvalues are given by $E_{n_1,n_2} = \omega(n_1 + n_2 + 1) = \omega(n + 1)$.

We can also work out the action of the basic operators a_i, $a_i^\dagger$ on these states as

$$\begin{aligned} a_1\,|n_1,\, n_2\rangle &= \sqrt{n_1}\,|n_1 - 1,\, n_2\rangle\,, \\ a_2\,|n_1,\, n_2\rangle &= \sqrt{n_2}\,|n_1,\, n_2 - 1\rangle\,, \\ a_1^\dagger\,|n_1,\, n_2\rangle &= \sqrt{n_1 + 1}\,|n_1 + 1,\, n_2\rangle\,, \\ a_2^\dagger\,|n_1,\, n_2\rangle &= \sqrt{n_2 + 1}\,|n_1,\, n_2 + 1\rangle\,. \end{aligned} \tag{1.23}$$

We have thus obtained a representation of the two-dimensional Heisenberg algebra on the vector space spanned by the vectors $|n_1,\, n_2\rangle$.

Going back to the discussion of the states, notice that the second one among the states in (1.21) is obtained from the first by replacing one $a_1^\dagger$ by $a_2^\dagger$, or by the application of the operator $a_2^\dagger a_1$. More generally, the operators which connect the degenerate states $|n_1, n_2\rangle$, with $n = n_1 + n_2$ fixed, are given by

$$K_- = a_2^\dagger a_1\,, \quad K_+ = a_1^\dagger a_2\,. \tag{1.24}$$

Actions of these operators can be described explicitly as below:

$$(a_1^\dagger)^n|0\rangle \underset{a_1^\dagger a_2}{\overset{a_2^\dagger a_1}{\rightleftharpoons}} (a_1^\dagger)^{n-1}a_2^\dagger|0\rangle \underset{a_1^\dagger a_2}{\overset{a_2^\dagger a_1}{\rightleftharpoons}} (a_1^\dagger)^{n-2}(a_2^\dagger)^2|0\rangle \underset{a_1^\dagger a_2}{\overset{a_2^\dagger a_1}{\rightleftharpoons}} \cdots$$

We can also calculate the commutator of K_+ and K_- to obtain

$$[K_+, K_-] = a_1^\dagger a_1 - a_2^\dagger a_2 \equiv 2K_3\,. \tag{1.25}$$

where we defined a new combination $K_3 = \frac{1}{2}(n_1 - n_2)$. Further K_3 is easily verified to satisfy the relations

$$[K_3, K_+] = K_+\ , \quad [K_3, K_-] = -K_-\,. \tag{1.26}$$

The operators $K_\pm$ and K_3 are all that we need to describe the degenerate states. The algebra formed by these operators is the three-dimensional angular momentum algebra or the $O(3)$ algebra. Notice that there are no physical pictures of angular momenta here since we are considering the two-dimensional oscillator. It is just that the the same mathematical structure arises as the algebra connecting various degenerate states. As we shall see below, the angular momentum algebra is also equivalent (or isomorphic) to the $SU(2)$ algebra. The maximum value of $K_3 = \frac{1}{2}(a_1^\dagger a_1 - a_2^\dagger a_2)$ among the set of degenerate states is given by $n/2 \equiv j$. Drawing on our knowledge of angular momentum theory, it is easy to see that there is a Casimir operator $K^2 = K_3^2 + \frac{1}{2}(K_+K_- + K_-K_+)$ which commutes with $K_\pm$ and K_3. In terms of j, this quadratic Casimir operator can be expressed as

$$K^2 = K_3^2 + \frac{1}{2}(K_+K_- + K_-K_+) = j(j+1)\,. \tag{1.27}$$

Notice that $K_\pm$ are not hermitian operators, but are hermitian conjugate to each other; hermitian combinations can be obtained as

$$K_1 = \frac{K_+ + K_-}{2}\,, \quad K_2 = \frac{K_+ - K_-}{2i}\,. \tag{1.28}$$

In terms of the K_i $(i = 1, 2, 3)$, the algebra takes the more familiar form

$$[K_i, K_j] = i\epsilon_{ijk}K_k\,. \tag{1.29}$$

We now consider matrix representation of the angular momentum algebra, equivalently, the algebra (1.29). The lowest nontrivial case is given by $n = 1$, with the degeneracy being 2. The corresponding two states can be labeled by $a_1^\dagger|0\rangle \equiv |1\rangle$ and $a_2^\dagger|0\rangle \equiv |2\rangle$. In terms of these, the matrix elements of the operator $K_+ = a_1^\dagger a_2$ are expressed as

$$\begin{pmatrix} \langle 1|K_+|1\rangle & \langle 1|K_+|2\rangle \\ \langle 2|K_+|1\rangle & \langle 2|K_+|2\rangle \end{pmatrix} = \begin{pmatrix} 0 & \langle 1|1\rangle \\ 0 & \langle 2|1\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = K_+ \,. \tag{1.30}$$

where we use the relations

$$\begin{aligned} K_+|1\rangle &= a_1^\dagger a_2 a_1^\dagger|0\rangle = 0\,, \\ K_+|2\rangle &= a_1^\dagger a_2 a_2^\dagger|0\rangle = a_1^\dagger|0\rangle = |1\rangle \end{aligned} \tag{1.31}$$

and the orthogonality of the two states, *i.e.*, $\langle 1|2\rangle = \langle 2|1\rangle = 0$. Similarly we can easily find

$$K_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\,, \quad K_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}. \tag{1.32}$$

This means that for $n = 1$ the matrix representation of K_i $(i = 1, 2, 3)$ is essentially given by the Pauli matrices σ_i, *i.e.*, $K_i = \frac{1}{2}\sigma_i$ which obey the algebra (1.29).

For $n = 2$, there are three degenerate states which are given by

$$|1\rangle \equiv \frac{(a_1^\dagger)^2}{\sqrt{2}}|0\rangle, \quad |2\rangle \equiv a_1^\dagger a_2^\dagger|0\rangle, \quad |3\rangle \equiv \frac{(a_2^\dagger)^2}{\sqrt{2}}|0\rangle\,. \tag{1.33}$$

By use of these states, in an analogous way to what was done for $n = 1$, we can calculate K_i as (3×3) matrices with the result,

$$K_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}, \; K_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}, \; K_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{1.34}$$

As is well-known, these expressions form a (3×3)-matrix representation of the angular momentum algebra.

More generally, the degree of degeneracy for the n-th level of the two-dimensional isotropic oscillator is given by $n + 1$ and, in this case, K_i become $(n + 1) \times (n + 1)$ matrices which obey the angular momentum algebra

$$[K_i, K_j] = i\epsilon_{ijk}K_k\,. \tag{1.35}$$

Alternatively we can begin with the algebra (1.35), say $\mathcal{A}$, viewed as an abstract algebra, the K_i not necessarily being matrices. This algebra has many matrix realizations, of dimensions $(n + 1)$, $n = 0, 1, 2$, etc. The K_i are hermitian $(n+1) \times (n+1)$ matrices. Thus the transformations generated by the K_i which are of the form $\exp(iK_i\theta^i)$ are unitary matrices. We may thus say that the matrix representations we found are unitary representations of the algebra $\mathcal{A}$.

In summary, we find that degenerate states of a two-dimensional harmonic oscillator form a unitary representation of the angular momentum algebra. In other words, we find that the degenerate states give a unitary representation of the algebra (1.35).

Lie's theorems

We shall now show how the algebra (1.35) is the same as the $SU(2)$ algebra. We first consider the group $SU(2)$ which can be defined as

$$SU(2) \;=\; \{\text{set of all } 2\times 2 \text{ unitary matrices with det} = 1\} \tag{1.36}$$

As mentioned before, a 2×2 unitary matrix is expressed as an exponential of a 2×2 hermitian matrix times i. A general 2×2 hermitian matrix can be parametrized by $\phi\mathbb{1} + \theta^i\frac{\sigma_i}{2}$ where ϕ and θ^i $(i=1,2,3)$ are real parameters and the σ_i are the 2×2 Pauli matrices. Thus a general 2×2 unitary matrix can be expressed as

$$g = \exp\left(i\left(\phi\mathbb{1} + \theta^i\frac{\sigma_i}{2}\right)\right)\,. \tag{1.37}$$

From the condition $\det g = 1$, we find that we need $\phi = 0$. For this it is easiest to use the relation $\det\left(e^M\right) = e^{\mathrm{Tr}M}$ which holds for an arbitrary matrix M. The most general $SU(2)$ matrix is thus parametrized as

$$g(\theta) = \exp\left(i\frac{\sigma_i}{2}\theta^i\right) \tag{1.38}$$

where θ^i are continuous real parameters. Thus each element of the group is labeled by the value of the continuous parameters θ^i, hence $SU(2)$ is a continuous group. The group composition law which takes the form of matrix multiplication gives the rule for composing the values of the parameters as well. In other words, the relation

$$g(\theta)\,g(\theta^{'}) = g(\theta^{''}) \tag{1.39}$$

specifies $\theta^{''}$ as a function of θ and $\theta^{'}$. With some analyticity requirements on this function, the continuous group becomes a Lie group. (For a proper definition, see Chapter 12.) The corresponding algebra, which is formed by a set of generators of the Lie group, is called the Lie algebra.

In general, generators t_i of a Lie group G are defined by specifying a general group element as $g = \exp(it_i\theta^i)$. Here $i = 1, 2, \cdots, dim\,G$, with $dim\,G$ denoting the dimension of G. For the group $SU(2)$, we have used its definition which naturally leads to a 2×2 matrix representation. Such a representation is called the defining representation of the group. For the $SU(2)$, the generators in the defining representation may be taken to be $\frac{\sigma_i}{2}$, $(i=1,2,3)$. This also means that any element of the algebra is a linear combination of these basis elements. Any other basis in the algebra will do as well; this is just a convenient choice.

The algebra of the generators $t_i = \frac{\sigma_i}{2}$ can be worked out from its matrix form. In this case, we get

$$[t_i, t_j] = i\epsilon_{ijk}\,t_k\,. \tag{1.40}$$

The idea is then to interpret this as a more general algebra, the t_i being not just 2×2 matrices and look for other representations. The algebraic relations (1.40) obviously show that the $SU(2)$ algebra is equivalent to the angular momentum algebra.

In general, for any Lie group, there exists a Lie algebra

$$[t_i, t_j] = C_{ij}^k t_k \tag{1.41}$$

where C_{ij}^k are constants that are independent of the parameters θ^i. The t_i in the defining representation can be found from the definition of the group, by examining the elements near the identity element. Also, if we are given a Lie algebra, we can reconstruct the Lie group by exponentiating the generators with arbitrary parameters (as in $g = \exp(it_i\theta^i)$). This will give the group elements, but of a group which is the original group we started with up to some global identifications. In other words, different groups which differ by global identifications all give the same Lie algebra. Therefore exponentiating the generators to construct the group elements will give one of these possible groups, the simply connected universal covering group of the set of groups which only differ by the global identifications.

These results on the Lie algebra, particularly the fact that C_{ij}^k in (1.41) are constants and on how the algebra is related to the Lie group are theorems due to Sophus Lie. (See Chapter 12 for details on Lie's theorem).

Once we have the Lie algebra, we can construct a unitary irreducible representation of the algebra. For a compact group, this can be done using the products of its defining (or fundamental) representation. For most physical problems of interest, we need unitary representations. This is primarily due to the fact that the allowed transformations of the Hilbert space in quantum mechanics are unitary transformations.

Unitary irreducible representation

What we have found so far is that a degenerate multiplet of states of a two-dimensional harmonic oscillator is given by a unitary representation of the $SU(2)$ algebra. To be more specific, the degenerate states give us an *irreducible* representation of $SU(2)$, which means that all the states of the multiplet are connected by the transformations generated by the algebra and further that the action of the transformations is closed, not giving any brand new state which has not been counted in the multiplet.

1.4 Theorems on symmetries

We will now formulate these ideas a bit more systematically by formulating a general theorem on symmetries and conservation laws and the degeneracy of states. Consider a general quantum system with a Hamiltonian H and a

unitary transformation U defined by a continuous parameter θ so that we may write $U = \exp(i\theta\, G)$ for some hermitian operator G. (G is the generator of the transformation.) The new Hamiltonian obtained by unitary transformation of the Hamiltonian H is $H' = U^\dagger\, H\, U$. Since θ is a continuous parameter, we can consider the case when it is infinitesimal. In this case, we may write

$$H' = e^{-i\theta G}\, H\, e^{i\theta G} \approx (1 - i\theta G)\, H\, (1 + i\theta G) = H - i\theta\, [G, H] \tag{1.42}$$

Thus the change in the Hamiltonian is $\delta H = -i\theta\, [G, H]$. We say that U is a symmetry or that G generates a symmetry if $\delta H = 0$, *i.e.*, if $[G, H] = 0$.

Now by the Heisenberg equation of motion, we have, for any operator G,

$$i\frac{\partial G}{\partial t} = [G, H]\,. \tag{1.43}$$

Thus if G generates a continuous symmetry (so that $[G, H] = 0$), the G is conserved. Obviously, the result extends to any number of symmetries with the corresponding parameters θ_i and generators G_i. This is Noether's theorem.

Theorem 1.2 *(Noether's Theorem) For every continuous symmetry of the quantum system, there is a conserved observable which is the generator of the symmetry transformation.*

The reverse of this statement that if we have a conserved quantum number, it generates a symmetry of the system is true to some extent. The exception would be for some quantity which is conserved for topological reasons and then the symmetry is not always obtained. There are examples of such quantities in physics, we will return to this question when we talk about solitons later.

Continuing with the analysis of symmetry, consider an eigenstate $|\alpha\rangle$ of the Hamiltonian with eigenvalue E_α. Thus $H\,|\alpha\rangle = E_\alpha\,|\alpha\rangle$. If U is a symmetry transformation, which may or may not be continuous, we have $U^\dagger\, H\, U = H$ or $H\, U = U\, H$. Thus

$$H\,(U\,|\alpha\rangle) = U\, H\,|\alpha\rangle = E_\alpha\,(U\,|\alpha\rangle) \tag{1.44}$$

We see that $U\,|\alpha\rangle$ is also an eigenstate of the Hamiltonian with the same energy. Thus all states which are related to each other by a symmetry transformation have the same energy. In other words, they form a degenerate multiplet of states. How many are there? That depends on the transformation. The state $U\,|\alpha\rangle$ can be different from the starting state $|\alpha\rangle$ or it could be the same. There may be several U's, as for example, in the continuous case where we get a U for each value of the parameter θ. Imagine we start from one state and get all possible states connected to it by the symmetry transformations, counting only the distinct states. We get a set of states $\{|\alpha_i\rangle\}$ which have the following property. All these states are degenerate having the same eigenvalue for H. Further, the action of any U on any state in this set will produce a linear

combination of states in the same set. Thus the set is closed under the action of U. We get a matrix realization of the U's of the form $\langle\alpha_i|U|\alpha_j\rangle$ on this set of states. We may thus conclude that the degenerate set of states furnish an irreducible representation of the symmetry transformations. If U is continuous we may express this in terms of the algebra of the generators of the symmetry, namely, the G's. We may express this as another theorem.

Theorem 1.3 *The eigenstates of the Hamiltonian of a quantum system can be classified into irreducible representations of the symmetry transformations, the states within each irreducible representation being degenerate.*

It should be noted that one can have many copies of the same irreducible representation, the multiplets being distinguished by some other quantum number which is not part of the symmetry under consideration. Also, there may be accidental degeneracies where the states from two distinct irreducible representations may be degenerate, although it is not required by the theorem. Generally, such degeneracies can be traced to symmetries which are not obvious, but somewhat hidden.

As an example, consider the Hydrogen atom with the Hamiltonian

$$H = \frac{p^2}{2m} - \frac{e^2}{r}\,. \tag{1.45}$$

This Hamiltonian has an obvious symmetry under rotations, since it only involves the magnitudes of p_i and x_i. Thus the generators of rotations, namely the angular momentum, commutes with H. Thus energy eigenstates fall into irreducible representations of the angular momentum algebra. Recall that the energy eigenstates of the Hamiltonian (1.45) can be labeled as $|n,l,m\rangle$, where $n = 1, 2$, etc., is the principal quantum number. The angular momentum quantum number l can take values $0, 1, \cdots, (n-1)$ for a given n. Further, the azimuthal quantum number m takes values $-l, -l+1, \cdots, l-1, l$, giving $2l+1$ possibilities for each l. The theorem then tells us that the $2l+1$ states for various m values, with n and l being fixed, are degenerate. And this is indeed true. For example, at $n = 2$, we have the states

$$|2,0,0\rangle, \qquad |2,1,1\rangle,\ |2,1,0\rangle,\ |2,1,-1\rangle\,.$$

The theorem tells us the eigenvalues

$$E_{2,1,1} = E_{2,1,0} = E_{2,1,-1} \tag{1.46}$$

while $E_{2,0,0}$ can be different. Now it so happens for the Hamiltonian (1.45), we also have $E_{2,0,0} = E_{2,1,1}$, etc. This would be an example of the accidental degeneracy, but it can be traced to an enhanced symmetry of (1.45) related to the Runge-Lenz vector. We will take this up in the next chapter.

1.5 The higher dimensional oscillators

Extension of the above analysis to the three-dimensional oscillator is straightforward. The Hamiltonian is given by

$$H = \omega\left(a_1^\dagger a_1 + a_2^\dagger a_2 + a_3^\dagger a_3 + \frac{3}{2}\right) = \omega\left(a_i^\dagger a_i + \frac{3}{2}\right) \tag{1.47}$$

where $i = 1, 2, 3$. The total number operator is given by $n = n_1 + n_2 + n_3 = a_i^\dagger a_i$. The degeneracy can be counted as $\frac{1}{2}(n+1)(n+2)$. Operators that connect the degenerate states include $a_2^\dagger a_1$, $a_3^\dagger a_1$, $a_3^\dagger a_2$, as shown in Figure 1.1, and the conjugates of these, *i.e.*, $a_1^\dagger a_2$, $a_1^\dagger a_3$, $a_2^\dagger a_3$.

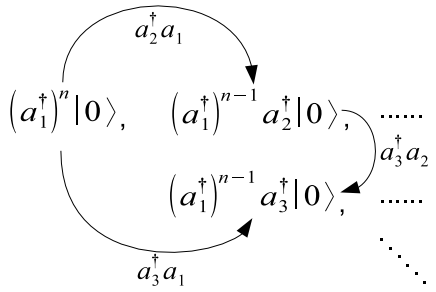

Figure 1.1: Examples of degenerate states and operators that connect them

In order to have a closed operator algebra, we need to add other operators to these six operators. Taking the commutator of two of the operators listed,

$$\begin{aligned}
[a_1^\dagger a_2, a_2^\dagger a_1] &= a_1^\dagger [a_2, a_2^\dagger a_1] + [a_1^\dagger, a_2^\dagger a_1] a_2 \\
&= a_1^\dagger a_1 - a_2^\dagger a_2 \\
&= n_1 - n_2 \,.
\end{aligned} \tag{1.48}$$

This shows that we will need to include $n_1 - n_2$ to our list of operators. Similarly, we need $(n_2 - n_3)$ and $(n_3 - n_1)$. The latter can be expressed as $(n_3 - n_1) = -(n_1 - n_2) - (n_2 - n_3)$. Thus we need two additional operators to close the operator algebra. The algebra formed by the total of eight operators is an $SU(3)$ algebra. This will become evident in the following.

Representation of $SU(3)$

As in the case of $SU(2)$, elements of the $SU(3)$ group are parametrized as $g = e^{it}$ where t is a (3×3) hermitian traceless matrix. Since a 3×3 matrix has 9 independent matrix elements possible, but tracelessness gives one condition, we

expect 8 independent parameters to specify an $SU(3)$ element. Thus the matrix t can be parametrized as $t = t^a\theta^a$ $(a = 1, 2, \cdots, 8)$ where t^a's represent a basis of hermitian traceless matrices and θ^a's are real parameters. (Alternatively, we can parametrize t by three complex numbers which are the parameters for $a_i^\dagger a_j$ $i \neq j$, and two real numbers which are the parameters corresponding to the diagonal generators.) Conventionally, the basis is taken as $t^a = \frac{\lambda^a}{2}$ where λ^a are the so-called Gell-Mann matrices given below.

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{1.49}$$

The normalization of these matrices is such that $\mathrm{Tr}(t^a t^b) = \frac{1}{2}\delta^{ab}$. The commutator $[t^a, t^b]$ is anti-hermitian. So it is expressed as $i \times$ (hermitian matrix). Since the commutator is traceless as well, it can be expanded in terms of the t^a themselves, so that we can write

$$[t^a, t^b] = i f^{abc} t^c \,. \tag{1.50}$$

f^{abc}'s are called the structure constants. They are a set of numbers; the specific value are not important to us at this stage. These commutation relations (1.50) form the Lie algebra of $SU(3)$.

We now consider the operators

$$\hat{T}^a = a_i^\dagger t_{ij}^a a_j = a^\dagger t^a a \tag{1.51}$$

where t_{ij}^a denote matrix components of the (3×3) matrices t^a $(i, j = 1, 2, 3; a = 1, 2, \cdots, 8)$. In what follows, we shall see that these operators correspond to those that connect degenerate states of a three-dimensional harmonic oscillator. First, we find that $\hat{T}^a$ are hermitian,

$$\begin{aligned} (\hat{T}^a)^\dagger &= a_j^\dagger \, (t^a)_{ij}^* \, a_i \\ &= a_j^\dagger t_{ji}^a a_i = \hat{T}^a \end{aligned} \tag{1.52}$$

where we use the hermiticity of t^a, $(t^a)_{ij}^* = t_{ji}^a$. Secondly, we can check that $\hat{T}^a$ obey the $SU(3)$ algebra:

$$\begin{aligned} [\hat{T}^a, \hat{T}^b] &= [a_i^\dagger t_{ij}^a a_j, a_k^\dagger t_{kl}^b a_l] \\ &= a_i^\dagger t_{ij}^a [a_j, a_k^\dagger t_{kl}^b a_l] + [a_i^\dagger, a_k^\dagger t_{kl}^b a_l] t_{ij}^a a_j \end{aligned}$$

$$
\begin{aligned}
&= a_i^\dagger t_{ij}^a \delta_{jk} t_{kl}^b a_l - \delta_{il} a_k^\dagger t_{kl}^b t_{ij}^a a_j \\
&= a_i^\dagger t_{ij}^a t_{jl}^b a_l - a_k^\dagger t_{ki}^b t_{ij}^a a_j \\
&= a_i^\dagger \left(t_{ij}^a t_{jl}^b - t_{ij}^b t_{jl}^a \right) a_l \\
&= a^\dagger [t^a, t^b] a = i f^{abc} a^\dagger t^c a = i f^{abc} \hat{T}^c \,. \qquad (1.53)
\end{aligned}
$$

Lastly, we find that $[\hat{T}^a, a_i^\dagger a_i]$ vanishes:

$$
[\hat{T}^a, a^\dagger \mathbb{1} a] = a^\dagger [t^a, \mathbb{1}] a = 0 \qquad (1.54)
$$

where $\mathbb{1}$ denotes the (3×3) identity matrix. Equation (1.54) naturally leads to the relation $[\hat{T}^a, H] = 0$. This means that the total number $n = n_1 + n_2 + n_3 = a_i^\dagger a_i$ is invariant under the actions of $\hat{T}^a$. In other words, the operators $\hat{T}^a$ will not take the system out of degeneracy.

From these arguments, we find that the degenerate states of a three-dimensional isotropic harmonic oscillator form a representation of $SU(3)$. Let $|\alpha\rangle$ denote the degenerate states with $\alpha = 1, 2, \cdots, N = \frac{1}{2}(n+1)(n+2)$. Notice that a general form of $|\alpha\rangle$ is expressed as

$$
\frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} \frac{(a_3^\dagger)^{n_3}}{\sqrt{n_3!}} |0\rangle, \qquad n_1 + n_2 + n_3 = n \,. \qquad (1.55)
$$

We then consider the $(N \times N)$ matrices $\langle \alpha | \hat{T}^a | \beta \rangle$. These will obey the $SU(3)$ algebra. Thus a specific matrix representation of the $SU(3)$ algebra is given by $\langle \alpha | \hat{T}^a | \beta \rangle$.

The formalism introduced in this chapter can easily be generalized to the $SU(N)$ algebra in general. To briefly sketch the construction, we first consider operators of the form $T^a = a^\dagger t^a a$ where t^a $(a = 1, 2, \cdots, N^2 - 1)$ form a base of $(N \times N)$ hermitian traceless matrices, with $a_i^\dagger$ and a_i $(i = 1, 2, \cdots N)$ being creation and annihilation operators of an N-dimensional harmonic oscillator, respectively. Transitions between degenerate states of the N-dimensional harmonic oscillator are described by the operators $T^a = a^\dagger t^a a$, and these operators obey the $SU(N)$ algebra.

Lastly, we should notice that not all unitary irreducible representations of $SU(N)$ $(N \geq 3)$ are obtained from this formalism. As we can easily see from the construction of degenerate states, we can obtain only *symmetric* representations. Note that all representations of $SU(2)$ are symmetric. But this case of $N = 2$ is rather exceptional. All the unitary irreducible representations of the $SU(2)$ algebra are obtained from the harmonic oscillator analysis, but only the symmetric representations of the higher $SU(N)$ can be constructed in this way.

1.6 A bit of physics

So far, what we have discussed is the general formalism of harmonic oscillators. One might ask whether the results we find can be directly applied to any physi-

cal system. One such example is provided by the shell model of the nucleus. In this model, we consider the protons and neutrons in the nucleus to be considered as single particles moving in a central potential. The potential may be viewed as some sort of average self-consistently generated potential. The one-particle approximation to the dynamics is rather drastic, since we know there are strong forces of interaction between the particles. The hope is that the self-consistent potential captures some of the effect of the strong interactions. A good approximation to the potential is then taken as the three-dimensional harmonic oscillator potential,

$$V(r) = -V_0 + \frac{1}{2}M\omega^2(x^2 + y^2 + z^2) \tag{1.56}$$

From our analysis we can see that the degenerate states of nucleons (neutrons or protons) can be classified by representations of $SU(3)$, with degeneracy given by $N = \frac{1}{2}(n+1)(n+2)$. Thus the first few states have degeneracies $1, 3, 6, 10$. If we fill these, taking account of spin, we get the number of particles in the filled shells as 2, $2 + 6 = 8$, $2 + 6 + 12 = 20$, $2 + 6 + 12 + 20 = 40$, etc. The magic numbers, which are the numbers of protons or neutrons at which one gets better stability for the nucleus, are 2, 8, 20, 28, 50, 82, 126. We see that our counting does reproduce the first few magic numbers. This does not hold as we go to higher values of n; one reason for this is the increasing importance of spin-orbit coupling for the nucleons. If that effect is included, the shell model does give a good explanation of the magic numbers for nuclei.

Chapter 2

Bound and scattering states of the Hydrogen atom

2.1 The Hydrogen atom

A solution to the problem of the Hydrogen atom, *i.e.*, a quantum mechanical derivation of the energy spectrum of the Hydrogen atom, was first given by Pauli as early as 1926, before the better known solution using the Schrödinger equation. This problem is mathematically the same as the Kepler problem, the problem of planetary orbits. The Hamiltonian for either problem has an identical form

$$H = \frac{p^2}{2m} - \frac{\kappa}{r} \tag{2.1}$$

where $\vec{p}$ is a momentum of a planet (or an electron) with $p^2 = \vec{p}^2$, r is a radial distance from the sun (or the nucleus), and m is a mass of the planet (or the electron). κ can be written as

$$\kappa = \begin{cases} Ze^2 & \text{for Coulomb potentials,} \\ GMm & \text{for gravitational potentials.} \end{cases} \tag{2.2}$$

Of course, this is the idealized Hydrogen atom. The real atom has many corrections to be taken account of, such as the motion of the nucleus, relativistic corrections, spin-orbit interaction, etc., and even radiative corrections from quantum field theory. Nevertheless, (2.1) gives a very good first approximation.

In classical mechanics, the Kepler problem is exactly solvable. In this regard, it is worth mentioning an interesting theorem, called Bertrand's theorem, which states that *in three spatial dimensions, the only two central potentials for which all bounded orbits are also closed are the harmonic oscillator potential* $V = \frac{1}{2}\omega r^2$ *and the Coulomb/gravitational potential* $V = -\frac{\kappa}{r}$. We have already considered

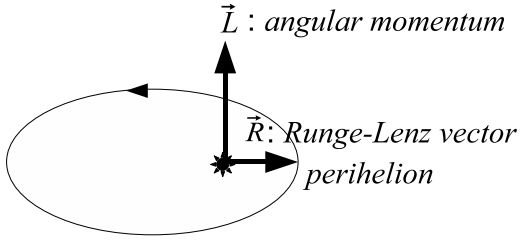

Figure 2.1: Angular momentum and Runge-Lenz vector in the Kepler problem

the harmonic oscillator, so it is entirely natural to take the Hydrogen atom as the next step.

The Hamiltonian (2.1) has obvious rotational invariance, so we know immediately that the angular momentum L_i, $i = 1, 2, 3$, is conserved. Further, in the classical Kepler problem of planetary orbits, we know that there is no precession of the orbits. This means that the placement of the orbit in the orbital plane is conserved; equivalently a vector from the origin to perihelion is conserved. This was already known to Laplace and more thoroughly analyzed by Runge and Lenz. The Runge-Lenz vector R_i and the angular momentum L_i form the conserved quantities for the problem (in addition to H itself, of course). These vectors are shown in Figure 2.1. The Runge-Lenz vector is usually defined as

$$R_i = \frac{1}{m}\epsilon_{ijk}p_j L_k - \kappa\hat{x}_i \tag{2.3}$$

where the momentum p_i and the position x_i are the basic phase space variables for the system. Also L_i is the angular momentum defined as $L_i = \epsilon_{ijk}x_j p_k$, and $\hat{x}_i$ is the unit vector $\hat{x}_i = \frac{x_i}{r}$. p_i and L_i do not commute each other in general. This means that the form of R_i in (2.3) is not hermitian. We thus define the quantum version of the Runge-Lenz vector by symmetrizing in p_j and L_k as

$$R_i = \frac{1}{2m}\epsilon_{ijk}(p_j L_k + L_k p_j) - \kappa\hat{x}_i\,. \tag{2.4}$$

This is manifestly hermitian.

In the following, we will consider the algebra of L_i, R_i and H. From the definition of L_i and the commutation relation $[x_i, p_j] = i\delta_{ij}$, we can easily find that

$$[L_i, H] = 0\,, \tag{2.5}$$
$$[L_i, x_j] = i\epsilon_{ijk}x_k\,, \tag{2.6}$$
$$[L_i, p_j] = i\epsilon_{ijk}p_k\,. \tag{2.7}$$

The last two equations may be viewed as stating that x_i and p_i transform as vectors under rotations. Using (2.4) and the relation $\epsilon_{ijk}\epsilon_{kjl} = -\epsilon_{ijk}\epsilon_{ljk} = -2\delta_{il}$, we can then rewrite (2.4) as

$$\begin{aligned} R_i &= \frac{1}{2m}\epsilon_{ijk}(p_j L_k + i\epsilon_{kjl}p_l + p_j L_k) - \kappa\hat{x}_i \\ &= \frac{1}{m}\epsilon_{ijk}p_j L_k - \frac{i}{m}p_i - \kappa\hat{x}_i\,. \end{aligned} \tag{2.8}$$

Commutation relations between the phase variables and the Hamiltonian can be computed as

$$[p_i, H] = \left[-i\frac{\partial}{\partial x^i}, -\frac{\kappa}{r}\right] = i\kappa\frac{\partial}{\partial x^i}\frac{1}{r} = -\frac{i\kappa}{r^2}\frac{x_i}{r} = -\frac{i\kappa\hat{x}_i}{r^2}\,, \tag{2.9}$$

$$[\hat{x}_i, H] = \left[\frac{x_i}{r}, \frac{p^2}{2m}\right] = \left[\frac{\nabla^2}{2m}, \frac{x_i}{r}\right]\,. \tag{2.10}$$

In order to compute (2.10), we need the following relations

$$\nabla_j\left(\frac{x_i}{r}\right) = \frac{\delta_{ij}}{r} - \frac{x_i}{r^2}\frac{x_j}{r} = \frac{\delta_{ij} - \hat{x}_i\hat{x}_j}{r}\,, \tag{2.11}$$

$$\begin{aligned} \nabla_j\nabla_j\left(\frac{x_i}{r}\right) &= -\frac{1}{r^2}(\delta_{ij} - \hat{x}_i\hat{x}_j)\hat{x}_j - \frac{1}{r}\frac{\delta_{ij} - \hat{x}_i\hat{x}_j}{r}\hat{x}_j - \frac{1}{r}\hat{x}_i\frac{\delta_{jj} - \hat{x}_j\hat{x}_j}{r} \\ &= -2\frac{\hat{x}_i}{r^2} \end{aligned} \tag{2.12}$$

where we use $\delta_{jj} = 3$ and $\hat{x}_j\hat{x}_j = 1$. Then (2.10) can be simplified as

$$\begin{aligned} [\hat{x}_i, H]\,\psi &= \frac{\nabla^2}{2m}\left(\frac{x_i}{r}\psi\right) - \frac{x_i}{r}\frac{\nabla^2\psi}{2m} \\ &= \frac{\nabla_j}{2m}\left[(\nabla_j\hat{x}_i)\psi + \hat{x}_i\nabla_j\psi\right] - \hat{x}_i\frac{\nabla^2}{2m}\psi \\ &= -\frac{1}{mr^2}\hat{x}_i\psi + \frac{1}{m}(\nabla_j\hat{x}_i)\nabla_j\psi \\ &= -\frac{\hat{x}_i}{mr^2}\psi + \frac{1}{m}\frac{\delta_{ij} - \hat{x}_i\hat{x}_j}{r}\,ip_j\,\psi \end{aligned} \tag{2.13}$$

where we introduce a wavefunction ψ for convenience. The commutator (2.10) is then written as

$$[\hat{x}_i, H] = -\frac{\hat{x}_i}{mr^2} + \frac{i}{mr}(p_i - \hat{x}_i\hat{x}\cdot p)\,. \tag{2.14}$$

From these results given above, we can explicitly show that R_i commutes with the Hamiltonian,

$$\begin{aligned} [R_i, H] &= \frac{1}{m}\epsilon_{ijk}[p_j, H]L_k - \frac{i}{m}[p_i, H] - \kappa[\hat{x}_i, H] \\ &= -\frac{1}{m}\epsilon_{ijk}\left(-\frac{i\kappa}{r^2}\hat{x}_j\right)L_k - \frac{i}{m}\left(-\frac{i\kappa}{r^2}\hat{x}_i\right) + \frac{\kappa}{mr^2}\hat{x}_i - \frac{i\kappa}{mr}(p_i - \hat{x}_i\hat{x}\cdot p) \end{aligned}$$

$$
\begin{aligned}
&= -\frac{i\kappa}{mr^2}\epsilon_{ijk}\hat{x}_j\epsilon_{kmn}x_m p_n - \frac{i\kappa}{mr}(p_i - \hat{x}_i\hat{x}\cdot p) \\
&= -\frac{i\kappa}{mr^2}\hat{x}_j x_m p_n(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) - \frac{i\kappa}{mr}(p_i - \hat{x}_i\hat{x}\cdot p) \\
&= -\frac{i\kappa}{mr^2}(\hat{x}_j x_i p_j - \hat{x}_j x_j p_i) - \frac{i\kappa}{mr}(p_i - \hat{x}_i\hat{x}\cdot p) \\
&= 0\,. \qquad (2.15)
\end{aligned}
$$

Thus, so far, we have the commutation rules

$$
\begin{aligned}
[R_i, H] = [L_i, H] &= 0\,, \\
[L_i, L_j] &= i\epsilon_{ijk}L_k\,. \qquad (2.16)
\end{aligned}
$$

From (2.6)-(2.8), we also find

$$
[L_i, R_j] = i\epsilon_{ijk}R_k\,. \qquad (2.17)
$$

This can be verified directly, or we could argue that R_i being a vector, this relation must hold true. To obtain the full algebra, we need to calculate one more commutator, $[R_i, R_j]$. This quantity must be antisymmetric in i, j, so it must be proportional to ϵ_{ijk}. We may thus write

$$
[R_i, R_j] = i\epsilon_{ijk}\Lambda_k\,. \qquad (2.18)
$$

which defines the hermitian operator Λ_k. We may think of this also as $i\Lambda_k = \epsilon_{ijk}R_iR_j$. In the following, we shall calculate this factor by use of (2.8).

$$
\begin{aligned}
\epsilon_{ijk}R_i &= \epsilon_{ijk}\left(\frac{1}{m}\epsilon_{iab}p_aL_b - \frac{i}{m}p_i - \kappa\hat{x}_i\right) \\
&= \frac{1}{m}(p_jL_k - p_kL_j) - \frac{i}{m}\epsilon_{ijk}p_i - \kappa\epsilon_{ijk}\hat{x}_i \\
i\Lambda_k = \epsilon_{ijk}R_iR_j &= \left(\frac{1}{m}(p_jL_k - p_kL_j) - \frac{i}{m}\epsilon_{ijk}p_i - \kappa\epsilon_{ijk}\hat{x}_i\right) \\
&\quad \times\left(\frac{1}{m}\epsilon_{jab}p_aL_b - \frac{i}{m}p_j - \kappa\hat{x}_j\right) \qquad (2.19)
\end{aligned}
$$

Since terms like $\epsilon_{ijk}p_ip_j$ and $\epsilon_{ijk}\hat{x}_i\hat{x}_j$ vanish, there are seven terms in (2.19). We shall calculate them one by one in the following.

$$
\begin{aligned}
\text{(Term1)} &= \frac{1}{m^2}(p_jL_k - p_kL_j)\epsilon_{jab}p_aL_b \\
&= \frac{1}{m^2}(p_jL_k\epsilon_{jab}p_aL_b - p_kL_j\epsilon_{jab}p_aL_b) \\
&= \frac{1}{m^2}\left[p_j\epsilon_{jab}(i\epsilon_{kam}p_m + p_aL_k)L_b - p_k\epsilon_{jab}(i\epsilon_{jam}p_m + p_aL_j)L_b\right] \\
&= \frac{1}{m^2}\left[ip_jp_mL_b(\delta_{jk}\delta_{bm} - \delta_{jm}\delta_{bk})\right.
\end{aligned}
$$

$$
\begin{aligned}
& - 2ip_k \underbrace{p_b L_b}_{0} + \frac{1}{2} p_k p_a \epsilon_{ajb} \, (L_j L_b - L_b L_j)] \\
&= -\frac{i}{m^2} p^2 L_k + \frac{1}{2m^2} p_k \underbrace{p_a \epsilon_{ajb} i \epsilon_{jbm} L_m}_{0} = -\frac{i}{m^2} p^2 L_k \qquad (2.20)
\end{aligned}
$$

$$
\begin{aligned}
(\text{Term2}) &= -\frac{i}{m^2}(p_j L_k - p_k L_j) p_j \\
&= -\frac{i}{m^2} p_j L_k p_j = -\frac{i}{m^2} p_j (p_j L_k + i \epsilon_{kjm} p_m) \\
&= -\frac{i}{m^2} p^2 L_k \qquad (2.21) \\
(\text{Term3}) &= -\frac{\kappa}{m}(p_j L_k - p_k L_j)\hat{x}_j \\
&= -\frac{\kappa}{m} p_j L_k \hat{x}_j = -\frac{\kappa}{m} p_j (\hat{x}_j L_k + i \epsilon_{kjm} \hat{x}_m) \\
&= -\frac{\kappa}{m} p_j \hat{x}_j L_k - i \frac{\kappa}{m} \epsilon_{kjm} \left(\hat{x}_m p_j - i \frac{\partial}{\partial x^j} \frac{x_m}{r} \right) \\
&= -\frac{\kappa}{m} p \cdot \hat{x} L_k + i \frac{\kappa}{mr} L_k \qquad (2.22) \\
(\text{Term4}) &= -\frac{i}{m} \epsilon_{ijk} p_i \epsilon_{jab} p_a L_b \\
&= \frac{i}{m} p_i (p_i L_k - p_k L_i) = \frac{i}{m} p^2 L_k \qquad (2.23) \\
(\text{Term5}) &= \frac{i\kappa}{m} \epsilon_{ijk} p_i \hat{x}_j = \frac{i\kappa}{m} \epsilon_{ijk} \left(\hat{x}_j p_i - i \frac{\partial}{\partial x^i} \frac{x_j}{r} \right) \\
&= -\frac{i\kappa}{mr} L_k \qquad (2.24) \\
(\text{Term6}) &= -\frac{\kappa}{m} \epsilon_{ijk} \hat{x}_i \epsilon_{jab} p_a L_b = \frac{\kappa}{m} \hat{x}_i (p_i L_k - p_k L_i) \\
&= \frac{\kappa}{m} \hat{x} \cdot p L_k - \frac{\kappa}{mr} \underbrace{x_i p_k L_i}_{i \delta_{ik} L_i} \\
&= \frac{\kappa}{m} \hat{x} \cdot p L_k - \frac{i\kappa}{mr} L_k \qquad (2.25) \\
(\text{Term7}) &= \frac{i\kappa}{m} \epsilon_{ijk} \hat{x}_i p_j = \frac{i\kappa}{mr} L_k \qquad (2.26)
\end{aligned}
$$

Putting together the terms in (2.20)-(2.26), we find

$$
\begin{aligned}
i\Lambda_k &= -i \frac{p^2}{m^2} L_k + \frac{\kappa}{m} (\hat{x} \cdot p - p \cdot \hat{x}) L_k \\
&= -i \frac{p^2}{m^2} L_k + \frac{2i\kappa}{mr} L_k \\
&= -i \frac{2}{m} \left(\frac{p^2}{2m} - \frac{\kappa}{r} \right) L_k
\end{aligned}
$$

$$= \quad -i\frac{2}{m}HL_k \tag{2.27}$$

where we use

$$\begin{aligned} \hat{x}\cdot p - p\cdot\hat{x} &= \frac{x_i}{r}p_i - p_i\frac{x_i}{r} \\ &= \frac{x_i}{r}p_i - i\frac{\partial}{\partial x^i}\left(\frac{x_i}{r}\right) - \frac{x_i}{r}p_i \\ &= i\left(\frac{3}{r} - \frac{x_i}{r^2}\frac{x_i}{r}\right) = \frac{i2}{r}\,. \end{aligned} \tag{2.28}$$

Thus, we finally find $\Lambda_k = -\frac{2H}{m}L_k$.

Summary

The algebra of observables of interest in the Kepler problem is then summarized as

$$\left\{\begin{array}{rcl} [L_i, L_j] &=& i\epsilon_{ijk}L_k \\ [L_i, R_j] &=& i\epsilon_{ijk}R_k \\ [R_i, R_j] &=& i\epsilon_{ijk}\left(-\frac{2H}{m}\right)L_k \\ [L_i, H] &=& [R_i, H] = 0 \end{array}\right. \tag{2.29}$$

The sign of H will be important in the further analysis of this algebra. In terms of the states of electrons in the Hydrogen atom, $H \leq 0$ corresponds to the bound states and $H > 0$ corresponds to the scattering states. The two cases have to be analyzed separately.

2.2 Bound states

In this section, we consider the case of $H \leq 0$ in the algebra (2.29). We first normalize the Runge-Lenz vector as

$$M_i = \sqrt{\frac{-m}{2H}}R_i\,. \tag{2.30}$$

The algebra of interest then becomes

$$\left\{\begin{array}{rcl} [L_i, L_j] &=& i\epsilon_{ijk}L_k \\ [L_i, M_j] &=& i\epsilon_{ijk}M_k \\ [M_i, M_j] &=& i\epsilon_{ijk}L_k \end{array}\right. \tag{2.31}$$

This is the $O(4)$ algebra, corresponding to the generators of rotations in four dimensions. We can rewrite this algebra in terms of

$$S_i = \frac{L_i + M_i}{2}\,, \quad T_i = \frac{L_i - M_i}{2}\,. \tag{2.32}$$

$$[S_i, T_j] \quad = \quad \frac{1}{4}\,[L_i + M_i, L_j - M_j]$$

$$= \frac{i}{4}\epsilon_{ijk}(L_k - M_k + M_k - L_k) = 0 \tag{2.33}$$

$$[S_i, S_j] = \frac{1}{4}[L_i + M_i, L_j + M_j]$$

$$= \frac{i}{4}\epsilon_{ijk}(L_k + M_k + M_k + L_k) = i\epsilon_{ijk}S_k \tag{2.34}$$

$$[T_i, T_j] = \frac{1}{4}[L_i - M_i, L_j - M_j]$$

$$= \frac{i}{4}\epsilon_{ijk}(L_k + L_k - M_k - M_k) = i\epsilon_{ijk}T_k \tag{2.35}$$

Thus the $O(4)$ algebra can be described by two independent (or mutually commute) angular-momentum like operators S_i, T_i.

To understand the spectrum of the Hydrogen atom we must obtain the Hamiltonian H in terms of S_i and T_i. For this, we first calculate R^2 as

$$\begin{aligned} R^2 &= \left(\frac{1}{m}\epsilon_{ijk}p_jL_k - \frac{i}{m}p_i - \kappa\hat{x}_i\right)\left(\frac{1}{m}\epsilon_{iab}p_aL_b - \frac{i}{m}p_i - \kappa\hat{x}_i\right) \\ &= \left[\frac{1}{m^2}p_jL_k(p_jL_k - p_kL_j) - \frac{i}{m^2}\epsilon_{ijk}p_jL_kp_i - \frac{\kappa}{m}\epsilon_{ijk}p_jL_k\hat{x}_i \right. \\ &\qquad \left. + \frac{i\kappa}{m}p\cdot\hat{x} - \frac{\kappa}{m}\epsilon_{iab}\hat{x}_ip_aL_b + \frac{i\kappa}{m}\hat{x}\cdot p + \kappa^2 - \frac{p^2}{m^2}\right] \\ &= \frac{1}{m^2}p^2L^2 + \frac{2p^2}{m^2} - \frac{2\kappa}{mr}L^2 + \frac{i\kappa}{m}(\hat{x}\cdot p - p\cdot\hat{x}) + \kappa^2 - \frac{p^2}{m^2} \\ &= \frac{p^2}{m^2}(L^2+1) - \frac{2\kappa}{mr}L^2 - \frac{2\kappa}{mr} + \kappa^2 \\ &= \left(\frac{p^2}{m^2} - \frac{2\kappa}{mr}\right)(L^2+1) + \kappa^2 \\ &= \frac{2}{m}H(L^2+1) + \kappa^2 \end{aligned} \tag{2.36}$$

where we have used (2.28) in simplifying the factor $(\hat{x}\cdot p - p\cdot\hat{x})$. Using (2.30), we can then express the Hamiltonian as

$$H = -\frac{m\kappa^2}{2}\frac{1}{M^2 + L^2 + 1}\,. \tag{2.37}$$

Going back to (2.32), we find

$$S^2 = \frac{1}{4}(L^2 + M^2 + L\cdot M + M\cdot L)\,. \tag{2.38}$$

The factors $L\cdot M$ and $M\cdot L$ vanish. This can easily be seen from

$$R\cdot L = \frac{1}{m}\epsilon_{ijk}p_jL_kL_i = \frac{1}{m}\epsilon_{ijk}p_j\frac{L_kL_i - L_iL_k}{2}$$

$$= \frac{i}{2m}\epsilon_{ijk}p_j\epsilon_{kim}L_m = 0\,. \tag{2.39}$$

Thus we find

$$S^2 + T^2 = \frac{L^2 + M^2}{2} \tag{2.40}$$

with $S^2 = T^2$. The Hamiltonian is therefore expressed as

$$H = -\frac{m\kappa^2}{2} \frac{1}{2(S^2 + T^2) + 1} \,. \tag{2.41}$$

Since S_i's obey the angular momentum algebra, we can write $S^2 + T^2$ as $S^2 + T^2 = 2S^2 = 2s(s+1)$ where $s = \frac{1}{2}q$ $(q = 0, 1, 2, \cdots)$, *i.e.*,

$$S^2 + T^2 \;=\; q\left(\frac{q}{2} + 1\right). \tag{2.42}$$

Thus the Hamiltonian can be rewritten as

$$H = -\frac{m\kappa^2}{2} \frac{1}{n^2} \qquad (n \equiv q + 1 = 1, 2, 3, \cdots). \tag{2.43}$$

This result gives eigenvalues of H, which correspond to the correct energy spectrum of a Hydrogen atom. This is obviously in agreement with the Bohr theory as well as the solution obtained from the Schrödinger equation. We can identify n as the principal quantum number.

Degeneracy

Since the angular momentum operator is written as $L_i = S_i + T_i$, we can count the degeneracy directly in terms of S_i and T_i. Eigenstates are labeled by $|s, m_s, t, m_t\rangle$ with $s = t$ and $m_s, m_t = s, s-1, \cdots, -s$. The number of states is $(2s+1)^2 = (q+1)^2 = n^2$, being the same as the result in quantum mechanics. The quantum number of L_i takes the values of $l = 0, 1, 2, \cdots, q = n - 1 (= 2s)$. Thus the degeneracy also agrees with what we know from the Schrödinger theory.

2.3 Scattering states

We now consider the case of $H > 0$ in the algebra (2.29). The normalized Runge-Lenz vector analogous to (2.30) is written as

$$N_i = \sqrt{\frac{m}{2H}} R_i \,. \tag{2.44}$$

The algebra of interest then becomes

$$\left\{ \begin{array}{rcl} [L_i, L_j] &=& i\epsilon_{ijk} L_k \\ [L_i, N_j] &=& i\epsilon_{ijk} N_k \\ [N_i, N_j] &=& -i\epsilon_{ijk} L_k \end{array} \right. \tag{2.45}$$

This forms an $O(3,1)$ algebra. The $O(3,1)$ group is the group of orthogonal linear transformations of x, y, z, t that keep the value of $x^2+y^2+z^2-t^2$ invariant. Thus the $O(3,1)$ algebra is also called the Lorentz algebra.

The N_i-part of the algebra can also be written as

$$\begin{array}{rcl}[L_i, iN_j] &=& i\epsilon_{ijk}(iN_k)\,,\\ [iN_i, iN_j] &=& i\epsilon_{ijk}L_k\end{array} \tag{2.46}$$

where $iN_i = \sqrt{-1}\,N_i$. Thus, in analogy with the previous section, we find that the operators

$$S_i^{(1)} = \frac{L_i + iN_i}{2}\,,\quad S_i^{(2)} = \frac{L_i - iN_i}{2} \tag{2.47}$$

give two copies of the $SU(2)$-type algebra ($i=1,2,3$). However, these operators are no longer hermitian. Elements of the transformation S generated by L_i, N_i can be parametrized as

$$S = \exp\left(i\theta^i L_i + i\alpha^i N_i\right) \tag{2.48}$$

for real parameters θ^i, α^i. A particular realization of the algebra (2.45) is given by $L_i = \frac{1}{2}\sigma_i$, $N_i = -\frac{i}{2}\sigma_i$, so that we may write

$$S = \exp\left(i\theta^i\frac{\sigma_i}{2} + \alpha^i\frac{\sigma_i}{2}\right) = \exp\left(i(\theta^i - i\alpha^i)\frac{\sigma_i}{2}\right). \tag{2.49}$$

This shows that we may view the transformations generated by the algebra (2.45) as a version of the $SU(2)$ transformations, but with complex parameters. This complexified $SU(2)$ group is called $SL(2,\mathbb{C})$.

Going back to (2.37), the Hamiltonian is now expressed as

$$H = \frac{m\kappa}{2}\frac{1}{N^2 - L^2 - 1}\,. \tag{2.50}$$

Introducing $S_\pm^{(1)} = S_1^{(1)} \pm iS_2^{(1)}$, we can consider the algebra of $S_i^{(1)}$ in terms of $S_\pm^{(1)}$ and $S_3^{(1)} = (L_3 + iN_3)/2$. We then define the "lowest" state $|\Omega\rangle$ by

$$S_-^{(1)}|\Omega\rangle = 0\,. \tag{2.51}$$

Action of $S_3^{(1)}$ on $|\Omega\rangle$ can be written as

$$S_3^{(1)}|\Omega\rangle = (a+ib)|\Omega\rangle \tag{2.52}$$

where a, b are real numbers corresponding to quantum number of the operator L_3 and N_3, respectively, on the lowest state. Notice that, since $[N_i, N_j]$ gives L_k, it is not consistent to set b equal to zero, for nonzero L_k. (If both L_i and N_i vanish on the state, it is invariant under the whole set of transformations and the representation is trivial.) We will therefore take $b \neq 0$, with $b \to 0$ as a limiting case. The quadratic Casimir operator of $S_i^{(1)}$ can be calculated as

$$S_i^{(1)} S_i^{(1)} = S_+^{(1)} S_-^{(1)} + {S_3^{(1)}}^2 - S_3^{(1)} \,. \tag{2.53}$$

This leads to the relation

$${S^{(1)}}^2 |\Omega\rangle = [(a+ib)^2 - (a+ib)]|\Omega\rangle \,. \tag{2.54}$$

Similarly, we have

$$S_3^{(2)} |\Omega\rangle = (a-ib)|\Omega\rangle \,, \tag{2.55}$$

$${S^{(2)}}^2 |\Omega\rangle = [(a-ib)^2 - (a-ib)]|\Omega\rangle \,. \tag{2.56}$$

Further, as seen before in (2.39), $L \cdot N = N \cdot L = 0$. Thus we obtain

$${S^{(1)}}^2 = {S^{(2)}}^2 = \frac{L^2 - N^2}{4} \,. \tag{2.57}$$

From (2.54) and (2.56), we can then solve for a and b:

$$\begin{aligned} (a+ib)^2 - (a+ib) &= (a-ib)^2 - (a-ib) \\ ib(2a-1) &= 0 \\ \Longrightarrow \; a &= \frac{1}{2} \;\; (b \neq 0) \end{aligned} \tag{2.58}$$

The other solution $b = 0$ gives a trivial solution discussed before. Using this result in (2.54) or (2.56), we find

$$N^2 - L^2 = 4b^2 + 1 \,. \tag{2.59}$$

The Hamiltonian (2.50) is therefore written as

$$H = \frac{m\kappa}{2} \frac{1}{(2b)^2} \; > \; 0 \,. \tag{2.60}$$

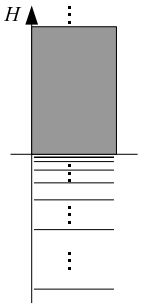

Figure 2.2: Energy spectra for bound states and scattering states

It is important to notice that there is no restriction on b, that is, b is continuous, taking any nonzero real value, *i.e.*, $-\infty < b < \infty$.

The representation we have obtained is unitary, with both L_i and N_i realized as hermitian operators. In particular, this means that $S_+^{(1)}$ and $S_-^{(2)}$ are conjugate to each other. There is no quantization requirement on the parameter b and because b is not quantized, H is not quantized either. The magnitude of b can be related to the impact parameter for the scattering states, with the sign of b distinguishing incoming and outgoing states in the scattering.

Since the parameter b is not quantized, if we consider the action of $S_+^{(1)}$, $S_+^{(2)}$ to build up the various states starting from $|\Omega\rangle$, the series does not terminate in general. We have infinite-dimensional representations. This is in conformity with a general theorem that unitary representations of a noncompact Lie group are infinite-dimensional. The group of transformations $O(3,1)$ (or $SL(2,\mathbb{C})$) is noncompact.

In this section, we found that there is no quantization of energy spectrum for scattering states. Combined with the results in the previous section, energy spectrum for the bound states ($H \leq 0$) and the scattering states ($H > 0$) can be schematically sketched as in Figure 2.2.

Chapter 3

Landau problem and quantum Hall effect

3.1 Introduction to QHE

In physics, the so-called Landau problem refers to the study of charged-particle dynamics in a uniform magnetic field. A realistic and important example of the Landau problem is given by quantum Hall effect (QHE) that has been observed in the motion of electrons inside a conducting material at low temperature. In the following two chapters, we shall make algebraic analyses of the QHE by making a certain semiclassical approximation.

We first begin with basics of QHE. The schematics of the experimental set-up are shown in Fig. 3.1. We are interested in how the Hall current J_2 is related

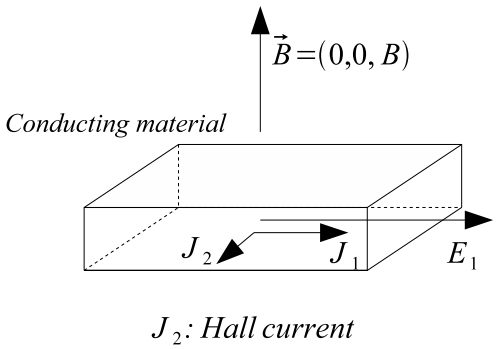

Figure 3.1: Quantum Hall effect at low temperature

to the electric field E_1. Generally, we have Ohm's law $J_i = \sigma_{ij} E_j$ $(i, j = 1, 2)$ where σ_{ij} is the conductivity tensor. Thus, in the present case, we have the following relations.

$$J_1 = \sigma_{11} E_1 \quad J_1\text{: normal current} \tag{3.1}$$
$$J_2 = \sigma_{21} E_1 \quad J_2\text{: Hall current} \tag{3.2}$$

The Hall conductivity σ_{21} must depend on B since the Hall current is due to the Lorentz force, which deflects the charge-carriers as they try to move along the direction of the applied electric field. Thus classically, the Hall current will be proportional to B giving $\sigma_{21} \sim B$, a straight line behavior for the graph of σ_{21} versus B, as shown in the first graph of Fig. 3.2. At low temperatures, what is observed is that σ_{21} is independent of B for some range of B, then makes a jump to a new value which is maintained for another range of B, then makes another jump to a new value and so on. This is illustrated in the second graph of Fig. 3.2. To restate the result, the graph of σ_{21} versus B shows plateaux. Further, the values of σ_{21} at the plateaux are quantized, hence the name quantum Hall effect. The integer QHE refers to the plateau-values being quantized in integer units of $e^2/2\pi$, e being the charge of the electron. The plateau-values can be fractionally quantized as well; this is the fractional QHE, which we shall consider in the next chapter.

In the following, we analyze the QHE with two approximations:

1. The dynamics of electrons along the x_3-axis is irrelevant; the system is effectively two-dimensional in (x_1, x_2)-plane with $\vec{B} = (0, 0, B)$ as in Fig. 3.1.

2. The edge potential inside the conducting material can be approximated by a two-dimensional harmonic oscillator.

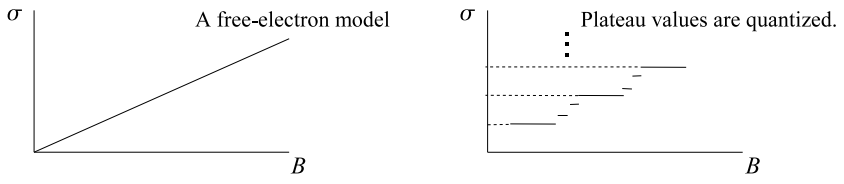

Figure 3.2: Linear σ-B relation for free electrons (left) and quantization of the Hall conductivity (right)

Thus the single-particle Hamiltonian of the system is given by

$$\begin{aligned} H &= \frac{(p_1 - eA_1)^2 + (p_2 - eA_2)^2}{2m} + \frac{m\omega^2}{2}(x_1^2 + x_2^2)\,, && (3.3)\\ A_1 &= -\frac{Bx_2}{2}\,, \quad A_2 = \frac{Bx_1}{2}\,, && (3.4)\\ B_3 &= \partial_1 A_2 - \partial_2 A_1 = B && (3.5) \end{aligned}$$

where m denotes an effective mass of the drifting electron and ∂_i $(i = 1, 2)$ denotes a derivative with respect to x_i. (The effective mass is usually denoted by m^*, but we will use just m, to avoid too much clutter in notation.) We now define

$$\Pi_i = p_i - eA_i\,. \tag{3.6}$$

Commutation relations between these operators are given by

$$\begin{aligned} [\Pi_i, \Pi_j] &= [p_i - eA_i, p_j - eA_j] \\ &= ie(\partial_i A_j - \partial_j A_i) = ieB\,\epsilon_{ij} \end{aligned} \tag{3.7}$$

where ϵ_{ij} is the rank-2 Levi-Civita tensor; $\epsilon_{12} = 1$, $\epsilon_{21} = -1$ and $\epsilon_{11} = \epsilon_{22} = 0$. Next we consider rewriting the Hamiltonian in terms of

$$\mathcal{A} = \frac{\Pi_1 + i\Pi_2}{\sqrt{2eB}}\,, \quad \mathcal{A}^\dagger = \frac{\Pi_1 - i\Pi_2}{\sqrt{2eB}}\,. \tag{3.8}$$

For these combinations, we find

$$\begin{aligned} [\mathcal{A}, \mathcal{A}^\dagger] &= \frac{1}{2eB}[\Pi_1 + i\Pi_2, \Pi_1 - i\Pi_2] = \frac{-i2}{2eB}[\Pi_1, \Pi_2] = 1\,, && (3.9)\\ \mathcal{A}^\dagger\mathcal{A} &= \frac{1}{2eB}(\Pi_1 - i\Pi_2)(\Pi_1 + i\Pi_2) \\ &= \frac{1}{2eB}(\Pi_1^2 + \Pi_2^2 + i[\Pi_1, \Pi_2]) = \frac{\Pi_1^2 + \Pi_2^2}{2eB} - \frac{1}{2}\,. && (3.10) \end{aligned}$$

Thus the Hamiltonian can be re-expressed as

$$\begin{aligned} H &= \frac{1}{2m}(\Pi_1^2 + \Pi_2^2) + \frac{m\omega^2}{2}(x_1^2 + x_2^2) \\ &= \underbrace{\frac{eB}{m}\left(\mathcal{A}^\dagger\mathcal{A} + \frac{1}{2}\right)}_{\equiv H_0} + \frac{m\omega^2}{2}(x_1^2 + x_2^2)\,. \end{aligned} \tag{3.11}$$

We will first consider the eigenstates of H_0, ignoring the harmonic oscillator potential for now. The potential energy $\frac{1}{2}m\omega^2(x_1^2 + x_2^2)$ will be included in section 3.3. For the Hamiltonian H_0, the ground state $|0\rangle$ is defined by

$$\mathcal{A}|0\rangle = 0\,. \tag{3.12}$$

Physical states are then constructed by the action of $\mathcal{A}^\dagger$'s on $|0\rangle$ and are easily seen to be

$$\mathcal{A}^\dagger|0\rangle\,,\,\frac{(\mathcal{A}^\dagger)^2}{\sqrt{2}}|0\rangle\,,\cdots\cdots\,,\,\frac{(\mathcal{A}^\dagger)^n}{\sqrt{n!}}|0\rangle\,. \tag{3.13}$$

The corresponding energy is given by

$$E_n = \frac{eB}{m}\left(n+\frac{1}{2}\right) \tag{3.14}$$

where $n = 0, 1, 2, \cdots$. These energy levels are called Landau levels. The lowest case $n = 0$ corresponds to the lowest Landau level (LLL) with zero-point energy

$$E_0 = \frac{eB}{2m} \equiv \omega_L\,. \tag{3.15}$$

This value ω_L is called the Larmor frequency. Energy separation in (3.14) is then given by $2\omega_L$.

Magnetic translations

We now introduce the operators

$$\widetilde{\Pi}_i = p_i + eA_i \tag{3.16}$$

where, as before, $A_i = -\frac{1}{2}Bx_j\epsilon_{ij}$. The commutation relations between Π_i and $\widetilde{\Pi}_j$ are given by

$$\begin{aligned}[\Pi_i, \widetilde{\Pi}_j] &= [p_i - eA_i, p_j + eA_j] \\ &= -ie\partial_i A_j - ie\partial_j A_i \,=\, -ie(\epsilon_{ji}+\epsilon_{ij})\frac{B}{2} \,=\, 0\,.\end{aligned} \tag{3.17}$$

From this, we can easily find that $[\widetilde{\Pi}_i, H_0] = 0$. The action of $\widetilde{\Pi}_i$'s on any physical states will thus lead to the degenerate states (in the absence of the oscillator potential). The commutators between $\widetilde{\Pi}_i$'s are calculated as

$$[\widetilde{\Pi}_i, \widetilde{\Pi}_j] = [p_i + eA_i, p_j + eA_j] = -ieB\,\epsilon_{ij}\,. \tag{3.18}$$

The operator $\widetilde{\Pi}_i$ is called the *magnetic translation operator*. Degeneracy of Landau levels is described by this operator. It is useful to visualize this in terms of the classical motion of the particle. Classically, the charged particle moves in a circle around the lines of force of the magnetic field and the trajectory of the form

$$\begin{aligned} x_1 - \tilde{x}_1 &= -\frac{\sqrt{2mE}}{eB}\cos\frac{eBt}{m}\,, \\ x_2 - \tilde{x}_2 &= \frac{\sqrt{2mE}}{eB}\sin\frac{eBt}{m}\,, \\ (x_1-\tilde{x}_1)^2 + (x_2-\tilde{x}_2)^2 &= \frac{2mE}{e^2B^2}\,. \end{aligned} \tag{3.19}$$

In this equation, $(\tilde{x}_1, \tilde{x}_2)$ are constants of integration. We see that the trajectory describes a circle centered at $(\tilde{x}_1, \tilde{x}_2)$ with a radius given by the energy E and the value of the magnetic field. The magnetic field is uniform, so there should be translational symmetry in the problem. This is reflected in the fact that the placement of the orbit in the (x_1, x_2)-plane is arbitrary, namely, any choice of $(\tilde{x}_1, \tilde{x}_2)$. The action of the magnetic translation operators corresponds to shifting these values.

From the point of view of dynamical variables, we have introduced new canonical pairs $\left(\frac{\Pi_1}{\sqrt{eB}}, \frac{\Pi_2}{\sqrt{eB}}\right)$ and $\left(\frac{\widetilde{\Pi}_1}{\sqrt{eB}}, \frac{\widetilde{\Pi}_2}{\sqrt{eB}}\right)$ in place of (x_i, p_i). The Hamiltonian depends on one pair $\left(\frac{\Pi_1}{\sqrt{eB}}, \frac{\Pi_2}{\sqrt{eB}}\right)$ while the other set constitutes a symmetry of the system. The degeneracy of any Landau level should thus be given, by the general theorem on symmetry, by a representation of the $\widetilde{\Pi}_i$ operators. This can be done at the operator level, but for most of what we want to do, since we will be discussing large numbers of fermions, a semiclassical analysis will suffice. We now turn to this question.

Semiclassical argument for degeneracy

We consider the degenerate states of the LLL. The number of states is generally given by

$$(\text{Number of states}) = \frac{(\text{Phase space volume})}{(2\pi\hbar)^d} \tag{3.20}$$

where d denotes the number of canonical pairs. In the present case, we are interested in the degeneracy due to the canonical pair $\left(\frac{\widetilde{\Pi}_1}{\sqrt{eB}}, \frac{\widetilde{\Pi}_2}{\sqrt{eB}}\right)$, so $d = 1$. Thus we have

$$(\text{Number of states}) = \frac{1}{2\pi} \frac{d\widetilde{\Pi}_2}{\sqrt{eB}} \frac{d\widetilde{\Pi}_1}{\sqrt{eB}} \tag{3.21}$$

where we set $\hbar = 1$. Now, classically, the Hamiltonian of the LLL is written as

$$H_0 = \frac{\Pi_1^2 + \Pi_2^2}{2m} \tag{3.22}$$

and the zero-point energy can be ignored. In the ground state, minimizing H_0, we get $\Pi_i = p_i - eA_i \approx 0$ $(i = 1, 2)$. This is adequate semi-classically, since the correction is just the zero-point energy. What this means is that, semi-classically, we may evaluate $\widetilde{\Pi}_i$ as

$$\widetilde{\Pi}_i = p_i + eA_i \approx 2eA_i = -eB\epsilon_{ij}x_j \,. \tag{3.23}$$

The number of states given by (3.21) can now be expressed as

$$(\text{Number of states}) \approx \frac{eB}{2\pi} dx_1 dx_2 \,. \tag{3.24}$$

We see that the degeneracy per unit area is therefore given by the fixed value $\frac{eB}{2\pi}$ in the semiclassical limit.

3.2 Integer QHE

We now consider a Landau level with all states filled with electrons, neglecting spin effects such as a Zeeman polarization. (If the magnetic field is strong, all electrons are in a spin-polarized state and the dynamics of the spin degrees of freedom is frozen. Therefore, it can be ignored for our considerations here.) As mentioned in the previous section, the QHE is realized in the quantization of Hall conductance σ_{ij} $(i, j = 1, 2)$ which is defined by

$$\langle J_i \rangle = \sigma_{ij} E_j \tag{3.25}$$

where $\langle J_i \rangle$ denotes the expectation value of the current J_i evaluated on the Landau level, which is occupied by a number of electrons. The main goal of this section is to calculate $\langle J_i \rangle$ and to see the quantization of σ_{ij} by use of the above-mentioned semiclassical analyses on Landau levels.

From the above results, we find that a total charge of degenerate states per unit area, *i.e.*, a charge density, is given by $\frac{e^2 B}{2\pi}$. If ν Landau levels are filled (ν being integer), then the total charge density of the state becomes

$$\langle J_0 \rangle = \langle \psi^{(0)} | J_0 | \psi^{(0)} \rangle = \nu \frac{e^2 B}{2\pi} \tag{3.26}$$

where $\psi^{(0)}$ is a wave function for the filled Landau levels of interest. ν is generally referred to as the "filling fraction", even though, in this case, it is an integer. (More generally, a formula similar to (3.26) holds with ν not necessarily an integer, as in the case of the fractional quantum Hall effect we discuss later.)

We now consider a perturbation due to a slight change of the magnetic field $A_i \longrightarrow A_i + \delta A_i$. Since the interaction term becomes $A_i J_i \longrightarrow A_i J_i + \delta A_i J_i$, the perturbation to the Hamiltonian is given by

$$H_{int} = \int d^2 y \delta A_i(y) J_i(y) \,. \tag{3.27}$$

Standard Rayleigh-Schrödinger perturbation theory tells us that the perturbed wave function can be expanded as

$$|\psi^{(0)}\rangle \longrightarrow |\psi^{(0)}\rangle + \frac{1}{E^{(0)} - H} H_{int} |\psi^{(0)}\rangle + \cdots \,. \tag{3.28}$$

where $E^{(0)}$ denotes the energy eigenvalue for $|\psi^{(0)}\rangle$, $H|\psi^{(0)}\rangle = E^{(0)}|\psi^{(0)}\rangle$. Note that we consider the first-order approximation in δA_i. Evaluated on this perturbed state, the charge density (3.26) is expressed as

$$\begin{aligned} & \left(\langle \psi^{(0)}| + \int \langle \psi^{(0)} | \delta A_i J_i \frac{1}{E^{(0)} - H} + \cdots \right) J_0 \left(|\psi^{(0)}\rangle + \frac{1}{E^{(0)} - H} \int \delta A_i J_i |\psi^{(0)}\rangle + \cdots \right) \\ = \; & \langle J_0 \rangle + \int d^2 y \delta A_i(y) \langle \psi^{(0)}| \left(J_i(y) \frac{1}{E^{(0)} - H} J_0(x) + J_0(x) \frac{1}{E^{(0)} - H} J_i(y) \right) |\psi^{(0)}\rangle + \cdots \,. \end{aligned}$$

(3.29)

where we omit the integral variables d^2y in the first line. Then in the first-order perturbation, $\delta\langle J_0(x)\rangle$ can be expressed as

$$\delta\langle J_0(x)\rangle = \int d^2y\ \delta A_i(y) F_i(x,y)\,, \tag{3.30}$$

$$F_i(x,y) = \langle\psi^{(0)}|\left(J_i(y)\frac{1}{E^{(0)}-H}J_0(x) + J_0(x)\frac{1}{E^{(0)}-H}J_i(y)\right)|\psi^{(0)}\rangle \tag{3.31}$$

Similarly, we can consider a perturbation due to $A_0 \longrightarrow A_0 + \delta A_0$. In this case, we can easily obtain

$$H_{int} = \int d^2y J_0(y)\delta A_0(y)\,, \tag{3.32}$$

$$\begin{aligned}\delta\langle J_i(x)\rangle &= \int d^2y \delta A_0(y)\langle\psi^{(0)}|\left(J_0(y)\frac{1}{E^{(0)}-H}J_i(x) + J_i(x)\frac{1}{E^{(0)}-H}J_0(y)\right)|\psi^{(0)}\rangle \\ &= \int d^2y \delta A_0(y) F_i(y,x)\end{aligned} \tag{3.33}$$

where $\delta\langle J_i(x)\rangle$ is evaluated in the first-order approximation in δA_0. From (3.26), we also know that

$$\langle J_0(x)\rangle = \nu\frac{e^2}{2\pi}(\partial_1 A_2 - \partial_2 A_1)\,, \tag{3.34}$$

so that we may write

$$\begin{aligned}\delta\langle J_0(x)\rangle &= \nu\frac{e^2}{2\pi}(\partial_1\delta A_2 - \partial_2\delta A_1) \\ &= \nu\frac{e^2}{2\pi}\int d^2y\left[\frac{\partial}{\partial x_1}\delta^{(2)}(x-y)\delta A_2(y) - \frac{\partial}{\partial x_2}\delta^{(2)}(x-y)\delta A_1(y)\right].\end{aligned} \tag{3.35}$$

Comparing (3.30) and (3.35), we see that we may identify $F_i(x,y)$ as given by

$$F_i(x,y) = -\frac{\nu e^2}{2\pi}\epsilon_{ij}\frac{\partial}{\partial x_j}\delta^{(2)}(x-y)\,. \tag{3.36}$$

Substituting this into (3.33), we can write $\delta\langle J_i(x)\rangle$ as

$$\begin{aligned}\delta\langle J_i(x)\rangle &= \int d^2y\delta A_0(y)\left(-\frac{\nu e^2}{2\pi}\right)\epsilon_{ij}\frac{\partial}{\partial y_j}\delta^{(2)}(y-x) \\ &= \frac{\nu e^2}{2\pi}\epsilon_{ij}\frac{\partial}{\partial x_j}\delta A_0(x)\end{aligned} \tag{3.37}$$

where we carry out a partial integration in the last step. We can therefore express $\langle J_i(x)\rangle$ as

$$\langle J_i(x)\rangle = \frac{\nu e^2}{2\pi}\epsilon_{ij}\frac{\partial A_0}{\partial x_j} = -\frac{\nu e^2}{2\pi}\epsilon_{ij}E_j \tag{3.38}$$

where we use $\frac{\partial A_0}{\partial x_j} = -E_j$. This is nothing but the relation $\vec{E} = -\nabla\phi$ where $\phi = A_0$ is the electrostatic potential. The result (3.38) gives an explicit form of (3.25). We have thus derived quantized conductivity

$$\sigma_{ij} = \nu \frac{e^2}{2\pi}\epsilon_{ij} \,. \tag{3.39}$$

The conductivity is quantized in units of $e^2/2\pi$, the quantization being given by the filling fraction ν.

Since ν takes integer values such as $1, 2, \cdots$, our result is for the integer QHE. Notice that σ_{ij} is independent of B. This suggests that the integer quantization of σ_{ij} is a non-perturbative effect. The integer QHE is a well established phenomenon and data from QHE experiments have been utilized to determine the Planck constant $\hbar$ or the fine structure constant $\alpha = (e^2/\hbar c)$ to a high degree of accuracy.

3.3 Laughlin wave function for integer QHE

In the previous section, we neglected the effects of the harmonic oscillator potential in the Hamiltonian. In this section, we will reinstate this potential and obtain a specific form of a LLL wave function Ψ_{LLL}. By use of Ψ_{LLL}, we then aim to obtain wave functions for integer QHE.

We first carry out the reparametrization of the Hamiltonian (3.11) with the complex variables

$$z = \sqrt{\frac{eB}{2}}(x_1 + ix_2) \,, \quad \bar{z} = \sqrt{\frac{eB}{2}}(x_1 - ix_2) \,. \tag{3.40}$$

Using these, we can express $\mathcal{A}$ as

$$\begin{aligned} \mathcal{A} &= \frac{\Pi_1 + i\Pi_2}{\sqrt{2eB}} = \frac{\left(p_1 + \frac{eB}{2}x_2\right) + i\left(p_2 - \frac{eB}{2}x_1\right)}{\sqrt{2eB}} \\ &= \frac{-i}{\sqrt{2eB}}\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right) - \frac{i}{2}\sqrt{\frac{eB}{2}}(x_1 + ix_2) \\ &= -i\left(\frac{\partial}{\partial \bar{z}} + \frac{z}{2}\right) \end{aligned} \tag{3.41}$$

where we use the relation

$$\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)\bar{z} = \left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)(x_1 - ix_2)\sqrt{\frac{eB}{2}} = \sqrt{2eB} \tag{3.42}$$

to define the derivative $(\partial/\partial\bar{z})$. Similarly we can obtain the conjugate operator $\mathcal{A}^\dagger$ as

$$\mathcal{A}^\dagger = -i\left(\frac{\partial}{\partial z} - \frac{\bar{z}}{2}\right) \,. \tag{3.43}$$

We now define new operators

$$
\begin{aligned}
C &= \mathcal{A} - \frac{i\alpha}{2} z = -i\left(\frac{\partial}{\partial \bar{z}} + \frac{1+\alpha}{2} z\right), && (3.44)\\
C^\dagger &= \mathcal{A}^\dagger + \frac{i\alpha}{2}\bar{z} = -i\left(\frac{\partial}{\partial z} - \frac{1+\alpha}{2}\bar{z}\right), && (3.45)\\
T &= \mathcal{A} + iz + \frac{i\alpha}{2} z = -i\left(\frac{\partial}{\partial \bar{z}} - \frac{1+\alpha}{2} z\right), && (3.46)\\
T^\dagger &= \mathcal{A}^\dagger - i\bar{z} - \frac{i\alpha}{2}\bar{z} = -i\left(\frac{\partial}{\partial z} + \frac{1+\alpha}{2}\bar{z}\right). && (3.47)
\end{aligned}
$$

where α is a constant parameter. Notice that the operators T and $T^\dagger$ can be interpreted as the extension of the magnetic-translation type of operators. Using these new operators, the Hamiltonian is expressed as

$$
\begin{aligned}
H &= 2\omega_L\left[\left(C^\dagger - \frac{i\alpha}{2}\bar{z}\right)\left(C + \frac{i\alpha}{2} z\right) + \frac{1}{2}\right] + \frac{\omega^2}{2\omega_L}\bar{z}z \\
&= 2\omega_L\left[C^\dagger C - i\frac{\alpha\bar{z}}{2} C + i\frac{\alpha}{2} C^\dagger z + \frac{\alpha^2}{4}\bar{z}z + \frac{1}{2}\right] + \frac{\omega^2}{2\omega_L}\bar{z}z \qquad (3.48)
\end{aligned}
$$

where ω_L is the Larmor frequency in (3.15). From (3.45) and (3.47), we can rewrite the factor $iC^\dagger z$ as

$$
\begin{aligned}
iC^\dagger z &= \left(\frac{\partial}{\partial z} - \frac{\bar{z}}{2} - \frac{\alpha}{2}\bar{z}\right) z \\
&= 1 + z\frac{\partial}{\partial z} - \frac{1+\alpha}{2}\bar{z}z \\
&= 1 + z\left(\frac{\partial}{\partial z} + \frac{1+\alpha}{2}\bar{z} - (1+\alpha)\bar{z}\right) \\
&= 1 + izT^\dagger - (1+\alpha)\bar{z}z\,. \qquad (3.49)
\end{aligned}
$$

Thus the Hamiltonian can be rewritten as

$$
H = 2\omega_L\left[\left(C^\dagger C - \frac{i\alpha\bar{z}}{2} C\right) + \frac{\alpha}{2}(1 + izT^\dagger - (1+\alpha)\bar{z}z) + \frac{\alpha^2}{4}\bar{z}z + \frac{1}{2}\right] + \frac{\omega^2}{2\omega_L}\bar{z}z\,. \qquad (3.50)
$$

From this expression, we find the factor $\bar{z}z$ cancels for

$$
\begin{aligned}
&-\omega_L\alpha(1+\alpha) + \frac{\alpha^2}{2}\omega_L + \frac{\omega^2}{2\omega_L} = 0 \\
&\Longrightarrow \quad \alpha = \frac{-\omega_L \pm \sqrt{\omega_L^2 + \omega^2}}{\omega_L} \qquad (3.51)
\end{aligned}
$$

Notice that we can require $\alpha \to 0$ as $\omega \to 0$. Thus the value of α is uniquely fixed by

$$
\alpha = \frac{-\omega_L + \sqrt{\omega_L^2 + \omega^2}}{\omega_L}\,. \qquad (3.52)
$$

In this case, the Hamiltonian (3.50) becomes

$$H = 2\omega_L \left(C^\dagger C - \frac{i\alpha\bar{z}}{2} C \right) + (1+\alpha)\omega_L + \alpha\omega_L z \left(\frac{\partial}{\partial z} + \frac{1+\alpha}{2}\bar{z} \right) \tag{3.53}$$

where the last term represents a magnetic-translation type operator. From this, we can define the state of LLL by

$$C|LLL\rangle = 0\,. \tag{3.54}$$

Thus the LLL wave function Ψ_{LLL} obeys

$$\left(\frac{\partial}{\partial \bar{z}} + \frac{1+\alpha}{2} z \right) \Psi_{LLL} = 0 \tag{3.55}$$

which leads to

$$\Psi_{LLL} = \exp\left(-\frac{1+\alpha}{2}\bar{z}z \right) f(z)\,. \tag{3.56}$$

$f(z)$ is an arbitrary holomorphic function. Using this expression, we find

$$\left(\frac{\partial}{\partial z} + \frac{1+\alpha}{2}\bar{z} \right) \Psi_{LLL} = \exp\left(-\frac{1+\alpha}{2}\bar{z}z \right) \frac{\partial f(z)}{\partial z}\,. \tag{3.57}$$

Schrödinger equation is then written as

$$H\Psi_{LLL} = e^{-\frac{1+\alpha}{2}\bar{z}z} \left[(1+\alpha)\omega_L + \alpha\omega_L z \frac{\partial}{\partial z} \right] f(z) = e^{-\frac{1+\alpha}{2}\bar{z}z} E f(z)\,. \tag{3.58}$$

Thus the equation for $f(z)$ is

$$(1+\alpha)\omega_L f(z) + \alpha\omega_L z \frac{\partial}{\partial z} f(z) = E f(z)\,. \tag{3.59}$$

This has the solution

$$f(z) = z^k, \qquad E = E_k = (1+\alpha)\omega_L + k\alpha\omega_L \quad (k = 0, 1, 2, \cdots). \tag{3.60}$$

Since the value of the parameter α is given by (3.52), we see that the the energy eigenvalues E_k behave as

$$\begin{aligned} E_k &\to \omega_L \qquad \text{as} \;\; \omega \to 0\,, \\ E_k &\approx (k+1)\omega \qquad \text{for} \;\; \omega \geq \omega_L\,. \end{aligned}$$

In other words, when the oscillator potential ω can be ignored, all eigenstates are degenerate, while in the large ω limit the eigenstates can be approximated to those of a two-dimensional oscillator. This gives a consistency check on our solution. The important point is that while the degeneracy is lifted as expected, and consistent with the electrons being confined inside the sample, the wave functions, apart from the pre-factor, are holomorphic in z. We can now use this to construct the many-electron wave function.

Laughlin wave function

From (3.60) we find that the LLL has multiple energy levels corresponding to $k = 0, 1, 2, \cdots$. We can thus consider the situation where N of the lowest of these states are filled by N electrons. Since we neglect spin, by Pauli principle, we allow one electron per state. The many-electron wave function must be antisymmetric under the exchange of any two, and hence it must be given by a Slater determinant. Thus it should be of the form

$$\Psi = \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \cdots & \psi_N(x_2) \\ \vdots & & \cdots & \vdots \\ \psi_1(x_N) & \psi_2(x_N) & \cdots & \psi_N(x_N) \end{vmatrix} \tag{3.61}$$

where $\psi_k(x_i)$ denotes a 1-particle wave function for i-th particle for the state labeled by k. In the present case, ψ_k $(k = 0, 1, 2, \cdots, N-1)$ are given by

$$f(z) = 1, z, z^2, \cdots, z^{N-1}. \tag{3.62}$$

Thus for two electrons, the corresponding determinant is then given by

$$\begin{vmatrix} 1 & z_1 \\ 1 & z_2 \end{vmatrix} = (z_2 - z_1) \tag{3.63}$$

where we ignore a normalization factor. Similarly, for three electrons we have

$$\begin{vmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{vmatrix} = (z_1 - z_2)(z_2 - z_3)(z_3 - z_1)\,. \tag{3.64}$$

For N electrons, one can generalize this argument to show that the determinant is given by

$$\begin{vmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{N-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{N-1} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & z_N & z_N^2 & \cdots & z_N^{N-1} \end{vmatrix} = \prod_{i>j}(z_i - z_j) = (-1)^{\frac{N(N-1)}{2}} \prod_{i<j}(z_i - z_j)\,. \tag{3.65}$$

This is called the Vandermonde determinant. This determinant is relevant to holomorphic part of the ground state wave functions. Including the pre-factor from the single-particle wave functions, we can take the many-electron wave function for the filled Landau level to be

$$\Psi = \mathcal{N} \exp\left(-\frac{1+\alpha}{2} \sum_{i=1}^{N} \bar{z}_i z_i\right) \prod_{1 \le i < j \le N} (z_i - z_j) \tag{3.66}$$

where $\mathcal{N}$ is a normalization constant. This wave function is known as the Laughlin wave function for integer QHE.

The highest occupied 1-particle state is proportional to $e^{-\frac{1+\alpha}{2}\bar{z}z}z^{N-1}$. The square of this is then given by

$$|(\text{highest 1-particle state})|^2 \sim e^{-(1+\alpha)\bar{z}z}(\bar{z}z)^{N-1}\,. \tag{3.67}$$

This value has a maximum at

$$-(1+\alpha)\bar{z}z + (N-1) = 0\,. \tag{3.68}$$

From (3.40), we find $\bar{z}z = \frac{eB}{2}r^2$ where r is the radius of a droplet of occupied states within the LLL state. The above relation then becomes

$$N - 1 = (1+\alpha)\frac{eB}{2\pi}(\pi r^2) = (1+\alpha)\frac{eB}{2\pi}(\text{area})\,. \tag{3.69}$$

The factor α is due to the oscillator potential ω. This is a more exact derivation of the result (3.24) relating N and the B and the area of the occupied states. In the limits of $\omega \to 0$ and N being large, (when the semiclassical argument is acceptable), this recovers the result in (3.24), *i.e.*,

$$N \approx \frac{eB}{2\pi}(\text{area})\,. \tag{3.70}$$

The harmonic oscillator potential is only to show how the degeneracy gets lifted. In a real system, the average potential inside the sample can be taken to be zero, except very close to the edge of the sample. Thus, if we try to model this by an oscillator potential, the corresponding ω is very small. The result (3.70) is thus obtained in a fairly straightforward way.

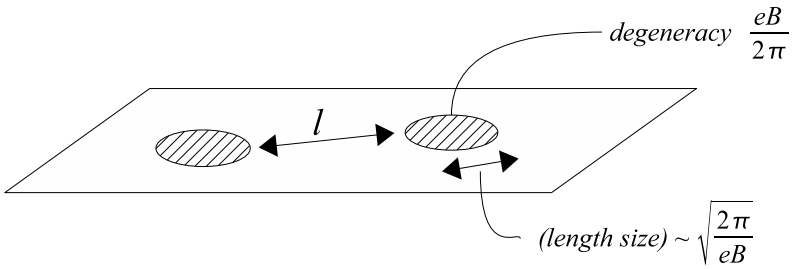

Figure 3.3: Quantum Hall droplets for integer QHE — In the integer QHE, the separation distance l is sufficiently large in comparison with the linear extent of the droplets that effects of Coulomb interactions can be neglected.

Notice that we have ignored electron-electron interactions. In the case of the integer QHE, the electron-electron interaction does not play a significant role. This is not to say that it can be neglected. Rather, because the states are all occupied and the exclusion principle prevents the totally free motion of any electron, the influence of the electrostatic repulsion on the dynamics is not significant. Ultimately, this is a question of the size of the Landau orbits and the range of the Coulomb interaction. (The latter has a finite screening length in the solid due to the ions and Thomas-Fermi effect.) If however, the individual Landau wave functions are broad enough to cause overlap of wave functions and hence a higher Coulomb energy, it can force a situation where not all states are occupied. (See Fig. 3.3, where the average separation distance is denoted by l. At some values of B, l is in the order of a screening length and Coulomb repulsion effects become important.) This is the case for fractional QHE which we shall discuss in the next section.

Chapter 4

Fractional QHE

In this chapter, we shall consider the fractional quantum Hall effect, where, as mentioned in the last chapter, the electron-electron Coulomb forces are important. Qualitatively, one would expect that the system would try to minimize the Coulomb energy by trying to keep the electrons away from each other. Thus compared to the integer Hall effect, say with $\nu = 1$, where essentially every available state in the lowest Landau level is occupied, we should expect that the filling fraction is less than 1. However, so far, there is no fully satisfactory derivation of the fractional quantum Hall states from a theoretical point of view.

The problem is easy enough to pose: We have a many-electron system with an external magnetic field and the electrostatic electron-electron repulsion; an overall confining potential (such as the oscillator potential we introduced in the last chapter) could be added to keep the electrons within the sample. Starting from this Hamiltonian, one would like to derive the fractional QHE state as the ground state. This would constitute a "first principles" derivation of fractional quantum Hall effect. In the absence of such a derivation, essentially, the entire analysis of the problem is done using the wave function which was suggested by Laughlin. In numerical simulations, Laughlin was able to argue for such a state to minimize the Coulomb energy. We will first discuss this Laughlin wave function and then go on to the dynamics of holes in fractional QHE and then present some algebraic considerations related to the hole dynamics. Our focus will be primarily on states with filling fractions of the form (odd integer)$^{-1}$, which can be discussed using the Laughlin wave function. Some of the algebraic features we want to discuss are most easily brought out in terms of these states. However, it should be kept in mind that these are not the only fractional quantum Hall states possible. There are also other many-body QHE states such as the Jain states and the Moore-Read state. We will not discuss these here.

4.1 Laughlin wave function for fractional QHE

Along the lines of analyses made in the previous chapter, fractional QHE is best understood by the following wave function proposed by Laughlin.

$$\Psi = (\text{prefactor}) \prod_{1\le i<j\le N} (z_i - z_j)^{2p+1} \tag{4.1}$$

where $p = 0, 1, 2, \cdots$. While it has not been possible to derive this from any kind of "first principles", there are some features which are worth emphasizing. Since electrons are fermions, the wave function should be antisymmetric under the exchange of any two positions, this is taken care of by the odd power $2p+1$ for the $(z_i - z_j)$ factors. Further, the Coulomb interaction would want to keep the electrons further from each other, so that we may expect a stronger vanishing condition as the positions of two electrons approach each other. Thus a higher power for the $(z_i - z_j)$ factors may be expected. This should also be reflected in the filling fraction, to which we now turn.

As in the case of integer QHE, the highest 1-particle state of the system is proportional to $z^{(N-1)(2p+1)}$, with its prefactor being $\exp(-\frac{\bar{z}z}{2})$, given that we can ignore potential factors. A maximum of the 1-particle wave function is then obtained at

$$-\bar{z}z + (N-1)(2p+1) = 0 \tag{4.2}$$

which leads to the relation

$$N - 1 = \frac{eB}{2\pi}\frac{1}{2p+1}(\text{area})\,. \tag{4.3}$$

These are analogous to the results of the integer QHE in (3.68) and (3.69). In the large N limit, a charge density of a Landau level filled with N electrons is expressed as

$$\langle J_0\rangle = \frac{Ne}{(\text{area})} \approx \nu\frac{e^2B}{2\pi}\,, \tag{4.4}$$

$$\nu = \frac{1}{2p+1} \tag{4.5}$$

($p = 0, 1, 2, \cdots$). Following the same line of reasoning as in Section 3.2, electric currents corresponding to (4.5) are given by

$$\langle J_i\rangle = -\nu\frac{e^2}{2\pi}\epsilon_{ij}E_j \tag{4.6}$$

where i, j take values of 1, 2. Since $\nu = \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \cdots$, excluding the case of $\nu = 1$, the above results show fractional quantization of the Hall conductivity. In reality, other quantization values, *e.g.*, $\nu = \frac{2}{5}$, are observed. There are even states with even denominators, for example, for $\nu = \frac{1}{2}$. We will not discuss these here, but refer the reader to the extensive literature on the subject.

It is useful to consider a picture of the filled states before we embark on considering the excited states. Recall that classically, we have Landau orbits, each having a certain radius. Quantum mechanically, we do not have well defined orbits, but each state is spread out a little with an approximate radius which is the same as the classical radius. Thus each state occupies an area equal to $2\pi/eB$. The number of states in an area A is thus $A/(2\pi/eB) = (eB/2\pi)A$. This is in agreement with the counting of states we have given. Further, we can only have one electron for each state due to the exclusion principle. With even a small additional potential, it is preferable to occupy states close to the minimum values of the potential, so the occupied states form a contiguous region, with one electron per state. This is thus a quantum Hall droplet of electrons. The exclusion principle prevents the compression of this droplet, so that we have an incompressible droplet.

It is now easy to see that there are two types of excitations possible for the quantum Hall droplet. The first type would consist of deformations of the boundary of the droplet, keeping the total area fixed, since that is proportional to the number of electrons we have. These are the so-called edge excitations. They can be studied in a simple hydrodynamic formulation where we consider the edge-deformations of a liquid droplet which is incompressible.

The other type of excitation would involve taking an electron from deeper inside the droplet and putting it outside of the droplet. This leaves a hole inside the droplet. This is the hole excitation. These two types of excitations are shown in Fig. 4.1.

Hole excitation

For integer QHE, the ground-state wave function is proportional to the Vandermonde determinant,

$$\begin{vmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{N-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{N-1} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & z_N & z_N^2 & \cdots & z_N^{N-1} \end{vmatrix} = \prod_{i>j} (z_i - z_j)\,. \tag{4.7}$$

This is the Slater determinant of the fermions, with 1, z, z^2, etc., being the single-particle states. Consider a hole at the center of a droplet formed from these. The state at the center corresponds to 1, and hence we can write a wave function for the system with hole by omitting this state from the Slater determinant. However, we do have N electrons still, and the lowest energy configuration would be to put the electron which was removed into the next

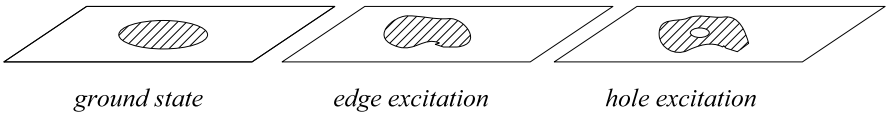

Figure 4.1: Excitations of quantum Hall droplets

available state, which is z^N. Thus the new state would be described by

$$\begin{vmatrix} z_1 & z_1^2 & z_1^3 & \cdots & z_1^N \\ z_2 & z_2^2 & z_2^3 & \cdots & z_2^N \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ z_N & z_N^2 & z_N^3 & \cdots & z_N^N \end{vmatrix} = z_1 z_2 \cdots z_N \prod_{i>j} (z_i - z_j)\,. \tag{4.8}$$

This gives the hole excitation for the QHE droplet. This wave function for the hole-excited state can also be written as

$$\Psi_{hole} = \prod_{i=1}^{N} (z_i - w) \Psi_{Laughlin} \tag{4.9}$$

where $\Psi_{Laughlin}$ is the Laughlin wave function for integer QHE shown in (3.66) and w denotes the position of a hole. The position of the hole was $z = 0$ for the case in (4.8) since we took the electron out of the single-particle state $\exp(-\frac{1}{2} z\bar{z})$.

In a way similar to this, for fractional QHE, we can then express a wave function of a 1-hole excited state by

$$\Psi_{hole} = \mathcal{N} \exp\left(-\frac{1}{2} \sum_{i=1}^{N} \bar{z}_i z_i \right) \prod_{i=1}^{N} (z_i - w) \prod_{1 \le i < j \le N} (z_i - z_j)^{2p+1} \tag{4.10}$$

where $\mathcal{N}$ is a normalization coefficient. Notice that we ignore potential factors in the above expression.

For the integer case (4.9), we can easily put an electron back to the system; this can be carried out simply by shifting N to $N + 1$. For the fractional case (4.10), however, the powers of $(z_i - w)$ and $(z_i - z_j)$ do not match to each other. This suggests that we need as many as $(2p + 1)$ holes at same position in order to fill the hole with an electron. A charge per hole is then given by a fractional charge $\frac{e}{2p+1}$. Notion of fractional charges therefore naturally arises in a system of fractional QHE.

4.2 Effective theory of fractional QHE

In this section, we consider an effective theory of fractional QHE using the following action

$$S = \int d^3x \left[\frac{k}{4\pi} \epsilon^{\mu\nu\alpha} a_\mu \partial_\nu a_\alpha + a_\mu \left(j^\mu - \frac{e}{2\pi} B^\mu \right) \right] \tag{4.11}$$

where a^μ ($\mu = 0, 1, 2$) is a new auxiliary field and j^μ denotes a hole current. k is a constant to be determined later. The field strength B^μ is defined by $B^\mu = \epsilon^{\mu\nu\alpha} \partial_\nu A_\alpha$. (We are using a three-dimensional covariant notation now. B^0 is the magnetic field, B^i will refer to the electric fields.) The electromagnetic-coupling term in (4.11) can be more explicitly written out as

$$\begin{aligned} -\frac{e}{2\pi} \int d^3x \; a_\mu \epsilon^{\mu\nu\alpha} \partial_\nu A_\alpha &= \frac{e}{2\pi} \int d^3x \; \partial_\nu a_\mu \epsilon^{\mu\nu\alpha} A_\alpha \\ &= -\frac{e}{2\pi} \int d^3x \; \epsilon^{\alpha\mu\nu} \partial_\mu a_\nu A_\alpha \\ &\equiv \int d^3x \; A_\alpha J^\alpha \,, \qquad (4.12) \\ J^\alpha &= -\frac{e}{2\pi} \epsilon^{\alpha\mu\nu} \partial_\mu a_\nu \qquad (4.13) \end{aligned}$$

where J^μ denotes the electromagnetic current.

Variation of a_μ in (4.11) leads to the equation of motion

$$\frac{k}{2\pi} \epsilon^{\mu\nu\alpha} \partial_\nu a_\alpha + j^\mu - \frac{e}{2\pi} B^\mu = 0 \,. \tag{4.14}$$

From (4.13) and (4.14), we then find

$$\begin{aligned} J^\mu &= \frac{e}{k} \left(j^\mu - \frac{e}{2\pi} B^\mu \right) \\ &= \frac{e}{k} j^\mu - \frac{e^2}{2\pi k} \epsilon^{\mu\nu\alpha} \partial_\nu A^\alpha \,. \qquad (4.15) \end{aligned}$$

The first term shows that the charge per hole is e/k. Thus, from the argument given earlier, we can determine k as

$$k = 2p + 1 \quad (p = 1, 2, \cdots) \tag{4.16}$$

for fractional QHE.

Dynamics of a hole

We can discuss dynamics of a hole by introducing time-dependence to the position of a hole, *i.e.*, $w^\mu = w^\mu(t)$ ($\mu = 0, 1, 2$, $t = x_0$). The hole current j^μ is written as

$$j^\mu = \dot{w}^\mu \delta^{(2)}(\vec{x} - \vec{w}(t)) \tag{4.17}$$

where $\vec{x}$ and $\vec{w}$ are 2-dimensional vectors $\vec{x} = (x_1, x_2)$, $\vec{w} = (w_1, w_2)$; we shall denote these without arrows when it is obvious. Using this current in (4.11), the effective action for one-hole dynamics is then expressed as

$$S_{hole} = \int d^3x \frac{k}{4\pi} \epsilon^{\mu\nu\alpha} a_\mu \partial_\nu a_\alpha + \int a_\mu(w) \dot{w}^\mu dt + \int \frac{m\dot{w}^2}{2} dt \,. \tag{4.18}$$

We have also added a kinetic energy term for the motion of the hole. Variation of this action with respect to a_μ gives the equation of motion

$$\frac{k}{2\pi} \epsilon^{\mu\nu\alpha} \partial_\nu a_\alpha + \dot{x}^\mu \delta^{(2)}(x - w) = 0 \,. \tag{4.19}$$

For $\mu = 0$, this becomes

$$\frac{k}{2\pi} (\partial_1 a_2 - \partial_2 a_1) + \delta^{(2)}(x - w) = 0 \,. \tag{4.20}$$

It is convenient to introduce complex coordinates by

$$z = x_1 + i x_2 \,, \quad \bar{z} = x_1 - i x_2 \,, \tag{4.21}$$

$$\partial_z = \tfrac{1}{2}(\partial_1 - i\partial_2) \,, \quad \partial_{\bar{z}} = \tfrac{1}{2}(\partial_1 + i\partial_2) \,, \tag{4.22}$$

$$a_z = \tfrac{1}{2}(a_1 - i a_2) \,, \quad a_{\bar{z}} = \tfrac{1}{2}(a_1 + i a_2) \,. \tag{4.23}$$

The expression $(\partial_1 a_2 - \partial_2 a_1)$ is given in terms of these combinations as

$$\begin{aligned} \partial_1 a_2 - \partial_2 a_1 &= (\partial_z + \partial_{\bar{z}}) \left(\frac{a_{\bar{z}} - a_z}{i} \right) - \frac{\partial_{\bar{z}} - \partial_z}{i} (a_z + a_{\bar{z}}) \\ &= \frac{2}{i} (\partial_z a_{\bar{z}} - \partial_{\bar{z}} a_z) \,. \end{aligned} \tag{4.24}$$

The equation of motion (4.20) takes the form

$$\partial_z a_{\bar{z}} - \partial_{\bar{z}} a_z = -i \frac{\pi}{k} \delta^{(2)}(x - w) \,. \tag{4.25}$$

Using the relations

$$\partial_z \frac{1}{\bar{z} - \bar{w}} = \pi \delta^{(2)}(x - w) \,, \quad \partial_{\bar{z}} \frac{1}{z - w} = \pi \delta^{(2)}(x - w) \,, \tag{4.26}$$

we can easily find that a solution of (4.25) is given by

$$a_{\bar{z}} = -\frac{i}{2k} \frac{1}{\bar{z} - \bar{w}} \,, \quad a_z = \frac{i}{2k} \frac{1}{z - w} \,. \tag{4.27}$$

We now go back to the action for the hole and write it in a gauge where $a_0 = 0$, to get the dynamics for a hole at the position $z (\neq w)$ as given by the action

$$S_{hole} = \int dt \left(\frac{1}{2} m \dot{\bar{z}} \dot{z} + a_z \dot{z} + a_{\bar{z}} \dot{\bar{z}} \right)$$

$$= \int dt \left(\frac{1}{2}m\dot{x}^2 + a_1\dot{x}^1 + a_2\dot{x}^2\right). \tag{4.28}$$

The canonical momenta for this action are given by

$$p_i = m\dot{x}_i + a_i \equiv -i\frac{\partial}{\partial x_i}\,. \tag{4.29}$$

In terms of the complex combinations we have been using, this can be written as

$$\begin{aligned} m\dot{z} &= \frac{2}{i}\left(\partial_{\bar{z}} - ia_{\bar{z}}\right), \\ m\dot{\bar{z}} &= \frac{2}{i}\left(\partial_z - ia_z\right). \end{aligned} \tag{4.30}$$

We will now show how this can be interpreted as a new representation of the Heisenberg algebra, different from the usual Schrödinger representation.

4.3 Heisenberg algebra and fractional spins

The Heisenberg algebra is generally expressed by the commutation rules

$$\begin{aligned} &[\hat{x}_i^{(\alpha)}, \hat{x}_j^{(\beta)}] = 0 \\ &[\hat{x}_i^{(\alpha)}, \hat{p}_j^{(\beta)}] = i\,\delta_{ij}\,\delta^{\alpha\beta} \\ &[\hat{p}_i^{(\alpha)}, \hat{p}_j^{(\beta)}] = 0 \end{aligned} \tag{4.31}$$

where $\alpha, \beta = 1, 2, \cdots, N$ and i, j denote the spatial indices. With regard to this algebra, we have already mentioned the Stone–von Neumann theorem which states that the representation of this algebra is unique (up to unitary equivalence) if N is finite and the underlying phase space of all (x, p) is simply connected. The two premises are important. For example, if N is not finite, one can have distinct representations of the algebra which correspond to different phases; thus phase transitions take place in the $N \to \infty$ limit. (This limit is also often referred to as the thermodynamic limit.) This happens in the BCS theory of superconductivity. We shall consider this case in a separate chapter.

When there are holes in space, the assumption of simple connectivity fails. This, we shall now show, happens for the hole excitation states in the fractional QHE.

We shall first see how there is only one representation of the algebra (4.31) if the space of the x's is simply connected. Since the x's commute among themselves, we can find a representation in which the x-operators are all simultaneously diagonalized. This is obviously the standard representation in quantum mechanics given by

$$\hat{x}^i \to x^i\,, \quad \hat{p}_i \to -i\frac{\partial}{\partial x^i}\,. \tag{4.32}$$

This is the Schrödinger representation which clearly obeys the algebra (4.31). Another solution to the algebra is given by

$$\hat{x}^i \to x^i\,, \quad \hat{p}_i \to -i\frac{\partial}{\partial x^i} + a_i(x)\,. \tag{4.33}$$

where $a_i(x)$ has to be restricted by $[\hat{p}_i, \hat{p}_j] = 0$, *i.e.*,

$$\left[-i\frac{\partial}{\partial x^i} + a_i(x)\,, -i\frac{\partial}{\partial x^j} + a_j(x)\right] = -i\left(\frac{\partial a_j}{\partial x^i} - \frac{\partial a_i}{\partial x^j}\right) = 0\,. \tag{4.34}$$

If x^i are defined on simply connected space, then a_i can be parametrized as

$$a_i = \partial_i \Lambda \tag{4.35}$$

where Λ is an arbitrary function of x^i. In this case, the representation (4.33) is expressed as

$$\hat{x}^i = U^{-1}x^i U\,, \quad \hat{p}_i = U^{-1}\left(-i\frac{\partial}{\partial x^i}\right)U\,, \tag{4.36}$$

$$U = \exp(i\Lambda)\,. \tag{4.37}$$

Thus the representation (4.33) is equivalent to the Schrödinger representation (4.32) via unitary transformations. In this sense, the algebra has only one unitary representation.

If the space is not simply connected, that is, if the space has noncontractible loops, then the condition (4.34) has nontrivial solutions other than (4.35). More precisely, one can show that $\partial_i a_j - \partial_j a_i = 0$ *has a nontrivial solution if and only if the space has noncontractible loops.* For example, we can choose

$$a_z = ic\frac{1}{z-w}\,, \quad a_{\bar{z}} = -ic\frac{1}{\bar{z}-\bar{w}} \tag{4.38}$$

where c is a constant. For hole dynamics in the fractional QHE, we can set $c = \frac{1}{2k} = \frac{1}{2(2p+1)}$ where $p = 1, 2, \cdots$. By use of (4.26), we then find $\partial_z a_{\bar{z}} - \partial_{\bar{z}} a_z = 0$ on $\mathbb{R}^2 - \{w\}$. This space corresponds to the configuration space of one-hole dynamics in the fractional QHE. The key for us is that the point w is to be excluded from space in the sense that other particles do not have access to it. For example, the wave function vanishes if any z_i for any electron is equal to w. We can re-express the solution (4.38) as

$$\begin{aligned} &a_z = \partial_z \Lambda\,, \quad a_{\bar{z}} = \partial_{\bar{z}}\Lambda\,, \\ &\Lambda(z,w) = ic\left[\log(z-w) - \log(\bar{z}-\bar{w})\right]\,. \end{aligned} \tag{4.39}$$

In analogy with (4.36), we might also try to define a set of transformed operators $\hat{x}^i = U^{-1}x^i U$, $\hat{p}_i = U^{-1}(-i\partial/\partial x^i)U$, where

$$U(z,w) = e^{i\Lambda(z,w)}\,. \tag{4.40}$$

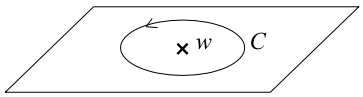

Figure 4.2: Contour C for the unitary transformation of U

The point, however, is that $U(z,w)$ is not a single-valued function on the space.

Consider moving along a curve C around the point w as in Fig. 4.2. The factor $\log(z-w)$ in (4.39) shifts by $2\pi i$ so that after one traversal of the curve, $U \to e^{-4\pi i c}U$. Alternatively,

$$U(e^{2\pi i}z, e^{2\pi i}w) = e^{-4\pi i c}U(z,w)\,. \tag{4.41}$$

This means that U is not a single-valued function unless $c = \frac{n}{2}$ where n is integer. For the case of hole dynamics, $c = \frac{1}{2(2p+1)}$, so that this condition is generally not obtained, except for $p = 0$ which corresponds to the integer Hall effect.

We have the result that $\partial_i a_j - \partial_j a_i = 0$ on the space under consideration. Thus the phase change in U is unaltered by deformations of the curve C. It is a topological invariant. Now, if space were simply connected, we can use a series of small deformations to shrink the curve to a point, where upon it is clear that there should be no change in U. Thus, for a simply connected space, the allowed U's must be single-valued. However, if we do not have simple connectivity, the the allowed U's do not need to be single-valued; there is no contradiction. However, a nonsingle-valued U does not qualify as an acceptable unitary transformation. For that the U's must be proper functions on the space, which means that it should have a unique value for each point. It must be single-valued.

The situation is now clear. For a simply connected space, we must have a unique representation (up to unitary equivalence), which is given by the Schrödinger representation. For the case of hole dynamics in fractional QHE, the representation given by (4.29) is not unitarily equivalent to the Schrödinger representation.

Two-hole dynamics

As discussed in (4.20), a_z and $a_{\bar{z}}$ obey the equation

$$-\frac{ik}{\pi}(\partial_z a_{\bar{z}} - \partial_{\bar{z}} a_z) + j_0 = 0 \tag{4.42}$$

where $j_0 = \delta^{(2)}(x-w)$ for a single hole system. For two holes, j_0 is given by

$$j_0 = \delta^{(2)}(x-w_1) + \delta^{(2)}(x-w_2)\,. \tag{4.43}$$

Since (4.42) is linear in the a's, a solution in this case is given by

$$a_{\bar{z}} = -\frac{i}{2k}\left(\frac{1}{\bar{z}-\bar{w}_1} + \frac{1}{\bar{z}-\bar{w}_2}\right), \tag{4.44}$$

$$a_z = \frac{i}{2k}\left(\frac{1}{z-w_1} + \frac{1}{z-w_2}\right). \tag{4.45}$$

Dynamics of two holes (at w_1 and w_2) can thus be described by the action

$$S = \int dt\left[\frac{m}{2}\dot{\bar{w}}_1\dot{w}_1 + \frac{m}{2}\dot{\bar{w}}_2\dot{w}_2 + a_{w_1}\dot{w}_1 + a_{\bar{w}_1}\dot{\bar{w}}_1 + a_{w_2}\dot{w}_2 + a_{\bar{w}_2}\dot{\bar{w}}_2\right] \tag{4.46}$$

where we remove singularities at two poles, *i.e.*,

$$a_{w_1} = \frac{i}{2k}\frac{1}{w_1-w_2}\,, \quad a_{w_2} = \frac{i}{2k}\frac{1}{w_2-w_1}\,, \tag{4.47}$$

$$a_{\bar{w}_1} = -\frac{i}{2k}\frac{1}{\bar{w}_1-\bar{w}_2}\,, \quad a_{\bar{w}_2} = -\frac{i}{2k}\frac{1}{\bar{w}_2-\bar{w}_1}\,. \tag{4.48}$$

We can interpret the singularities as something like Coulomb self-interactions so that the elimination of these singularities is physically reasonable. As in the case of one-hole dynamics, momentum operators of a hole are expressed as

$$m\dot{w}_1 = \frac{2}{i}\left(\frac{\partial}{\partial\bar{w}_1} - ia_{\bar{w}_1}\right) \equiv \frac{2}{i}\overline{\mathcal{D}}_1\,, \tag{4.49}$$

$$m\dot{\bar{w}}_1 = \frac{2}{i}\left(\frac{\partial}{\partial w_1} - ia_{w_1}\right) \equiv \frac{2}{i}\mathcal{D}_1\,, \tag{4.50}$$

$$\overline{\mathcal{D}}_1 = \frac{\partial}{\partial\bar{w}_1} - \frac{\partial}{\partial\bar{w}_1}\left[\frac{1}{2k}\log(\bar{w}_1-\bar{w}_2) - \frac{1}{2k}\log(w_1-w_2)\right], \tag{4.51}$$

$$\mathcal{D}_1 = \frac{\partial}{\partial w_1} - \frac{\partial}{\partial w_1}\left[\frac{1}{2k}\log(\bar{w}_1-\bar{w}_2) - \frac{1}{2k}\log(w_1-w_2)\right]. \tag{4.52}$$

Similar results can be obtained for the other hole,

$$\overline{\mathcal{D}}_2 = \frac{\partial}{\partial\bar{w}_2} - \frac{\partial}{\partial\bar{w}_2}\left[\frac{1}{2k}\log(\bar{w}_2-\bar{w}_1) - \frac{1}{2k}\log(w_2-w_1)\right], \tag{4.53}$$

$$\mathcal{D}_2 = \frac{\partial}{\partial w_2} - \frac{\partial}{\partial w_2}\left[\frac{1}{2k}\log(\bar{w}_2-\bar{w}_1) - \frac{1}{2k}\log(w_2-w_1)\right]. \tag{4.54}$$

The Hamiltonian of the system is given by

$$\mathcal{H} = \frac{m}{2}\dot{\bar{w}}_1\dot{w}_1 + \frac{m}{2}\dot{\bar{w}}_2\dot{w}_2\,. \tag{4.55}$$

In terms of $\overline{\mathcal{D}}_i$ and $\mathcal{D}_i$ $(i = 1, 2)$, this is written as

$$\mathcal{H} = -\frac{2}{m}\left(\mathcal{D}_1\overline{\mathcal{D}}_1 + \mathcal{D}_2\overline{\mathcal{D}}_2\right) \equiv \mathcal{H}_{\mathcal{D}_1} + \mathcal{H}_{\mathcal{D}_2}\,. \tag{4.56}$$

Notice that, as in (4.49) and (4.50), the momentum operators are defined by $m\dot{w}_i = -i2\overline{\mathcal{D}}_i$ and $m\bar{w}_i = -i2\mathcal{D}_i$.

We now consider a Schrödinger equation relevant to $\mathcal{H}_{\mathcal{D}_1}$, introducing an ansatz for the wave function, $\Psi = e^F\Phi$. The derivative of Ψ with respect to $\overline{\mathcal{D}}_1$ becomes

$$\begin{aligned}\overline{\mathcal{D}}_1\Psi &= \left(\frac{\partial}{\partial\bar{w}_1} - \frac{\partial}{\partial\bar{w}_1}\left[\frac{1}{2k}\log(\bar{w}_1 - \bar{w}_2) - \frac{1}{2k}\log(w_1 - w_2)\right]\right)e^F\Phi \\ &= e^F\left(\frac{\partial}{\partial\bar{w}_1}\Phi + \frac{\partial F}{\partial\bar{w}_1}\Phi\right. \\ &\qquad \left. -\frac{\partial}{\partial\bar{w}_1}\left[\frac{1}{2k}\log(\bar{w}_1 - \bar{w}_2) - \frac{1}{2k}\log(w_1 - w_2)\right]\Phi\right). \end{aligned} \tag{4.57}$$

Identifying F as

$$F(w_1, w_2) = \frac{1}{2k}\left[\log(\bar{w}_1 - \bar{w}_2) - \log(w_1 - w_2)\right]\,, \tag{4.58}$$

we can then reduce the Schrödinger equation $\mathcal{H}_{\mathcal{D}_1}\Psi = E_1\Psi$ to $\mathcal{H}_{\partial_1}\Phi = E_1\Phi$ where

$$\mathcal{H}_{\partial_1} = -\frac{2}{m}\frac{\partial}{\partial w_1}\frac{\partial}{\partial\bar{w}_1} \tag{4.59}$$

is an ordinary Hamiltonian for a free particle and Φ denotes an ordinary Schrödinger wave function for a hole. The wave function Ψ is therefore expressed as

$$\Psi(w_1, w_2) = \exp\left(\frac{1}{2k}\big[\log(\bar{w}_1 - \bar{w}_2) - \log(w_1 - w_2)\big]\right)\,\Phi\,. \tag{4.60}$$

An exchange of the hole positions can be carried out by a π-rotation of the two points and then translating the positions back to w_1 and w_2, as shown in Fig. 4.3. Notice that translations do not affect the form of Ψ. Under this operation, the wave function transforms as

$$\Psi(w_2, w_1) = \exp\left(\frac{1}{2k}(-i\pi - i\pi)\right)\;\Psi(w_1, w_2) = \exp\left(-i\frac{\pi}{k}\right)\;\Psi(w_1, w_2) \tag{4.61}$$

where, as mentioned earlier, $k = 2p + 1$ $(p = 1, 2, \cdots)$ for the fractional QHE. For the integer QHE, we have $p = 0$ corresponding to $k = 1$ so that (4.61) is consistent with the fact that a hole is a fermion with spin $\frac{1}{2}$. For $k > 1$, however, (4.61) shows that holes in the fractional QHE obey "fractional statistics".

In two spatial dimensions, it is possible to have fractional values for spin for a particle. The usual quantization for spin in terms of integer or half-integer is because the angular momentum operators do not commute among themselves

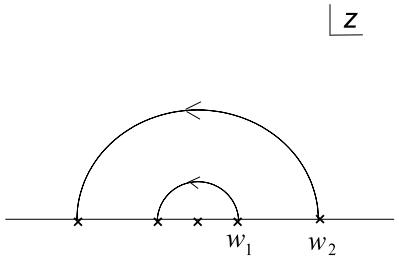

Figure 4.3: π-rotation on complex plane for the exchange of hole positions

and because we need a unitary representation. In two spatial dimensions, there is only one rotation and the issue of noncommutativity of the operators does not arise. Thus fractional values for spin are possible. This holds true even in a Lorentz-invariant theory, for which the rotation generator and the two generators of Lorentz transformations do not commute. This is because of the noncompact nature of the Lorentz group (and the full Poincaré group including translations), which allows unitary representations with fractional spin.

There is a spin-statistics connection in two spatial dimensions as well, according to which the result we have shown means that holes have fractional spins or that they are "anyons". These results can be considered as a consequence of the different representation of the Heisenberg algebra that arises from the nonsimply connected nature of the space of interest.

Chapter 5

Quarks and the spectroscopy of hadrons

5.1 Symmetry of hadrons

Fundamental particles, as we know them today, consist of four different types, the quarks, the leptons, the gauge bosons and the Higgs particle. Quarks and leptons are fermions and can be considered the basic building blocks from which all matter is made. The gauge bosons are the particles which mediate forces; they are necessary, among other things, to provide the binding forces for quarks and leptons so that composite particles and bound states can be built up. Finally, there is the Higgs particle which is needed for giving masses to some of the gauge particles and to the quarks and leptons. Up to now, six species (or flavors as they are called) of quarks and six species of leptons are known to exist. The six flavors of quarks are called up (u), down (d), strange (s), charm (c), top (t) and bottom (b). Bound states of quarks are called hadrons; they are subject to the strong nuclear forces (or the chromodynamic forces) and are the strongly interacting particles we see in the laboratory. In this chapter, we consider the spectroscopy of hadrons. There are two types of hadrons. One is called the mesons which are made of a quark and an anti-quark ($Q\overline{Q}$) and the other is called the baryons which consist of three quarks (QQQ). Protons (which are uud bound states) and neutrons (udd) are examples of the baryons. According to Particle Data Group (as of year 2013), masses of quarks are given by $m_u \approx 2.3$, $m_d \approx 4.8$, $m_s \approx 95$, $m_c \approx 1275$, $m_b \approx 4180$ and $m_t \approx 173000$ in the unit of MeV.

In constructing the bound states of the heavy quarks (c, b, t), it is adequate to treat them nonrelativistically as a good first approximation, with corrections which can then be added in. This is because the binding energy for such bound states, as measured by meson masses, is small compared to the masses of the

quarks. Thus the spectroscopy of heavy-quark hadrons can be analyzed by a simple Schrödinger equation. The interquark interaction can also be approximated by a nonrelativistic potential.

However, for hadrons made of the light quarks (u, d, s), a nonrelativistic approach is not applicable. In comparing the mass of the proton (938 MeV) with the masses of the up and down quarks which make up the proton, it is clear that the binding energy is far greater than the masses; the quarks are in ultra-relativistic motion. Not only is the relativistic dynamics important, but also the concept of potential, which is classical and non-relativistic, breaks down. In fact, single particle dynamics is also not adequate since there is enough energy to create a sea of virtual and real light quarks. The spectroscopy of light-quark hadrons then needs to be analyzed by some other means. One promising approach is to make numerical simulations by use of powerful computers. Another approach is to use symmetries of hadrons. We shall discuss the latter in this chapter.

According to the standard model of particle physics, quarks obtain masses via the spontaneous symmetry breaking of the electroweak symmetry. The field theoretic realization of this is called the Higgs mechanism. In the symmetry analysis we do not need the details of the Higgs mechanism. What is crucial for our analysis is the simple fact that we can write down an effective Lagrangian for quarks (including mass terms) as long as the energy level of interest is well below the electroweak scale of 246 GeV, the energy level at which the electroweak symmetry breaking takes place. The effective Lagrangian for light quarks is then expressed as

$$\mathcal{L}(Q) = \overline{Q}\, i\gamma \cdot (\partial - igA)\, Q + \overline{Q} \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} Q \tag{5.1}$$

where Q and $\overline{Q}$ are defined as

$$Q = \begin{pmatrix} u \\ d \\ s \end{pmatrix}, \quad \overline{Q} = \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix}. \tag{5.2}$$

An explicit form of the factor $\gamma \cdot (\partial - igA)$ in (5.1) is given by $\gamma^\mu(\partial_\mu - igt^a A^a_\mu)$ where $\mu = 0, 1, 2, 3$ is the space-time index and t^a ($a = 1, 2, \cdots, 8$) are the color matrices. (There is a color index for each species of quarks which also takes three possible values, say 1, 2, 3, which is not explicitly shown; thus u should really read u^i, $i = 1, 2, 3$, etc. $\{t^a\}$ denote 3×3 matrices (8 of them) which form a basis for the Lie algebra of $SU(3)$. They act on the color index of the quarks. This $SU(3)$ is the color symmetry group, which is not broken and which we will not explicitly show here. The $SU(3)$ to be introduce below is a flavor $SU(3)$ and is totally distinct from this group.) Further, in (5.1), γ^μ, g and A^a_μ denote the Dirac gamma matrices, the coupling constant for strong interactions, and the gluon fields, respectively. There are also terms involving gluons and electroweak interactions but they are irrelevant for our symmetry analysis.

We now consider a unitary transformation of Q, which can mix the different flavors u, d, s. This is realized by an action of a (3×3) unitary matrix U on Q:

$$Q \to Q' = UQ\,, \quad \overline{Q} \to \overline{Q}' = \overline{Q}U^\dagger\,. \tag{5.3}$$

Since $U^\dagger U = 1$, the first term in (5.1) is invariant under this transformation. It does not distinguish between different flavors of quarks, hence it is proportional to the identity matrix in the flavor labels. On the other hand, the second term breaks the $U(3)$ symmetry because quark masses are not identical. As mentioned before, they are related by $m_s \gg m_u \approx m_d$. We can therefore analyze the mass splittings among the mesons and baryons in terms of the breaking of this symmetry. Precursor of this idea can be found in a theory of *crystal field splitting* formulated by H. Bethe in the late 1920's.

Actually, if we neglect all masses, there is a large symmetry $U(3)_L \times U(3)_R$ since the Dirac operator can be split into left and right chiral components. In other words, we can write

$$\mathcal{L}(Q) = \overline{Q}\, i\gamma\cdot(\partial - igA)\, Q = \overline{Q}_L\, i\gamma\cdot(\partial - igA)\, Q_L + \overline{Q}_R\, i\gamma\cdot(\partial - igA)\, Q_R \tag{5.4}$$

where $Q_L = \frac{1}{2}(1+\gamma_5)Q$, $Q_R = \frac{1}{2}(1-\gamma_5)$. Thus it is possible to carry out independent $U(3)$ transformations on these components,

$$Q'_L = U_L\, Q_L, \qquad Q'_R = U_R\, Q_R \tag{5.5}$$

with the U in (5.3) being an element of the subgroup defined by $U_L = U_R = U$. This chiral symmetry is actually only $SU(3)_L \times SU(3)_R \times U(1)$, for reasons related to a quantum anomaly; but this will not concern us here. Further the curial symmetry $SU(3)_L \times SU(3)_R$ is spontaneously broken by the strong gluon interactions. The relevant energy scale of this chiral symmetry breaking is around 1 GeV. So the effective Lagrangian we are using should be valid for processes with transfer momenta below about 1 GeV.

The mass term in (5.1) breaks the chiral symmetry, and even the diagonal subgroup defined by $U_L = U_R = U$, explicitly. Thus the situation is a spontaneous symmetry breaking at a scale of about 1 GeV along with an explicit symmetry breaking due to the quark masses and mass differences, which is at a lower scale, $m_s \lesssim 150$ MeV. In the following, we discuss how we can realize this symmetry structure in quantum theory.

5.2 Realization of symmetry in quantum theory

We will now return to the consideration of symmetry in quantum theory making a refinement of the theorem 1.3 from Chapter 1. We start with the hamiltonian H and let $U(\theta)$ denote a unitary transformation which is a symmetry of the theory so that

$$U^\dagger(\theta)\, H\, U(\theta) = H\,, \qquad \text{or} \quad [H, U(\theta)] = 0\,. \tag{5.6}$$

The parameters θ (there may be several of them) labeling the transformations may be continuous as for a continuous symmetry (such as rotations or the $SU(3)$ transformations mentioned above) or discrete as for discrete symmetries such as parity (in which case we have only one U different form identity with $U^2 = 1$). Given two such transformations $U(\theta)$ and $U(\theta')$, we see that the product $U(\theta)\,U(\theta')$ is also, obviously, a symmetry,

$$[U(\theta)\,U(\theta')]^\dagger\; H\,U(\theta)\,U(\theta') = U^\dagger(\theta')\,U^\dagger(\theta)\,H\,U(\theta)\,U(\theta') = U^\dagger(\theta')\,H\,U(\theta') = H\,. \tag{5.7}$$

The the composition of two symmetries (given by their multiplication as operators) is a symmetry. Thus starting with a set of U's and taking products, we can eventually get a set of U's which is closed under multiplication; all of the elements of this set will commute with H. The identity transformation exists as the identity operator. $U(\theta)$, being unitary, has an inverse given by $U^\dagger(\theta)$. Finally, operator multiplication is associative. Thus all the conditions defining a group are satisfied and we can state that the set of all U's which commute with H form a group, which we may designate as the symmetry group G of the physical system.

Now consider the action of the symmetry group on the ground state $|\psi_0\rangle$ of the Hamiltonian. There are two possibilities for this,

$$\begin{array}{lll} 1. & U(\theta)\,|\psi_0\rangle = 0 & \text{Wigner realization,} \\ 2. & U(\theta)\,|\psi_0\rangle \neq 0 & \text{Goldstone realization.} \end{array} \tag{5.8}$$

First we consider the Wigner realization. Since the ground state is invariant under the symmetry, we can go on to look at the excited states. Let $|\alpha\rangle$ denote an eigenstate of H with eigenvalue E_α, *i.e.*,

$$H\,|\alpha\rangle = E_\alpha\,|\alpha\rangle\,. \tag{5.9}$$

Since $U(\theta)$ commutes with H, we find

$$H\,U(\theta)|\alpha\rangle = E_\alpha\,U(\theta)|\alpha\rangle \tag{5.10}$$

so that $U(\theta)|\alpha\rangle$ is an eigenstate of H with the same eigenvalue; in other words $U(\theta)|\alpha\rangle$ is degenerate with $|\alpha\rangle$. By considering all elements $U(\theta)$ of the symmetry group G, we will obtain a number of degenerate states which are all connected by the symmetry transformation. By construction, the subspace V of the Hilbert space spanned by these states $U(\theta)|\alpha\rangle$ for all $U(\theta)$ (including $|\alpha\rangle$ which corresponds to $U = 1$) is invariant under the action of the $U(\theta)$'s. (Individual states change but are transformed to the state within V.) Further, there cannot be a smaller subspace $V' \subset V$ which is invariant under G since V was constructed by including all states which are related to each other by transformations in G. Thus we have a unitary irreducible representation of G on the subspace V. Let $\{|i\rangle\}$ denote an orthonormal basis for V. Then the matrices $\langle i|U|j\rangle$ will reproduce the group composition laws under multiplication as matrices.

This argument may be extended to all eigenstates of H, so that they can all be grouped into multiplets, the states within each multiplet having the same energy. Different irreducible representations can have different energy eigenvalues. Symmetry does not give any reason why they should have the same eigenvalue, so this is the generic situation. However, it could also happen that there are several distinct irreducible multiplets which happen to have the same energy. This was explained in Chapter 1 with the example of the excited states in the Hydrogen atom.

We now turn to the Goldstone realization. In this case, we have to study the properties of $U(\theta)|\psi_0\rangle \neq 0$. The first question is whether this state is normalizable, so that it is an element of the Hilbert space. Assume it is normalizable. In this case, it will have the same eigenvalue for the Hamiltonian as $|\psi_0\rangle$. Then we can construct a state

$$|\Psi_0\rangle = \sum_\theta U(\theta)\,|\psi_0\rangle \tag{5.11}$$

where the summation (or integration) is over all values of θ. This state is invariant under transformations in G since

$$U(\theta')\,|\Psi_0\rangle = \sum_\theta U(\theta')\,U(\theta)\,|\psi_0\rangle = \sum_\theta U(\tilde\theta)\,|\psi_0\rangle = \sum_{\tilde\theta} U(\tilde\theta)\,|\psi_0\rangle = |\Psi_0\rangle\,. \tag{5.12}$$

Thus we have constructed a ground state that is invariant under the symmetry and the situation reverts to the case of the Wigner realization. (There are combinations of $U(\theta)|\psi_0\rangle$ which are orthogonal to the state $|\Psi_0\rangle$ in (5.12), but these have generally higher energies.) So the truly different case is when $U(\theta)\,|\psi_0\rangle$ is not normalizable. In that case, $U(\theta)\,|\psi_0\rangle$ is not an element of the Hilbert space anymore. In this case, we cannot have a unitary realization of the symmetry group at the level of the states. We have to choose a ground state which is not symmetric and then when we build excited states by acting on the ground state with suitable operators, the higher states also have no symmetry. We still have the symmetry at the level of the operator commutation rules, H still commutes with U, but it is not realized on the states. Thus the breakdown of symmetry is only because of the ground state not having the symmetry. This situation is known as spontaneous symmetry breaking.

A more general situation is that we can have a subgroup of transformations, say $h \in H \subset G$, which are still a symmetry of the ground state in the sense that $U(h)|\psi_0\rangle = 0$. In this case, we have a Wigner realization of the subgroup and states can be classified by irreducible representations of H. But the transformations in G which are not in H are not realized unitarily. In this case, we say the the symmetry is spontaneously broken from G to H.

If the symmetry which is spontaneously broken is a continuous symmetry, the parameters θ^a can be varied continuously. (There may be several parameters, so we label them by $a = 1, 2$, etc.) Taking $\theta^a = 0$ as corresponding to the identity, we may consider infinitesimal transformations of the form

$$U(\theta) = \exp(i\theta^a\,Q_a) \approx 1 + i\theta^a\,Q_a\ . \tag{5.13}$$

This corresponds to some group element $g = e^{i\theta^a t_a} \in G$, where t_a are the generators of the infinitesimal transformations. And Q_a in (5.13) are operators corresponding to the generators t_a. These generators will in general satisfy some algebra

$$[t_a, t_b] = C^c_{ab}\, t_c \tag{5.14}$$

where C^c_{ab} are the structure constants of the Lie algebra of the group G. Correspondingly, in terms of the operators Q_a we will have

$$[Q_a, Q_b] = C^c_{ab}\, Q_c \tag{5.15}$$

with the same structure constants, showing that the Q's give a representation of the algebra (5.14).

In the Goldstone realization, we can consider Q_a acting on the ground sate. The result that $U(\theta)$ is not normalizable then translates into the statement that $\langle\psi_0|Q_a|\psi_0\rangle$ is infinite. This statement is consistent with the non-normalizability of $|\alpha\rangle = Q_a|\psi_0\rangle$ by virtue of the inequality

$$\langle\psi_0|\psi_0\rangle\langle\alpha|\alpha\rangle \geq |\langle\psi_0|\alpha\rangle|^2 = |\langle\psi_0|Q_a|\psi_0\rangle|^2\,. \tag{5.16}$$

In terms of the action of U, we will have

$$\langle\psi_0|U(\theta)|\psi_0\rangle = 0 \tag{5.17}$$

so that transitions between states which may be thought of as connected by the symmetry transformation are not obtained.

For the action of Q_a on $|\psi_0\rangle$ to diverge and give a non-normalizable state, it will be necessary to have an infinite number of degrees of freedom, or consider systems in the thermodynamic limit of the number of degrees of freedom tending to infinity. Thus the Goldstone realization is in the realm of quantum field theory.

We may summarize the results on the realizations of symmetry as follows.

Theorem 5.1 *Let G denote the symmetry group of a physical system in the sense that the action of G leaves the Hamiltonian unchanged. Such a symmetry can be realized in two modes:*

1. *The ground state $|\psi_0\rangle$ is invariant under the action of G; this is the Wigner realization*
2. *The ground state is not invariant under the action of G, i.e., the symmetry is spontaneously broken. This is the Goldstone realization.*

In the Wigner realization, the eigenstates of the Hamiltonian of a quantum system can be classified into irreducible representations of the symmetry transformations, the states within each irreducible representation being degenerate. In the Goldstone realization, the symmetry cannot be realized unitarily; as a result, the states do not show the symmetry structure.

The breaking of the chiral $SU(3) \times SU(3)$ symmetry by strong interactions, which was allied to above, is an example of the Goldstone realization. Another case is superconductivity, which will be considered in the next chapter.

Now, coming back to the hadronic problem, we see that in the Wigner realization, the spectrum of the theory or the states describing mesons and baryons should fall into multiplets which carry irreducible representations of the $U(3)$ algebra. So we will now consider some of the irreducible representations of the $SU(3)$ group, or equivalently the $SU(3)$ algebra.

5.3 Irreducible representations of $SU(3)$

The group $SU(3)$, as the name indicates, is defined as the the set of all 3×3 unitary matrices U of unit determinant, $\det U = 1$. It is useful to think of this as acting as linear transformations of a vector space. For this, we first consider a 3-component complex column vector

$$\phi_i = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} \tag{5.18}$$

$(i = 1, 2, 3)$. A linear transformation of this by the action of U is given by $\phi' = U\phi$. In terms of the components, this is written as

$$\phi'_i = \sum_{k=1}^{3} U_{ik}\phi_k \,. \tag{5.19}$$

The 3×3 matrix U with matrix elements U_{ik} (and determinant 1) is the defining representation of $SU(3)$. Often, by a slight abuse of terminology, the space on which the transformations act is also referred to as the representation. Thus, we may also say that the space spanned by column vectors of the form (5.18) is the defining representation (or fundamental representation) of $SU(3)$.

We now consider a tensor version of the transformation, $T'_{ij} = U_{ik}U_{jl}T_{kl}$. This is definitely a representation, as may be seen by checking the group composition law. If we designate the direct product matrix $U_{ik}U_{jl}$ as $\mathbf{U}_{ij,kl}$ (with composite indices as shown), then it is to check that the composition of two group elements U_1 and U_2 to form an element U_3 is reproduced by the $\mathbf{U}$'s as well,

$$U_1U_2 = U_3 \quad \Longrightarrow \quad \mathbf{U}_1\,\mathbf{U}_2 = \mathbf{U}_3\,. \tag{5.20}$$

The product representation $\mathbf{U}$ is in general reducible. Equivalently, the vector space on which it acts splits into invariant subspaces for the action of the group. This can be seen by symmetrizing the tensor; the symmetric and antisymmetric components of the tensor transform separately. Transformations of

the symmetric part are given by

$$\left(\frac{T_{ij}+T_{ji}}{2}\right)' = \frac{1}{2}(U_{ij}U_{jl}T_{kl} + \underbrace{U_{jk}U_{il}T_{kl}}_{U_{jl}U_{ik}T_{lk}}) = U_{ij}U_{jl}\left(\frac{T_{kl}+T_{lk}}{2}\right). \tag{5.21}$$

This shows that the set of symmetric components transforms into itself; it forms an invariant subspace. This is one irreducible component, no further reduction is possible. Similarly, we find that the antisymmetric components transform into themselves and form another invariant subspace,

$$\left(\frac{T_{ij}-T_{ji}}{2}\right)' = U_{ij}U_{jl}\left(\frac{T_{kl}-T_{lk}}{2}\right). \tag{5.22}$$

This is also irreducible. The full 9 elements of T_{ij} are thus reduced to two irreducible representations, a six-dimensional one and a three-dimensional one.

By analyzing possible ways of reduction of a product, we will find three simple rules for $SU(3)$. The first rule, based on what we did so far is the following.

Rule 1: *Separate symmetric and antisymmetric indices.*

Now we consider the complex conjugate of ϕ_i, *i.e.*,

$$\phi_i^* = \begin{pmatrix} \phi_1^* \\ \phi_2^* \\ \phi_3^* \end{pmatrix}. \tag{5.23}$$

The conjugate of (5.19) is written as $\phi_i^{*\prime} = U_{ik}^*\phi_k^*$ where the sum over k is contracted. This can also be expressed as $\phi_i^{*\prime} = U_{ik}^*\phi_k^* = \phi_k^\dagger(U^\dagger)_{ki} = (\phi^\dagger U^\dagger)_i$, *i.e.*,

$$\phi^{\dagger\prime} = \phi^\dagger U^\dagger. \tag{5.24}$$

Notice that ϕ and ϕ^* are linearly independent because conjugation is not a linear transformation. Clearly $U^\dagger$ must define another representation; it is called the conjugate of the fundamental representation or the anti-fundamental representation. To avoid confusion and keep track of the conjugate representation, we label the conjugate by upper indices

$$\phi^{i^*} = U^{ik^*}\phi^{k^*}. \tag{5.25}$$

A tensor of the form T_j^i transforms as

$$T_j^{i\prime} = U^{ik^*}U_{jl}T_l^k. \tag{5.26}$$

Applying the Kronecker delta to the above transformation, we find

$$\delta^{i\prime}_j = U^{ik*}U_{jl}\delta^k_l = U^{ik*}U_{jk} = U_{jk}(U^\dagger)^{ki} = (UU^\dagger)^i_j = \delta^i_j \tag{5.27}$$

where the unitarity of U is used. δ^i_j is therefore an invariant tensor which can be used for the contraction of indices to make a lower rank tensor from a higher rank one. This operation is invariant in the sense that the quantity obtained by the action of the Kronecker delta will transform properly under the group transformations. We also have the relation

$$T^{i\prime}_i = U^{ik*}U_{il}T^k_l = (U^\dagger U)^k_l T^k_l = \delta^k_l T^k_l = T^k_k\,. \tag{5.28}$$

Thus T^i_i, or the trace of T, is invariant under U transformations; this is an identity representation. From these facts, we can decompose the tensor T^i_j as

$$T^i_j = (T_{\rm traceless})^i_j + \frac{1}{3}\,\delta^i_j\,T^k_k\,. \tag{5.29}$$

The traceless part and the trace would form two separate invariant subspaces or equivalently two irreducible representations. This leads us to the next rule.

Rule 2: Separate out all contractions of upper indices with lower indices.

We have seen that contractions can be carried out by applying the invariant tensor δ^i_j. There is another invariant tensor, the antisymmetric Levi-Civita tensor ϵ_{ijk}, arising from the condition $\det U = 1$:

$$\epsilon'_{ijk} = U_{ia}U_{jb}U_{kc}\epsilon_{abc} = \epsilon_{ijk}{\rm det}U = \epsilon_{ijk} \tag{5.30}$$

where we use the definition of the determinant as

$$\epsilon_{ijk}U_{ia}U_{jb}U_{kc} = \epsilon_{abc}\,(\det U)\,. \tag{5.31}$$

Thus we can use ϵ_{ijk} or ϵ^{ijk} for the contraction of tensor indices. For example, suppose that the indices of T_{ij} are antisymmetric. Then we can re-express these components $T^k = T_{ij}\epsilon^{ijk}$. Since ϵ^{ijk} is an invariant tensor, this shows that T_{ij}, equivalently T^k, transform as the antifundmental representation of $SU(3)$. Similarly, we can make contractions such as $T^{ij}\epsilon_{ijk} = T_k$. By successive use of the ϵ-tensors, we can make antisymmetric indices disappear for tensors of any ranks. This gives us rule 3.

Rule 3: Antisymmetric pairs of indices should be contracted with the invariant tensor ϵ_{ijk} or ϵ^{ijk}.

Based on the rules we have enunciated, it is easy to see what kind of irreducible tensor representations are possible for $SU(3)$. Consider applying these rules to a general tensor. A general tensor is of the form

$$T^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_q} \tag{5.32}$$

where each index can take values $1, 2, 3$. From the third rule, we can say that the tensor can be taken to be symmetric under permutations of all upper indices $(i_1 i_2 \cdots i_p)$ and symmetric under permutations of all lower indices $(j_1 j_2 \cdots j_q)$. The trace of this tensor is a tensor of lower rank transforming into itself under $SU(3)$. Thus in isolating irreducible components, we may set it to zero, taking it as having been already counted among the lower rank tensors. This is an application of our second rule. Thus we can say

$$\delta^{i_n}_{j_m} T^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_q} = 0\,. \tag{5.33}$$

We need to consider only the traceless part of a tensor like (5.32) in getting irreducible representations. The combination of the two conditions (5.32) and (5.33) gives all finite-dimensional irreducible representations of the $SU(3)$ algebra.

Let us now count the dimension of these tensor representations. First consider the set of indices $(i_1 i_2 \cdots i_p)$, each taking values $1, 2, 3$. Since this is symmetric, we may think of this as having k indices having the value 3, with $p-k$ indices taking values either 1 or 2. So we could have $p-k$ 1's, zero 2's, $p-k-1$ 1's, one 2, etc. Obviously we have $p-k+1$ possibilities. Allowing k to range from zero to p, we get $\sum_k (p-k+1) = \frac{1}{2}(p+1)(p+2)$ Thus, to begin with, the number of independent components in (5.32) is $\frac{1}{2}(p+1)(p+2) \times \frac{1}{2}(q+1)(q+2)$. Removing the trace (which corresponds to $p-1$ and $q-1$ indices), this number reduces to

$$\frac{(p+1)(p+2)}{2}\frac{(q+1)(q+2)}{2} - \frac{p(p+1)}{2}\frac{q(q+1)}{2} = \frac{(p+1)(q+1)(p+q+2)}{2}\,. \tag{5.34}$$

This number gives the dimension of $SU(3)$ in the irreducible representation which may be labeled by (p, q). Examples of these are given by the following table.

The representation $(p, q) = (1, 1)$ is called the adjoint representation. This corresponds to tensors of the form T^i_j. In terms of these dimensions the reduction

Table 5.1: Some low dimensional irreducible representations of $SU(3)$

(p, q)	Dimension	(p, q)	Dimension
(1, 0)	**3**	(3, 0)	**10**
(0, 1)	$\mathbf{3}^*$	(0, 3)	$\mathbf{10}^*$
(2, 0)	**6**	(2, 1)	**15**
(0, 2)	$\mathbf{6}^*$	(1, 2)	$\mathbf{15}^*$
(1, 1)	**8**	(2, 2)	**27**

(5.29) can be expressed as

$$\mathbf{3} \otimes \mathbf{3}^* = \mathbf{1} \oplus \mathbf{8} \tag{5.35}$$

where **1** denotes the identity representation.

Tensor analysis for $SU(2)$: a digression

The above analysis is also applicable to the $SU(2)$ algebra in which case we have a rank-2 invariant tensor ϵ_{ij}. Contraction of tensor indices with ϵ_{ij} or ϵ^{ij} leads to no distinction between upper and lower indices, *e.g.*, $T_i\epsilon^{ij} = T^j$. Thus it is sufficient to consider a tensor with lower indices only $T_{i_1 i_2 \cdots i_q}$ where $(i_1, i_2, \cdots i_q)$ are symmetric, with i's taking the values $1, 2$. The irreducible representation is then labeled by a single integer q; the value of spin j is given by $q/2$. The dimension of this representation corresponds to the number of independent components of $T_{i_1 i_2 \cdots i_q}$, which is obviously $q + 1 = 2j + 1$.

5.4 Mesons and baryons and mass formulae

We can now consider the application of the results of the previous section to the spectroscopy of hadrons. This is a fairly straightforward case, the flavor $SU(3)$ symmetry is realized in the Wigner mode with explicit symmetry breaking terms due to the mass differences among the quarks. First consider mesons which are bound states of a light quark Q and a light antiquark $\overline{Q}$. The quark content of the mesons can thus be given as $M^i_j = Q^i\overline{Q}_j$. Since quarks transform as the fundamental representation **3** of $SU(3)$, the mesons belong to the product $\mathbf{3}\otimes\mathbf{3}^*$. Using the decomposition (5.35), we can then split these mesons into singlets (trace part) and octets (traceless part). Thus we should expect 8 mesons of approximately equal mass and one meson which may have a significantly different mass. Further since quarks are spin-$\frac{1}{2}$ particles, we should expect at least a set of spin-0 and spin-1 mesons, with the bound quarks being in a state of zero orbital angular momentum. There could be higher ones with nonzero orbital angular momentum for the quarks as well. Further, parity is preserved by strong interactions, so we can expect parity even and parity odd states for each spin and $SU(3)$ representation. For typical spin-0 and parity-odd mesons (the so-called pseudoscalar mesons), their masses and quark contents are listed in Table 5.2. The upper scripts stand for the electric charges of the mesons, being consistent with the charges of light quarks $(+2/3, -1/3, -1/3)$ for (u, d, s). The table also lists the lowest lying spin-1 and parity-odd mesons (the so-called vector mesons).

Baryons are bound states of three quarks (QQQ). In terms of the tensor representation, this can be expressed as T_{ijk}. This is reducible since we can make contractions such as $T_{ijk}\epsilon^{jkl} = T^l_i$. Baryons are then classified by the decuplets B_{ijk} and the octets B^i_j. In fact, in terms of the dimension of $SU(3)$,

Table 5.2: Spin-0 and spin-1 mesons with the lowest masses

	Spin-0	Mass (MeV)	Spin-1	Mass (MeV)	Quark content
Singlet	η'	958	ω	783	$(u\bar{u}+d\bar{d}+s\bar{s})/\sqrt{3}$
Octet	π^0	135	ρ^0	775	$(u\bar{u}-d\bar{d})/\sqrt{2}$
	π^+	140	ρ^+	775	$u\bar{d}$
	π^-	140	ρ^-	775	$d\bar{u}$
	K^+	494	K^{*+}	892	$u\bar{s}$
	K^-	494	K^{*-}	892	$s\bar{u}$
	K^0	498	K^{*0}	896	$d\bar{s}$
	$\bar{K}^0$	498	$\bar{K}^{*0}$	896	$s\bar{d}$
	η	548	φ	1019	$(u\bar{u}+d\bar{d}-2\,s\bar{s})/\sqrt{6}$

this can be expressed as

$$\begin{aligned} \mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} &= (\mathbf{6}\oplus\mathbf{3}^*)\otimes\mathbf{3} \\ &= \mathbf{10}\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{1}\,. \end{aligned} \tag{5.36}$$

Since quarks have spin $\frac{1}{2}$, baryons have spin $\frac{1}{2}$ and $\frac{3}{2}$. First consider the decuplet of baryons. These correspond to the $(3,0)$ representation and is totally symmetric in the flavor indices. They are totally antisymmetric in color indices because we must always have a color singlet and the only way to make a singlet out of three quarks is to take the totally antisymmetric color combination. Thus, Pauli exclusion principle tells us that these must be symmetric in spin; they are spin-$\frac{3}{2}$ particles. The octets, by a similar reasoning, can be seen to be spin-$\frac{1}{2}$ particles. An example of the octets is given by a set of baryons $\{\Sigma^{\pm}, \Sigma^0, p, n, \Lambda, \Xi^-, \Xi^0\}$. Here p is the familiar proton and n is the neutron, while the others are baryons containing the strange quark as well. In a matrix representation, the octet of baryons can be expressed as

$$B = \begin{pmatrix} \frac{\Sigma^0}{\sqrt{2}}+\frac{\Lambda}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & -\frac{\Sigma^0}{\sqrt{2}}+\frac{\Lambda}{\sqrt{6}} & n \\ \Xi^- & \Xi^0 & -\sqrt{\frac{2}{3}}\Lambda \end{pmatrix} = \sum_{a=1}^{8}\psi^a\frac{\lambda^a}{\sqrt{2}} \tag{5.37}$$

where λ^a $(a=1,2,\cdots,8)$ are the Gell-Mann matrices explicitly shown in (1.49). ψ^a denote the field (or operator) corresponding to the particles, $\psi^3=\Sigma^0$, $(\psi^4-i\psi^5)/\sqrt{2}=p$, etc. The numerical coefficients in (5.37) are chosen from normalization of $\mathrm{Tr}(\overline{B}B)$,

$$\mathrm{Tr}(\overline{B}B) = \bar{p}p+\bar{n}n+\overline{\Sigma^-}\Sigma^-+\overline{\Sigma^0}\Sigma^0+\overline{\Sigma^-}\Sigma^-+\overline{\Xi^-}\Xi^-+\overline{\Xi^0}\Xi^0+\overline{\Lambda}\Lambda \tag{5.38}$$

where

$$\overline{B} = \begin{pmatrix} \frac{\overline{\Sigma^0}}{\sqrt{2}} + \frac{\overline{\Lambda}}{\sqrt{6}} & \overline{\Sigma^-} & \overline{\Xi^-} \\ \overline{\Sigma^+} & -\frac{\overline{\Sigma^0}}{\sqrt{2}} + \frac{\overline{\Lambda}}{\sqrt{6}} & \overline{\Xi^0} \\ \bar{p} & \bar{n} & -\sqrt{\frac{2}{3}}\overline{\Lambda} \end{pmatrix} . \tag{5.39}$$

The advantage of the matrix way of writing these out is that, under the flavor $SU(3)$ transformation, $B \to U\,B\,U^\dagger$ where U is the 3×3 matrix which is an element of $SU(3)$. Notice that we can also tell quark contents of each baryon from the matrix (5.37). For example, p, n and Σ^+ are expressed as

$$\begin{aligned} p &\to {B_1}^3 \sim \epsilon^{3ij} B_{1ij} \sim B_{112} \sim uud\,, \\ n &\to {B_2}^3 \sim B_{212} \sim udd\,, \\ \Sigma^+ &\to {B_1}^2 \sim B_{131} \sim uus\,. \end{aligned} \tag{5.40}$$

These baryons have spin-$\frac{1}{2}$ and parity-even. Their masses and quark contents are listed in Table 5.3. A similar list for the decuplet of baryons with their masses is given in Table 5.4.

Baryon masses

If the flavor $SU(3)$ symmetry were exact, all members of the same multiplet of baryons would have the same mass. But we know that the flavor symmetry is broken by the mass differences among the quarks, as well as by weak interaction effects. Thus the mass differences, say in the octet of Table 5.3, presumably can be understood in terms of such effects. It is actually possible to do better than this qualitative statement, by writing down a possible mass term which builds in the pattern of symmetry breaking. We can express the mass terms, say for the octet baryons, as

$$M \;=\; a\,\mathrm{Tr}(\overline{B}B) \;+\; (\text{terms not preserving the } SU(3) \text{ symmetry}) \tag{5.41}$$

where the coefficient a denotes a common mass in the octet because $\mathrm{Tr}(\overline{B}B)$ is invariant under the $SU(3)$ symmetry; since $B' = UBU^\dagger$ and $\overline{B}' = U\overline{B}U^\dagger$,

Table 5.3: The octet of baryons and their masses

	Baryon	Mass (MeV)	Quark content
Octet	p	938	uud
	n	940	udd
	Λ	1116	uds
	Σ^+	1189	uus
	Σ^0	1193	uds
	Σ^-	1197	dds
	Ξ^0	1315	uss
	Ξ^-	1322	dss

Table 5.4: The decuplet of baryons and their masses

	Baryon	Mass (MeV)	Quark content
Decuplet	Δ^{++}	1232	uuu
	Δ^{+}	1232	uud
	Δ^{0}	1232	udd
	Δ^{-}	1232	ddd
	Σ^{*+}	1383	uus
	Σ^{*0}	1384	uds
	Σ^{*-}	1387	dds
	Ξ^{*0}	1532	uss
	Ξ^{*-}	1535	dss
	Ω^{-}	1672	sss

we can easily check that $\mathrm{Tr}(\overline{B}'B') = \mathrm{Tr}(\overline{B}B)$. The first term in (5.41) thus represents a contribution from the binding energy among quarks. The remaining terms correspond to those that break the $SU(3)$ symmetry. As discussed in the beginning of this chapter, the quark mass term $\overline{Q}mQ$ breaks this symmetry due to mass differences, $m_u \approx m_d \ll m_s$. If we ignore the m_u-m_d mass difference, we can approximate the mass matrix in the action (5.1) as

$$m \approx \begin{pmatrix} m_u & & \\ & m_u & \\ & & m_s \end{pmatrix} = m_u \mathbf{1} + \begin{pmatrix} 0 & & \\ & 0 & \\ & & m_s - m_u \end{pmatrix}. \tag{5.42}$$

Thus, if we take the dominant source of symmetry breaking to be the mass differences among the quarks, (5.42) suggests that the remaining terms in (5.41) may be approximated by the $(3,3)$-element of 3×3 matrix $\overline{B}B$ or $B\overline{B}$. Thus we can write down a general mass formula for the octet baryons as

$$M = a\,\mathrm{Tr}(\overline{B}B) + b\,(\overline{B}B)_{33} + c\,(B\overline{B})_{33} \tag{5.43}$$

$$(\overline{B}B)_{33} = \bar{p}p + \bar{n}n + \frac{2}{3}\overline{\Lambda}\Lambda \tag{5.44}$$

$$(B\overline{B})_{33} = \overline{\Xi^{-}}\Xi^{-} + \overline{\Xi^{0}}\Xi^{0} + \frac{2}{3}\overline{\Lambda}\Lambda \tag{5.45}$$

where a, b and c are parameters to be fixed. (They are calculable, in principle, from the underlying quark dynamics, but that is very difficult. We will treat them as parameters to be fixed by comparison with some of the known masses.) There are 8 masses and 3 unknown parameters so that we can actually make predictions for masses from this formula. Using (5.38) and (5.43)-(5.45), we find the relations:

$$\begin{cases} M_p = M_n = a + b\,, \\ M_{\Sigma^0} = M_{\Sigma^+} = M_{\Sigma^-} = a\,, \\ M_{\Xi^0} = M_{\Xi^-} = a + c\,, \\ M_\Lambda = a + \frac{2}{3}(b+c)\,. \end{cases} \tag{5.46}$$

From these, we can obtain the following mass relations:

$$\begin{aligned} &\left\{ \begin{array}{rcl} M_p + M_{\Xi^0} &=& 2a + b + c\,, \\ \frac{1}{2}M_{\Sigma^0} + \frac{3}{2}M_\Lambda &=& 2a + b + c\,, \end{array} \right. \\ \Longrightarrow\quad & 2(M_p + M_{\Xi^0}) = M_{\Sigma^0} + 3M_\Lambda\,. \end{aligned} \tag{5.47}$$

Suppose we do not know one of the baryon masses in (5.47), then this can be used to predict unknown masses. We can check that this relation holds to a fairly good accuracy with the measured masses listed in Table 5.3. The data is also in accord with the relations $M_p = M_n$, $M_{\Sigma^0} = M_{\Sigma^+} = M_{\Sigma^-}$ and $M_{\Xi^0} = M_{\Xi^-}$, which are predicted from (5.46) as well.

For the decuplets, we can similarly obtain a mass formula by assuming $m_u = m_d$. An explicit formula is expressed as

$$M \;=\; a \,+\, b \times (\text{number of } \overline{Q}_3\text{'s}) \,+\, c \times (\text{number of } Q_3\text{'s}) \tag{5.48}$$

where Q_3 corresponds to the strange quark. Applying this to the decuplet baryons in Table 5.4, we find the relations:

$$M_{\Omega^-} - M_{\Xi^*} \;=\; M_{\Xi^*} - M_{\Sigma^*} = M_{\Sigma^*} - M_\Delta\,, \tag{5.49}$$

$$M_\Delta \;=\; M_{\Delta^{++}} = M_{\Delta^+} = M_{\Delta^0} = M_{\Delta^-}\,, \tag{5.50}$$

$$M_{\Sigma^*} \;=\; M_{\Sigma^{*+}} = M_{\Sigma^{*0}} = M_{\Sigma^{*-}}\,, \tag{5.51}$$

$$M_{\Xi^*} \;=\; M_{\Xi^{*0}} = M_{\Xi^{*-}}\,. \tag{5.52}$$

These are in good agreement with the real masses. Historically, Gell-Mann used the formula (5.49) to predict the mass of Ω^- (as around 1675 MeV) which had not been discovered by that time. The discovery of the Ω^- in 1964 with a mass of 1672 MeV was thus a real triumph of the symmetry-based analysis. The mass formulae such as (5.41) and (5.48) have provided useful criteria for predictions and explanations of the hadron spectroscopy. These were developed by Okubo and Gell-Mann in the early 1960's.

Meson masses

Since the mesons we have considered so far also belong to the octet representation, we can expect a similar pattern of masses for them. If we write an effective action for the mesons represented by appropriate fields, it would involve the square of the mass rather than the mass itself. Hence we expect formulae similar to (5.47) to hold for the $(\text{mass})^2$ of the mesons, *i.e.*,

$$\begin{aligned} 2\,(M_K^2 + M_{K^0}^2) &= M_{\pi^0}^2 + 3\,M_\eta^2\,, \\ 2\,(M_{K^*}^2 + M_{K^{0*}}^2) &= M_{\rho^0}^2 + 3\,M_\varphi^2\,. \end{aligned} \tag{5.53}$$

One can check that these too hold to some reasonable accuracy using the experimentally measured masses.

Symmetry can be used for a number of other predictions as well. For example, it has been applied to the magnetic moments, yielding relations known as the Coleman-Glashow relations.

Meson-Baryon interactions

Symmetry can also give some constraints on interactions. Consider, for example, the baryon-baryon-meson interactions for the octets of baryons and pseudoscalar mesons. If we think of writing down the most general term of the B-B-M type in an action, we could have different couplings for each choice of particles giving 8^3 couplings.

Of course, there will be some reduction based on charge conservation and other principles. Nevertheless, a lot of unknown couplings would be there. (Once again, they are all calculable from the fundamental theory, but this has been very difficult; so we are talking about what we can say in the absence of such calculations.) Based on the flavor $SU(3)$ symmetry, there can only be two basic couplings given by

$$S_{int} = \int d^4x \ \left[g_1 \mathrm{Tr}(\bar{\mathbf{B}}\,\gamma_5\,\mathbf{B}\,\mathbf{M}) + g_2 \mathrm{Tr}(\bar{\mathbf{B}}\,\gamma_5\,\mathbf{M}\,\mathbf{B}) \right] . \tag{5.54}$$

Here γ_5 is the Dirac γ_5 matrix. It is needed because the mesons are pseudoscalar and strong interactions conserve parity. And $\mathbf{M}$ stands for the matrix way of writing out the mesons, similar to (5.37) for the baryons,

$$\mathbf{M} = \begin{bmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2}{\sqrt{6}}\eta \end{bmatrix} . \tag{5.55}$$

Working out the scattering or decay amplitudes with the action (5.54), we can relate the decay rates and scattering amplitudes for a host of processes, since there are only two unknowns. (The technology is similar to the use of the Wigner-Eckart theorem for relating transition matrix elements, as for example, the absorption and emission of photons by different angular momentum states of the atom.)

It is interesting to write out the action (5.54) for the pion-nucleon sector. Retaining only the pions and the p and the n,

$$S_{int} = g_2 \int d^4x \ \left[\frac{\pi^0}{\sqrt{2}} \big(\bar{p}\gamma_5\, p - \bar{n}\,\gamma_5 n \big) + \bar{n}\,\gamma_5 p\,\pi^- + \bar{p}\,\gamma_5 n\,\pi^+ \right] . \tag{5.56}$$

This is essentially the original interaction proposed by Yukawa in his prediction of the π-mesons.

Chapter 6

Superconductivity and the BCS theory

6.1 Introduction to superconductivity

Superconductivity is a phenomenon which is observed in many materials at very low temperature. This has been an extensive area of research in both experimental and theoretical physics ever since its first observation in 1911. The basic framework for a theory of superconductivity was given by Bardeen, Cooper and Schrieffer (BCS) in 1957. In this chapter, we will consider some algebraic aspects of the BCS theory. As an introduction we shall first present some basic facts of superconductivity for our purposes.

1. As one cools the material, the electric resistance abruptly drops to zero at some temperature T_c and remains zero for all lower temperatures.

2. Superconductivity is observed in a variety of materials. This implies that details of the band structure of each material are not very important.

3. There is a gap in the spectrum for excited electrons.

4. The transition temperature T_c is related to the mass of nuclei $M_{\rm nuc}$ by $T_c \propto {M_{\rm nuc}}^{-\frac{1}{2}}$, showing that superconductivity involves the atomic nuclei, not just the electrons, in an essential way. As discussed below, this also suggests that lattice vibrations, *i.e.*, phonons or quantized sound waves, are crucial.

In a system of many electrons inside a lattice of nuclei, there are two types of electron-electron interaction which are important. One is, of course, the Coulomb repulsion and the other is an interaction mediated by lattice vibrations

or phonons. It is known that the Coulomb interactions are screened in the situation we are considering, namely, a multi-electron system with a background of atomic nuclei. The screening is described by a Yukawa potential $V_{\rm Yuk} \sim (e^{-ar})/r$ where a^{-1} represents an effective range of distance for the Coulomb interactions. In the momentum-space representation, the Yukawa potential is proportional to $(k^2+a^2)^{-1}$ where $\vec{k}$ ($k=\sqrt{|\vec{k}|^2}$) denotes the momentum vector in three spatial dimensions. One can easily check this relation from the Fourier transform

$$\int \frac{d^3k}{(2\pi)^3}\frac{1}{k^2+a^2}e^{i\vec{k}\cdot\vec{r}} = \frac{e^{-ar}}{4\pi r} \sim V_{\rm Yuk}(r)\,. \tag{6.1}$$

Thus, at some energy level where $V_{\rm Yuk}$ is comparatively small, electron-phonon interactions can become dominant as an attractive force and this attractive force creates pairs of electrons. A pair of electrons behaves as a boson so that we can interpret superconductivity as a consequence of the Bose-Einstein condensation. This is a main idea of the BCS theory. The pair of electrons is called the Cooper pair. Let $k_{\rm vib}$ be an analog of "spring constant" for the vibrations of nuclei. Then, roughly speaking, a typical energy level for the electron-phonon interaction is given by $\omega_{\rm vib} \sim \sqrt{k_{\rm vib}/M_{\rm nuc}}$. Thus the picture that electrons are paired via interactions with phonons has the potential to explain the above-mentioned property of $T_c \propto {M_{\rm nuc}}^{-\frac{1}{2}}$.

Fermion operators

We shall start by reviewing the formalism of creation and annihilation operators of an electron in a multi-electron system. One electron state with momentum $\vec{k}$ can be expressed as $|\cdots 1_k \cdots\rangle \equiv |1_k\rangle$ where we ignore spins. Let C_k be an annihilation operator of this electron. This means that we can write

$$C_k|1_k\rangle = |0_k\rangle\,. \tag{6.2}$$

This is equivalent to stating that $\langle 0_k|C_k|1_k\rangle = 1$. We can rewrite this relation as $\langle\psi|1_k\rangle = 1$ where $|\psi\rangle = C_k^\dagger|0_k\rangle$, which is basically the definition of the adjoint of C_k. This shows that

$$C_k^\dagger|0_k\rangle = |1_k\rangle\,. \tag{6.3}$$

Thus $C_k^\dagger$ denotes a creation operator for an electron of momentum $\vec{k}$. Since electrons are fermions, we have the Pauli exclusion principle which does not allow double occupancy of a state. We must therefore have

$$C_k^\dagger C_k^\dagger|\cdots 0_k \cdots\rangle = 0\,. \tag{6.4}$$

(Notice that here we ignore spin degrees of freedom.) Since a similar result must hold when we start with any state, generally, the exclusion principle is expressed as

$$C_k^\dagger C_k^\dagger = C_k C_k = 0\,. \tag{6.5}$$

From above relations, we also find

$$\begin{array}{l} C_k C_k^\dagger |0_k\rangle = C_k |1_k\rangle = |0_k\rangle \,, \\ C_k^\dagger C_k |0_k\rangle = C_k^\dagger C_k C_k |1_k\rangle = 0 \,, \\ C_k C_k^\dagger |1_k\rangle = C_k C_k^\dagger C_k^\dagger |0_k\rangle = 0 \,, \\ C_k^\dagger C_k |1_k\rangle = C_k^\dagger |0_k\rangle = |1_k\rangle \,. \end{array} \tag{6.6}$$

It is now easily verified that C_k and $C_k^\dagger$ obey the anticommutation rule

$$C_k C_k^\dagger + C_k^\dagger C_k = 1 \,. \tag{6.7}$$

These relations can easily be expanded to multiple electrons, allowing for states of different momenta, say, $\vec{k}$ and $\vec{l}$. Since a two-electron state is antisymmetric under exchange of the particles, we have $C_k^\dagger C_l^\dagger |0_k 0_l\rangle = |1_k 1_l\rangle = -|1_l 1_k\rangle$. This leads to the relation $C_k^\dagger C_l^\dagger = -C_l^\dagger C_k^\dagger$. From this and its conjugate, we find

$$C_k^\dagger C_l^\dagger + C_l^\dagger C_k^\dagger = 0 \,, \tag{6.8}$$

$$C_k C_l + C_l C_k = 0 \,. \tag{6.9}$$

The anticommutation relation (6.7) can also be generalized as

$$C_k C_l^\dagger + C_l^\dagger C_k = \delta_{kl} \,. \tag{6.10}$$

The relations (6.8)-(6.10) give the algebra for creation and annihilation of fermions. From (6.6) we find that the number operator is given by $C_k^\dagger C_k$.

More generally, we can use labels α, β, etc., to denote a general one-particle state. They are composite indices standing in for, say, momentum and spin labels; $|\alpha\rangle = |\vec{k}, \uparrow\rangle$, for example. Denoting the annihilation operator for fermions in such a state by C_α, the full fermionic algebra is given by

$$\begin{array}{rcl} C_\alpha \, C_\beta + C_\beta \, C_\alpha & = & 0 \,, \\ C_\alpha^\dagger \, C_\beta^\dagger + C_\beta^\dagger \, C_\alpha^\dagger & = & 0 \,, \\ C_\alpha \, C_\beta^\dagger + C_\beta^\dagger \, C_\alpha & = & \delta_{\alpha\beta} \,. \end{array} \tag{6.11}$$

Electron-phonon interaction

Let us now consider the electron-phonon interaction which is crucial for superconductivity, at least for the standard BCS-type superconductors. This interaction arises from the electrostatic energy of the electrons in the field of the oscillating ions in a crystal lattice. We begin with the approximate expression for this when the amplitude of the lattice vibrations is small compared to the lattice spacing.

The positions of the ions in a crystal can be written as

$$x_{\alpha i} = a_{\alpha i} \; + \; \xi_i(x^0, a_\alpha) \,. \tag{6.12}$$

Here $a_{\alpha i} = \vec{a}_\alpha$ are the lattice sites, $i = 1, 2, 3$, and $\alpha = 1, 2, \cdots, N$. The total number of ions is N. (Also we use x^0, y^0, etc. for time.) The lattice vibrations ξ_i are given by

$$\xi_i(x^0, a_\alpha) = \sum_\alpha \left(\frac{b_{ik}}{\sqrt{2\omega_k}} e^{i\vec{k}\cdot\vec{a}_\alpha - i\omega_k x^0} + \frac{b_{ik}^\dagger}{\sqrt{2\omega_k}} e^{-i\vec{k}\cdot\vec{a}_\alpha + i\omega_k x^0} \right) . \tag{6.13}$$

In this expression, b_{ik} and $b_{ik}^\dagger$ denote the annihilation and creation operators for phonons, $\omega_k = \omega(k)$ gives the energy-momentum relation for phonons. The charge density due to ions is

$$\rho(x) = Ze \sum_\alpha \delta^{(3)}(x - a_\alpha - \xi(a_\alpha)) . \tag{6.14}$$

Note that as in Chapter 4 on the fractional QHE, we omit arrows of vectors when it is obvious in the following. The electrostatic potential due to the ions is given by

$$\begin{aligned} A_0(x) &= \int d^3y \int \frac{d^3q}{(2\pi)^3} e^{i\vec{q}\cdot(\vec{x}-\vec{y})} \frac{1}{q^2}\ \rho(\vec{y}) \\ &= Ze \sum_\alpha \int \frac{d^3q}{(2\pi)^3} e^{i\vec{q}\cdot(\vec{x}-\vec{a}_\alpha)} \frac{e^{-i\vec{q}\cdot\xi}}{q^2} \\ &\approx Ze \sum_\alpha G(\vec{x} - \vec{a}_\alpha) - iZe \sum_\alpha \frac{\vec{q}\cdot\xi(a_\alpha)}{q^2} e^{i\vec{q}\cdot(\vec{x}-\vec{a}_\alpha)} + \cdots \end{aligned} \tag{6.15}$$

where $G(\vec{x} - \vec{y})$ is the Coulomb Green's function given by

$$G(\vec{x} - \vec{y}) = \int \frac{d^3q}{(2\pi)^3} \frac{1}{q^2} e^{i\vec{q}\cdot(\vec{x}-\vec{y})} . \tag{6.16}$$

The electrostatic interaction energy is given by $e\int \psi^*\psi A_0$. Here ψ denotes a non-relativistic electron field. When we use (6.15), the first term is a constant independent of the lattice vibrations, it is not important for the electron-phonon interaction. The second term gives

$$\begin{aligned} S_{int} &= -iZe^2 \int d^4x\ \psi^*\psi(x) \sum_\alpha \frac{\vec{q}\cdot\xi(a_\alpha)}{q^2} e^{i\vec{q}\cdot(\vec{x}-\vec{a}_\alpha)} + \cdots \\ &\approx -Ze^2 \int d^4x\ \psi^*\psi(x) \frac{\delta}{\delta x_i} \left(\sum_\alpha G(\vec{x} - \vec{a}_\alpha)\xi_i(a_\alpha) \right) . \end{aligned} \tag{6.17}$$

We now consider the long wavelength or continuum limit, with a_α becoming continuous, $\sum_\alpha \to (constant)\int d^3y$, to write

$$S_{int} = -Ze^2(constant) \int d^4x d^3y\ \psi^*\psi(x)\nabla_i G(\vec{x} - \vec{y})\xi_i(y) . \tag{6.18}$$

We also introduce the continuum version of the phonon field by

$$\begin{aligned} \phi_i(x) &= \sum_k \left(b_{ik} u_k(x) + b_{ik}^\dagger u_k^*(x) \right) , \\ u_k(x) &= \frac{1}{\sqrt{2\omega_k V}} \exp(-i\omega_k x^0 + i\vec{k}\cdot\vec{x}) . \end{aligned} \tag{6.19}$$

Absorbing all factors due to the changes in normalizations as a single factor F, we can write the interaction term as

$$S_{int} = F \int dx^0 d^3x d^3y \ \psi^*\psi(x) G(\vec{x}-\vec{y}) \ \nabla\cdot\phi . \tag{6.20}$$

There are many effects we have not treated carefully here; for example, the electrons may not feel the full electrostatic potential of the nuclei due to screening effects, to mention one. But the point is that the form of the interaction is as given in (6.20). The strategy is to treat F as summarizing all the unknowns and identify its value, not from detailed calculations of the electron-atom interactions in the solid but from relating it to measurable quantities such as the dielectric constant. We can see how this can be done in an elementary way as follows. We calculate the electron-electron interaction in perturbation theory due to the phonon exchange using the interaction (6.20). This contribution is given, in second order perturbation theory, by

$$\begin{aligned} \Gamma &= -\frac{i}{2!}\langle (iS_{int})^2 \rangle \\ &= \frac{i}{2} F^2 \int_{x,y} \psi^*\psi(x) \ \psi^*\psi(y) \int \frac{d^4k}{(2\pi)^4} \ \frac{ie^{-ik_0(x^0-y^0)}}{k_0^2-\omega_k^2+i\epsilon} \ \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{\vec{k}\cdot\vec{k}} . \end{aligned} \tag{6.21}$$

For nearly static charge densities, $\psi^*\psi(x)$ varies very slowly with x^0, so that we may do one of the time-integrations to get

$$\Gamma = \frac{i}{2} F^2 \int d^4x d^3y \ \frac{d^3k}{(2\pi)^3} \psi^*\psi(x) \ \psi^*\psi(y) e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \frac{i}{(-\omega_k^2)} \ \frac{1}{\vec{k}\cdot\vec{k}} . \tag{6.22}$$

For the polaron problem, optical modes of phonons are the relevant ones. For this, we may approximate $\omega_k \approx \omega_0$. Thus

$$\Gamma \approx \frac{F^2}{2\omega_0^2} \int d^4x d^3y \ \psi^*\psi(x) G(\vec{x}-\vec{y}) \psi^*\psi(y) . \tag{6.23}$$

The standard Coulomb interaction between electrons is given by

$$S_{Coul} = -\frac{e^2}{2} \int d^4x d^3y \ \psi^*\psi(x)\psi^*\psi(y) \int \frac{d^3k}{(2\pi)^3} \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{\epsilon(\omega)k^2} \tag{6.24}$$

where $\epsilon(\omega)$ is the dielectric constant for photon frequency ω. Let ϵ_∞ denote the dielectric constant without the contribution due to electron-phonon interaction.

Since the phonon effects are negligible at very high frequencies, this is essentially $\epsilon(\omega)$ as $\omega \to \infty$, which explains the notation. The dielectric constant at low frequencies, or $\epsilon(\omega)$ as $\omega \to 0$ (and which we denote by ϵ_0), includes the effects of the electron-phonon interaction. Comparing (6.23) and (6.24), we may thus write

$$\frac{1}{\epsilon_0} = \frac{1}{\epsilon_\infty} \; - \; \frac{F^2}{e^2\omega_0^2} \, . \tag{6.25}$$

The electron-phonon interaction can now be expressed in terms of the dielectric constant as

$$S_{int} = e\omega_0 \sqrt{\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}} \int d^4x d^3y \; \psi^*\psi(x) \; G(\vec{x}-\vec{y}) \; \nabla\cdot\phi(y) \, . \tag{6.26}$$

This is the form of the electron-phonon interaction obtained by Fröhlich in 1952.

The pairing interaction

In the interaction (6.26), the fields ψ and ψ^* denote the electron annihilation and creation operators and $\vec{\phi}$ is the phonon field. The propagator for the phonon field gives the two-point function

$$\begin{aligned} \langle \chi(x) \, \chi(y) \rangle = \int \frac{d^3q}{(2\pi)^3} \, \frac{|D_q|^2}{2\omega_q} \Big[& \theta(x^0-y^0) \, e^{-i\omega_q(x^0-y^0)+i\vec{q}\cdot(\vec{x}-\vec{y})} \\ & + \theta(y^0-x^0) \, e^{i\omega_q(x^0-y^0)-i\vec{q}\cdot(\vec{x}-\vec{y})} \Big] \end{aligned} \tag{6.27}$$

where $\chi(x) = \int d^3y \, G(x,y) \nabla\cdot\phi(y)$ and $|D_q|^2 = 1/q^2$ for us. This function could be different from $1/q^2$ when we take account of the details of the crystal structure. So basically, we will leave it as a function to be specified for each material of interest.

By expanding $e^{iS_{int}}$ to the quadratic order and carrying out the Wick contraction, we get the effective action due to the phonon exchange (to this order) as

$$\Gamma = i\frac{F^2}{2} \int d^4x d^4y \, \psi^*\psi(x) \; \psi^*\psi(y) \; \langle \chi(x)\chi(y) \rangle \, . \tag{6.28}$$

Introduce the mode expansions

$$\psi(x) = \frac{1}{\sqrt{V}} \sum_l C_l \, e^{-iE_l x^0 + i\vec{l}\cdot\vec{x}}, \qquad \psi^*(x) = \frac{1}{\sqrt{V}} \sum_k C_k^\dagger \, e^{iE_k x^0 - i\vec{k}\cdot\vec{x}} \tag{6.29}$$

where C_k, $C_k^\dagger$ are the annihilation and creation operators for electrons we mentioned earlier. Using this expansion in (6.28), the term with $x^0 > y^0$ becomes

$$\Gamma^{(1)} = i\frac{F^2}{2} \int_q \frac{|D_q|^2}{2\omega_q} \sum_{l,r} C_k^\dagger \, C_l \, C_p^\dagger \, C_r \; \delta_{\vec{p},\vec{r}-\vec{q}} \, \delta_{\vec{k},\vec{l}+\vec{q}} \, \frac{2\pi \, \delta(E_k + E_p - E_l - E_r)}{i(E_p - E_r + \omega_q - i\epsilon)}$$

$$= \frac{F^2}{2}\int_q \frac{|D_q|^2}{2\omega_q}\sum_{l,r}\int dt\,\frac{C^\dagger_{l+q}(t)\,C_l(t)\,C^\dagger_{r-q}\,(t)C_r(t)}{(E_{r-q}-E_r+\omega_q-i\epsilon)} \tag{6.30}$$

where we have used $C_p(t) = C_p\,e^{-iE_pt}$, etc. The energy conservation δ-function is reproduced by the integration over t. The term with $y^0 > x^0$ has a similar form except for the exchange of momenta; it is given by

$$\Gamma^{(2)} = -\frac{F^2}{2}\int_q \frac{|D_q|^2}{2\omega_q}\sum_{l,r}\int dt\,\frac{C^\dagger_{l+q}(t)\,C_l(t)\,C^\dagger_{r-q}\,(t)C_r(t)}{(E_{r-q}-E_r-\omega_q+i\epsilon)} \tag{6.31}$$

where we have also used the change of variable $\vec{q} \to -\vec{q}$. Combining the two terms,

$$\Gamma = \frac{F^2}{2}\int_q \frac{|D_q|^2}{2\omega_q}\sum_{l,r}\int dt\,C^\dagger_{l+q}(t)\,C_l(t)\,C^\dagger_{r-q}\,(t)C_r(t)\times \left[\frac{1}{(E_{r-q}-E_r+\omega_q-i\epsilon)}-\frac{1}{(E_{r-q}-E_r-\omega_q+i\epsilon)}\right]. \tag{6.32}$$

The real part of this expression, with $i\epsilon \to 0$, can be written as

$$\Gamma = \frac{F^2}{2}\int_q |D_q|^2\int dt\,\sum_{l,r}\frac{C^\dagger_{l+q}(t)\,C_l(t)\,C^\dagger_{r-q}\,(t)C_r(t)}{[\omega_q^2-(E_r-E_{r-q})^2]}\,. \tag{6.33}$$

This is like a potential energy term in the action, so we may identify the correction to the Hamiltonian as

$$\tilde{H} = -\frac{F^2}{2}\int_q |D_q|^2\sum_{l,r}\frac{C^\dagger_{l+q}\,C_l\,C^\dagger_{r-q}\,C_r}{[\omega_q^2-(E_r-E_{r-q})^2]}\,. \tag{6.34}$$

This has to be added to the screened Coulomb repulsion between electrons. The total four-electron interaction term is thus

$$H_{int} = \int_q \sum_{l,r} C^\dagger_{l+q}\,C_l\,C^\dagger_{r-q}\,C_r\left(\frac{e^2}{q^2+a^2}-\frac{F^2|D_q|^2}{2\,[\omega_q^2-(E_r-E_{r-q})^2]}\right) \tag{6.35}$$

where a^{-1} is the screening length for the Coulomb interaction. Depending on the strength of the electron-phonon interaction, the second term in (6.35) can overcome the Coulomb repulsion and produce an attractive interaction between electrons, if $(E_r - E_{r-q})^2 < \omega_q^2$.

This interaction Hamiltonian induces a transition from initial states of momenta $\vec{l}, \vec{r}$ into final states of momenta $\vec{l}+\vec{q}$, $\vec{r}-\vec{q}$. Since most of the states up to the Fermi level μ_F are filled at low temperatures, if the initial states are deep inside the Fermi sea, then the final states are already occupied and the scattering is forbidden by the Pauli principle, or mathematically, by the

anticommuting nature of the C_k, $C_k^\dagger$. Thus the electrons for which the interaction (6.35) applies are very near the Fermi surface or above it. Because of the condition $(E_r - E_{r-q})^2 < \omega_q^2$, we see that attractive interaction is obtained only for a small range of values of energy close the Fermi surface. For electrons with energies close to the Fermi energy, there is an enhancement if the initial particles have approximately opposite momenta due to the availability of more final states. To see how this can arise, consider a simple case where the Fermi surface is spherically symmetric, so that the filled states form a filled sphere in momentum space as shown in Fig. 6.1. Consider the total initial momentum $\vec{l} + \vec{r}$ with the incoming momenta $\vec{l}$ and $\vec{r}$ close to Fermi surface. If $\vec{l'}$, $\vec{r'}$ are the final momenta, the total final momentum $\vec{l'} + \vec{r'}$ must be equal to $\vec{l} + \vec{r}$ by conservation of momentum. Thus the allowed values of $\vec{l'}$ and $\vec{r'}$ must lie on a circle which corresponds to the rotation of $\vec{l}$ and $\vec{r}$ around the total value $\vec{l} + \vec{r}$. So if we pick a value for $\vec{l'}$ then $\vec{r'}$ is fixed. The freedom of final states is thus a circle of values corresponding to the orientations of one of the final momenta, say, $\vec{l'}$. However, there is one exception to this case. If $\vec{l}$ and $\vec{r}$ are of approximately opposite orientation, so that $\vec{l} + \vec{r} \approx 0$, then the final value of $\vec{l'}$ can be anywhere of the whole two-dimensional sphere, giving a full 4π solid angle worth of orientations (with $\vec{r'}$ being opposite to it). The number of available final states is substantially larger. Thus, when we have the attractive interaction with states near the Fermi surface, the most significant contribution

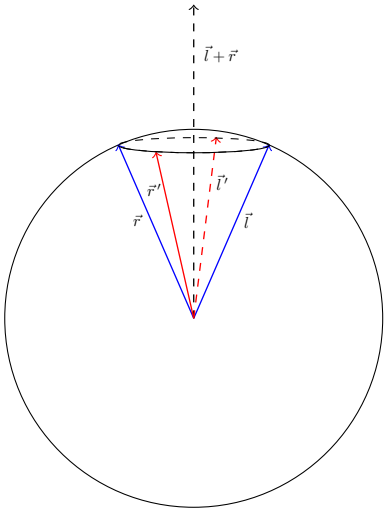

Figure 6.1: Available states for the scattering of electrons near the Fermi surface. The Fermi surface is taken as a two-sphere, with the incoming momenta shown in blue, labeled $\vec{l}$ and $\vec{r}$. The outgoing values lie on a circle generated by rotation of $\vec{l}$, $\vec{r}$ around $\vec{l} + \vec{r}$. When $\vec{l} + \vec{r} = 0$, the outgoing value of $\vec{l'}$ can be anywhere on the surface of the sphere, with $\vec{r'}$ being opposite to it.

arises from the values of $\vec{l}$ and $\vec{r}$ with $\vec{l}+\vec{r}\approx 0$. For these values, we can write

$$E_r - E_{r-q} = \tfrac{1}{2}\left(E_r - E_{r-q} + E_{l+q} - E_l\right) \approx \frac{\vec{q}\cdot(\vec{l}+\vec{r})}{2m_*} \approx 0 \tag{6.36}$$

where we used $E_k = k^2/2m_*$ for a spherically symmetric Fermi surface, m_* being the effective mass of the electrons. Further, since $\vec{l}+\vec{q}=\vec{l}'$ and $\vec{r}-\vec{q}=\vec{r}'$ have opposite orientation, but are otherwise unconstrained in direction, we see that a good approximation to (6.35) is

$$H_{int} \approx \sum_{k,k'} V_{kk'}\, C^\dagger_{-k'} C^\dagger_{k'}\, C_k C_{-k} \tag{6.37}$$

where we have dropped a term due to reordering of the operators (which is irrelevant for our consideration for superconductivity) and

$$V_{kk'} = \left(\frac{e^2}{q^2+a^2} - \frac{F^2|D_q|^2}{2\,\omega_q^2}\right)\Bigg]_{q=k'-k}. \tag{6.38}$$

We are considering the case where $V_{kk'}$ is negative. Notice that, expressed in terms of the electron field operators, we find

$$\sum_k C_k\, C_{-k} = \int d^3x\ \psi(\vec{x})\,\psi(\vec{x}) \tag{6.39}$$

where $\psi(\vec{x})$ is given by

$$\psi(\vec{x}) = \frac{1}{\sqrt{V}}\sum_k C_k\, e^{i\vec{k}\cdot\vec{x}}\,.$$

(6.39) will vanish by anticommutativity of the ψ's. Once the spin degree of freedom is taken into account, however, we can use field operators of different spins to get a nonzero value. Thus, including spin, the Hamiltonian (6.37) must be written as

$$H_{int} = \sum_{k,k'} V_{kk'}\, C^\dagger_{-k'\downarrow} C^\dagger_{k'\uparrow} C_{k\uparrow} C_{-k\downarrow}\,. \tag{6.40}$$

This is the BCS (Bardeen-Cooper-Schrieffer) pairing interaction. It was arrived at after making several approximations and should be viewed as a model capturing the key features of the attractive pairing due to the electron-phonon interaction. Since, generally, the qualitative features survive in more realistic approximations, it can be used as a suitable starting point for the theory of superconductivity.

6.2 The BCS theory

We have obtained the pairing interaction between electrons. Including the free part, the full BCS Hamiltonian is

$$H \;=\; H_0 + H_{int}\,, \tag{6.41}$$

$$H_0 = \sum_{k,\sigma} \epsilon_k C^\dagger_{k\sigma} C_{k\sigma} \qquad \sigma = (\uparrow, \downarrow)\,, \tag{6.42}$$

$$H_{int} = \sum_{k,k'} V_{kk'} C^\dagger_{-k'\downarrow} C^\dagger_{k'\uparrow} C_{k\uparrow} C_{-k\downarrow}\,. \tag{6.43}$$

For simplicity of notation, we shall use different letters for the operators corresponding to different spin states, i.e.,

$$\begin{aligned} b_k &= C_{k\uparrow}\,, \qquad & b^\dagger_k &= C^\dagger_{k\uparrow} \\ c_{-k} &= C_{-k\downarrow}\,, \qquad & c^\dagger_{-k} &= C^\dagger_{-k\downarrow} \end{aligned} \tag{6.44}$$

The Hamiltonian (6.41)-(6.43) is thus

$$\begin{aligned} H &= H_0 \;+\; H_{int} \\ H_0 &= \sum_k \epsilon_k \left[b^\dagger_k b_k \;+\; c^\dagger_k c_k \right] \\ H_{int} &= \sum_{k,k'} V_{kk'}\, b^\dagger_{k'} c^\dagger_{-k'} c_{-k} b_k \end{aligned} \tag{6.45}$$

We can carry out an approximate diagonalization of this Hamiltonian by introducing the combinations

$$\begin{aligned} A_k &= \alpha_k c_{-k} - \beta_k b^\dagger_k \qquad & B_{-k} &= \alpha_k b_k + \beta_k c^\dagger_{-k} \\ A^\dagger_k &= \alpha_k c^\dagger_{-k} - \beta_k b_k \qquad & B^\dagger_{-k} &= \alpha_k b^\dagger_k + \beta_k c_{-k} \end{aligned} \tag{6.46}$$

where α_k, β_k are functions of k which will be determined by the requirement of diagonalization. The basic idea is that we can interpret these operators as the creation and annihilation operators for a set of "quasiparticles" which form the new eigenstates, at least to the approximation we calculate. We require that the new operators obey the same anticommutation rules as the original operators. In other words,

$$\begin{aligned} A_k A_l + A_l A_k &= 0 \qquad & B_{-k} B_{-l} + B_{-l} B_{-k} &= 0 \\ A_k A^\dagger_l + A^\dagger_l A_k &= \delta_{kl} \qquad & B_{-k} B^\dagger_{-l} + B^\dagger_{-l} B_{-k} &= \delta_{kl} \\ A^\dagger_k A^\dagger_l + A^\dagger_l A^\dagger_k &= 0 \qquad & B^\dagger_{-k} B^\dagger_{-l} + B^\dagger_{-l} B^\dagger_{-k} &= 0 \end{aligned} \tag{6.47}$$

It is easily seen that this can be achieved by requiring

$$\alpha_k^2 \;+\; \beta_k^2 \;=\; 1\,. \tag{6.48}$$

This means that we can write

$$\alpha_k = \sin\theta_k\,, \qquad \beta_k = \cos\theta_k\,. \tag{6.49}$$

The transformation of the canonical operators as in (6.46), with the condition (6.48), which ensures the same commutation rules for the new operators as for

the old, is known as a Bogolyubov transformation. The inverse Bogolyubov transformation can be found from the relations (6.46) as

$$\begin{aligned} b_k &= \alpha_k B_{-k} - \beta_k A_k^\dagger & \qquad c_{-k} &= \alpha_k A_k + \beta_k B_{-k}^\dagger \\ b_k^\dagger &= \alpha_k B_{-k}^\dagger - \beta_k A_k & \qquad c_{-k}^\dagger &= \alpha_k A_k^\dagger + \beta_k B_{-k} \end{aligned} \tag{6.50}$$

We now rewrite the Hamiltonian in terms of $A_k, A_k^\dagger, B_{-k}, B_{-k}^\dagger$ using these relations.

$$H_0 = \sum_k \epsilon_k \left[2\beta_k^2 \; + \; (\alpha_k^2 - \beta_k^2)(N_k + \tilde{N}_k) - 2\alpha_k\beta_k(A_k B_{-k} + B_{-k}^\dagger A_k^\dagger) \right] \tag{6.51}$$

where $N_k, \tilde{N}_k$ are the number operators corresponding to the new creation and annihilation operators; they are given by

$$N_k = A_k^\dagger A_k \,, \qquad \tilde{N}_k = B_{-k}^\dagger B_{-k} \,. \tag{6.52}$$

In simplifying H_{int}, at the level of finding the ground state, we will only keep terms which are at most quadratic in the new creation and annihilation operators. Thus terms like $A_k B_{-k} N_k$, which are quartic in the operators, will be neglected. Using that $V_{kk'}$ is symmetric in k, k', we find that H_{int} can be simplified as

$$\begin{aligned} H_{int} \;\; &= \;\; \sum_{kk'} V_{kk'} \left[\alpha_{k'}\beta_{k'} \left(N_{k'} + \tilde{N}_{k'} - 1 \right) + \left(\alpha_{k'}^2 B_{-k'}^\dagger A_{k'}^\dagger - \beta_{k'}^2 A_{k'} B_{-k'} \right) \right] \\ &\qquad\qquad \times \left[\alpha_k\beta_k \left(N_k + \tilde{N}_k - 1 \right) + \left(\alpha_k^2 A_k B_{-k} - \beta_k^2 B_{-k}^\dagger A_k^\dagger \right) \right] \\ &\approx \;\; \sum_{kk'} V_{kk'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k - 2 \sum_{kk'} V_{kk'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k (N_k + \tilde{N}_k) \\ &\qquad - \sum_{kk'} V_{kk'} \alpha_{k'}\beta_{k'} (\alpha_k^2 - \beta_k^2)(A_k B_{-k} + B_{-k}^\dagger A_k^\dagger) \;\; + \;\; (\text{quartic terms}) \,. \end{aligned} \tag{6.53}$$

Our strategy will be to set the coefficient of the $(A_k B_{-k} + B_{-k}^\dagger A_k^\dagger)$-term to zero. The total Hamiltonian then involves only the number operators $N_k, \tilde{N}_k$, to this order, and is trivially diagonalized, since the diagonal basis for $N_k, \tilde{N}_k$ is well known and easily constructed. Setting the coefficient of the $(A_k B_{-k} + B_{-k}^\dagger A_k^\dagger)$-term to zero gives the condition

$$2\,\epsilon_k \alpha_k \beta_k + (\alpha_k^2 - \beta_k^2) \; \sum_{k'} V_{kk'} \alpha_{k'} \beta_{k'} = 0 \,. \tag{6.54}$$

Using (6.49), we can write this as

$$\epsilon_k \sin 2\theta_k + \Delta_k \cos 2\theta_k = 0 \tag{6.55}$$

where we have defined

$$\Delta_k = -\sum_{k'} V_{kk'} \alpha_{k'} \beta_{k'} = -\frac{1}{2} \sum_{k'} V_{kk'} \sin 2\theta_{k'} \,. \tag{6.56}$$

Since we are considering states very close to the Fermi energy, it is a reasonable approximation to take $V_{kk'}$ to be independent of k, k' for the small range of values involved, so that we may write $V_{kk'} = -V_0$. Thus Δ_k defined by (6.56) can be taken to be independent of k; we will henceforth denote it as Δ. Equation (6.55) leads to the expressions

$$\sin 2\theta_k = \frac{\Delta}{\sqrt{\epsilon_k^2 + \Delta^2}}\,, \qquad \cos 2\theta_k = -\frac{\epsilon_k}{\sqrt{\epsilon_k^2 + \Delta^2}}\,. \tag{6.57}$$

Using these expressions back in the definition of Δ_k, namely, in (6.56), we get the following implicit equation for Δ.

$$\begin{aligned} \Delta &= \frac{V_0}{2} \sum_k \frac{\Delta}{\sqrt{\epsilon_k^2 + \Delta^2}} \\ &\approx \frac{V_0}{2} \int d\epsilon\; G(\epsilon)\; \frac{\Delta}{\sqrt{\epsilon^2 + \Delta^2}} \end{aligned} \tag{6.58}$$

where we have approximated the summation over momenta by integration. We have further converted the k-integral to an energy integral with $d^3k \to d\epsilon\, G(\epsilon)$, $G(\epsilon)$ being the density of states. Equation (6.58) is known as the gap equation, for reasons which will become clear shortly. The attractive interaction is valid only for a small range of energy values near the Fermi level. Recalling that we are measuring energy from the Fermi level, this range may be taken as $-\omega_D \le \epsilon \le \omega_D$, where ω_D is the Debye frequency. The integration in (6.58) is limited to this range of ϵ-values. (We may think of Δ as k-dependent, being constant within a small range and then rapidly going to zero beyond this range.) For this range of values, $G(\epsilon)$ is approximately constant, so we can approximate the integral by replacing $G(\epsilon)$ by $G_0 = G(0)$, its value at the Fermi level. The gap equation then simplifies to

$$\Delta\; \left[1 -\; V_0 G_0 \sinh^{-1}(\omega_D/\Delta)\right] = 0\,. \tag{6.59}$$

We have two solutions to this equation, the trivial solution $\Delta = 0$ and the nontrivial one,

$$\begin{aligned} \Delta &= \omega_D\; \frac{1}{\sinh(1/V_0 G_0)} \\ &\approx 2\;\omega_D\; \exp\left(-\frac{1}{V_0 G_0}\right)\,. \end{aligned} \tag{6.60}$$

In the last line, we have used the fact that $V_0 G_0$ is usually very small. Notice that this solution is intrinsically nonperturbative. Since V_0 is the coupling constant or interaction strength, a perturbative expansion is a power series in V_0. Our expression for Δ has an essential singularity at $V_0 = 0$ and so cannot be expanded as a power series around that point. Of the two solutions of the gap equation, we shall see shortly that the nonzero solution is the one which minimizes the energy of the ground state.

6.3 The BCS ground state

With the solution of α_k, β_k given above, the Hamiltonian can be simplified as

$$\begin{aligned} H &= \sum_k 2\epsilon_k \beta_k^2 - V_0 \sum_{kk'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k \\ &\quad + \sum_k \left[-\epsilon_k(\alpha_k^2 - \beta_k^2) + 2V_0 \sum_{k'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k \right] (N_k + \tilde{N}_k) \; + \cdots \end{aligned} \tag{6.61}$$

where the ellipsis denotes terms which are quartic in the new creation and annihilation operators. Further simplification of this expression can be carried out as follows.

$$\begin{aligned} \sum_k 2\epsilon_k \beta_k^2 - V_0 \sum_{kk'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k &= \sum_k \left(2\epsilon_k \cos^2\theta_k - \frac{\Delta}{2}\sin 2\theta_k \right) \\ &= \int d\epsilon \; G(\epsilon) \; \left[\epsilon - \frac{\epsilon^2}{\sqrt{\epsilon^2+\Delta^2}} - \frac{1}{2}\frac{\Delta^2}{\sqrt{\epsilon^2+\Delta^2}} \right] \\ &= -G_0 \left[\frac{1}{2}\epsilon\sqrt{\epsilon^2+\Delta^2} \right]_{-\omega_D}^{\omega_D} \\ &= -G_0\,\omega_D \sqrt{\omega_D^2+\Delta^2} \end{aligned} \tag{6.62}$$

We also find

$$-\epsilon_k(\alpha_k^2 - \beta_k^2) + 2V_0 \sum_{k'} \alpha_{k'}\beta_{k'}\alpha_k\beta_k = \sqrt{\epsilon_k^2+\Delta^2}\,. \tag{6.63}$$

The Hamiltonian (6.61) is thus given by

$$H \;=\; -\,G_0\,\omega_D\sqrt{\omega_D^2+\Delta^2} \;+\; \sum_k \sqrt{\epsilon_k^2+\Delta^2}\;(N_k+\tilde{N}_k) \;+\cdots\,. \tag{6.64}$$

We now see that, to the approximation of neglecting the quartic terms, the ground state $|G\rangle$ may be defined by

$$A_k\;|G\rangle \;=\; B_{-k}\;|G\rangle \;=\; 0\,. \tag{6.65}$$

The ground state energy is then obtained as

$$\mathcal{E}_\Delta \;=\; -G_0\,\omega_D\sqrt{\omega_D^2+\Delta^2}\,. \tag{6.66}$$

This is to be compared with the energy at $\Delta = 0$; the difference is

$$\mathcal{E}_\Delta - \mathcal{E}_{\Delta=0} = -G_0\,\omega_D\left[\sqrt{\omega_D^2+\Delta^2} - \omega_D\right] \approx -\frac{1}{2}G_0\,\Delta^2\,. \tag{6.67}$$

We see that, indeed, the nontrivial solution to the gap equation is preferred on grounds of minimizing the energy.

We also see that the the spectrum of the theory consists of particle-like excitations of energy $\sqrt{\epsilon_k^2 + \Delta^2}$. In the normal phase, with $\Delta = 0$, the energy eigenvalues can be arbitrarily small; *i.e.*, arbitrarily close to the Fermi energy. However, once we have a nonzero Δ, the single particle energies have a gap equal to Δ. (This justifies the name 'gap equation' for (6.58).)

The expressions (6.60) for the gap Δ and the energy gain (6.67) show the "isotope effect". The Debye frequency ω_D depends on the masses of the atomic nuclei as $M_{nuc}^{-\frac{1}{2}}$. This means that superconductors of the same chemical type but with different isotopes for the nuclei will show slightly different values for the energy gap for electrons. A similar statement also applies to the transition temperature.

The determination of the ground state $|G\rangle$ can be done more explicitly as follows. Let $|0\rangle$ be the normal ground state obeying $b_k|0\rangle = c_{-k}|0\rangle = 0$. Since A_k and B_{-k} are fermionic operators, we have $A_k^2 = B_{-k}^2 = 0$. If $|0\rangle$ is the normal ground state obeying $b_k|0\rangle = c_{-k}|0\rangle = 0$, we see that a solution to (6.65) is given by

$$\begin{aligned} |G\rangle &\sim \prod_k f_k B_{-k} A_k |0\rangle \\ &\sim -\prod_k f_k (\alpha_k \beta_k + \beta_k^2 c_{-k}^\dagger b_k^\dagger)|0\rangle \,. \end{aligned} \tag{6.68}$$

The factors f_k can be chosen by the normalization condition $\langle G|G\rangle = 1$. This gives $f_k = 1/\beta_k$ and hence the normalized state is

$$|G\rangle = \prod_k (\alpha_k + \beta_k c_{-k}^\dagger b_k^\dagger)|0\rangle \,. \tag{6.69}$$

It is interesting to consider the inner product of this new ground state with the standard normal ground state $|0\rangle$. We find

$$\langle 0|G\rangle = \prod_k \alpha_k = \prod_k \sin\theta_k \,. \tag{6.70}$$

Since $\sin\theta_k \leq 1$, this overlap is very small and, if the number of modes for the system is infinite, this overlap is actually zero. We can also consider the overlap of any normal state (with a finite number of $b_k^\dagger$s and $c_{-k}^\dagger$s acting on $|0\rangle$) with any new state (with a finite number of $A_k^\dagger$s and $B_{-k}^\dagger$s acting on $|G\rangle$). Reducing the inner product via the commutation rules, we will get a prefactor depending on the specific states times the factor $\langle 0|G\rangle$. This will again be zero in the limit of infinite number of available modes for the system. For example, consider a state of the form $A_k^\dagger|G\rangle$. For the overlap with $|0\rangle$ we find

$$\begin{aligned} \langle 0|A_k^\dagger|G\rangle &= \langle 0|\alpha_k c_{-k} - \beta_k b_k^\dagger|G\rangle \\ &= \alpha_k \prod_l \langle 0|c_{-k}(\alpha_l + \beta_l c_{-l}^\dagger b_l^\dagger)|0\rangle \end{aligned}$$

$$= \left(\prod_{l\neq k}\alpha_l\right)\ \alpha_k\ \langle 0|c_{-k}c^\dagger_{-k}b^\dagger_k|0\rangle = 0\,. \tag{6.71}$$

where we have used $\langle 0|b^\dagger_k = 0$. A somewhat more nontrivial example would be to consider the overlap of the two-particle state $|1_k 1_{-k}\rangle = c^\dagger_{-k}b^\dagger_k|0\rangle$ with $|G\rangle$. This works out to

$$\begin{aligned}\langle 1_k 1_{-k}|G\rangle &= \prod_l \langle 0|b_k c_{-k}\,(\alpha_l + \beta_l\, c^\dagger_{-l}b^\dagger_l)|0\rangle \\ &= \left(\prod_{l\neq k}\alpha_l\right)\beta_k \to 0 \quad (\text{as } N\to\infty).\end{aligned} \tag{6.72}$$

where N denotes the number of modes for the system.

What we see here is that the Hilbert space built upon $|0\rangle$ by the action of any (finite) number of $b^\dagger_k, c^\dagger_{-k}$ and the Hilbert space built upon $|G\rangle$ by the action of any (finite) number of $A^\dagger_k$, $B^\dagger_{-k}$ are disjoint, with no overlap in the limit of $N\to\infty$. In particular, the new ground state cannot be unitarily connected to the normal state in this limit. Thus even though formally $|G\rangle$ is obtained by the action of an operator $U = \prod_k(\alpha_k + \beta_k c^\dagger_{-k}b^\dagger_k)$ on $|0\rangle$, U is not unitary since it has zero overlap with any state built on the Fock ground state $|0\rangle$. The two are not unitarily related.

Since the two sets of operators $\{b_k, b^\dagger_k, c_{-k}, c^\dagger_{-k}\}$ and $\{A_k, A^\dagger_k, B_{-k}, B^\dagger_{-k}\}$ have the same algebra of anticommutation rules, but lead to disjoint Hilbert spaces, we see that the two choices of ground state can be taken as leading to two different irreducible representations of the anticommutation rules.

As discussed in Section 4.3, the Stone–von Neumann theorem states that there exists only one unitary irreducible representation for the Heisenberg algebra (4.31) if (a) N is finite or (b) space is simply connected. In Chapter 4, we have considered how the conclusion of a single irreducible representation as given by this theorem is evaded when the condition (b) breaks in the fractional QHE.

In the present case, the algebra we are discussing is not the Heisenberg algebra, rather it is the N dimensional fermionic algebra of anticommutation rules. Nevertheless, the above analysis on the ground state of the BCS theory clearly shows how the canonical commutation rules can have inequivalent representations when the number of modes tends to infinity. A more specific realization of how the premise (a) of the theorem is evaded in the bosonic case (*i.e.*, the Heisenberg algebra) is given by spontaneous symmetry breaking which will be discussed later.

6.4 The transition temperature

We can also get an estimate of the transition temperature as follows. Going back to (6.53), we keep the terms involving N_k, $\tilde{N}_k$ in the coefficient of terms involving $A_k B_{-k}$ and its conjugate. Such terms in H_{int} are given by

$$\begin{aligned} H_{int} = & \sum_{kk'} V_{kk'} \left[\alpha_{k'} \beta_{k'} \left(N_{k'} + \tilde{N}_{k'} - 1 \right) \left(\alpha_k^2 A_k B_{-k} - \beta_k^2 B_{-k}^\dagger A_k^\dagger \right) \right] \\ & + \sum_{kk'} V_{kk'} \left[\left(\alpha_{k'}^2 B_{-k'}^\dagger A_{k'}^\dagger - \beta_{k'}^2 A_{k'} B_{-k'} \right) \alpha_k \beta_k \left(N_k + \tilde{N}_k - 1 \right) \right] + \cdots . \end{aligned} \tag{6.73}$$

The terms involving N_k and $\tilde{N}_k$ are zero at zero temperature. At a nonzero temperature, we can approximate them by the thermal occupation numbers. Then H_{int} simplifies and the gap equation becomes

$$\Delta - \frac{V_0}{2} \sum_k \frac{\Delta}{\sqrt{\epsilon_k^2 + \Delta^2}} \left(1 - \langle N_k \rangle - \langle \tilde{N}_k \rangle \right) = 0 . \tag{6.74}$$

The energies of the single particle excitations are now $\sqrt{\epsilon_k^2 + \Delta^2}$ as we see from (6.64). Thus the occupation numbers are given by the Fermi-Dirac distribution

$$\langle N_k \rangle = \langle \tilde{N}_k \rangle = \frac{1}{e^{E_k/T} + 1} , \qquad E_k = \sqrt{\epsilon_k^2 + \Delta^2} . \tag{6.75}$$

The gap equation becomes

$$\Delta \left[1 - \frac{V_0 G_0}{2} \int d\epsilon \, \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} \, \tanh(\sqrt{\epsilon^2 + \Delta^2}/2T) \right] = 0 . \tag{6.76}$$

As before, $\Delta = 0$ is one solution. The solution with the nonzero gap is what we are interested in. This is given by

$$1 - \frac{V_0 G_0}{2} \int d\epsilon \, \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} \, \tanh(\sqrt{\epsilon^2 + \Delta^2}/2T) = 0 . \tag{6.77}$$

At low temperatures, $\tanh(\sqrt{\epsilon^2 + \Delta^2}/2T) \approx 1$ and we revert to the equation we found earlier. If we start with the nonzero gap at $T = 0$ and raise the temperature, the gap Δ will vanish at a certain critical value of T. This signals the transition from the superconducting to the normal phase. We can thus find the transition temperature by setting Δ equal to zero in (6.77). The resulting equation is

$$\begin{aligned} \frac{1}{V_0 G_0} &= \int_{-\omega_D}^{\omega_D} \frac{d\epsilon}{2\sqrt{\epsilon^2}} \tanh(\sqrt{\epsilon^2}/2T_c) \\ &= \log(\omega_D/2T_c) \tanh(\omega_D/2T_c) - \int_0^{\omega_D/2T_c} ds \, \frac{\log s}{\cosh^2 s} . \end{aligned} \tag{6.78}$$

Usually $\omega_D \gg T_c$, so that the integral in the second term can be approximated by taking $\omega_D/2T_C \to \infty$. We then find

$$\frac{1}{V_0 G_0} \approx \log(\omega_D/2T_c) + \gamma - \log(\pi/4) \tag{6.79}$$

where γ is the Euler-Mascheroni constant ($\gamma \approx 0.5772$). Equation (6.79) gives the transition temperature as

$$T_c = \frac{2e^\gamma}{\pi}\omega_D \exp\left(-\frac{1}{V_0 G_0}\right) = \frac{e^\gamma}{\pi}\Delta_0\,. \tag{6.80}$$

Here $\Delta_0 = 2\omega_D \exp\left(-\frac{1}{V_0 G_0}\right)$ is the value of Δ at $T = 0$. Notice that this formula again displays the isotope effect via the factor ω_D.

Chapter 7

Bosonization and Kac-Moody algebra

7.1 Abelian bosonization

In 1+1 dimensions (one spatial and one temporal dimensions), a system of fermions is equivalent to that of bosons, *i.e.*, bosons and fermions are not distinguishable in 1+1 dimensions. One of the main objectives in this chapter is to show how this equivalence can be understood algebraically.

The notion of spin for particles in d+1 dimensions arises from d-dimensional spatial rotations described by the angular momentum algebra, $[J_i, J_j] = i\epsilon_{ijk}J_k$ $(i, j, k = 1, 2, \cdots, d)$. As well known, representations of the algebra J for $d = 3$ are characterized by the values of spin, $s = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2}, \cdots$. In this case, according to the spin-statistics theorem, *particles of integer spin must be bosons, while particles of half-odd-integer spin must be fermions.* For $d = 2$, however, there are no constraints on the value of spin, that is, s can take fractional values in 2+1 dimensions. In this case, the spin-statistics theorem reduces to the statement that *particles of spin s obey fractional s-statistics.* Such particles are often called anyons. For $d = 1$, which is of our interest here, there is *no notion of spin.* Thus there is no restriction on statistics of particles in 1+1 dimensions. In the following, we shall see how this is realized by showing the equivalence between bosons and fermions in 1+1 dimensions.

Klein transformations

Generally bosons are described by the following algebra

$$[a_k, a_l] = 0\,, \quad [a_k, a_l^\dagger] = \delta_{kl}\,, \quad [a_k^\dagger, a_l^\dagger] = 0 \tag{7.1}$$

where we consider that the bosons in a one-dimensional line segment of length

L so that the label k can be given by $k = \frac{2\pi n}{L}$, with $n \in \mathbf{Z}$ representing the number of modes (or the number of degrees of freedom).

The number operator is defined by

$$N_k = a_k^\dagger a_k \tag{7.2}$$

which satisfies the commutation relations

$$[N_k, a_k] = -a_k\,, \quad [N_k, a_k^\dagger] = a_k^\dagger\,. \tag{7.3}$$

Using the Baker-Cambbell-Hausdorff formula

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2}[A, [A, B]] + \cdots\,, \tag{7.4}$$

we can then obtain the relation

$$\begin{aligned} e^{i\alpha N} a e^{-i\alpha N} &= a + i\alpha[N, a] + \frac{(i\alpha)^2}{2!}[N, [N, a]] + \frac{(i\alpha)^3}{3!}[N, [N, [N, a]]] + \cdots \\ &= a - i\alpha a + \frac{(i\alpha)^2}{2!}a + \frac{(i\alpha)^3}{3!}(-a) + \cdots \\ &= e^{-i\alpha} a \end{aligned} \tag{7.5}$$

where we omit the label k for simplicity. α is a constant. Putting the labels, we can rewrite the above result as

$$a_k e^{i\alpha N_k} = e^{i\alpha} e^{i\alpha N_k} a_k\,. \tag{7.6}$$

Similarly, we find

$$a_k^\dagger e^{i\alpha N_k} = e^{-i\alpha} e^{i\alpha N_k} a_k^\dagger\,. \tag{7.7}$$

We now introduce the operators

$$C_i = \exp\left(i\pi \sum_{k<i} N_k\right) a_i\,, \tag{7.8}$$

$$C_j = \exp\left(i\pi \sum_{k<j} N_k\right) a_j \quad (i < j) \tag{7.9}$$

where require $1 \le i < j \le n$ without losing generality. $C_i C_j$ and $C_j C_i$ can be calculated as

$$\begin{aligned} C_i C_j &= \exp\left(i\pi \sum_{k<i} N_k\right) a_i \exp\left(i\pi \sum_{k<j} N_k\right) a_j \\ &= \exp\left(i\pi \sum_{k<i} N_k\right) \exp\left(i\pi \sum_{k<j} N_k\right) e^{i\pi} a_i a_j\,, \end{aligned} \tag{7.10}$$

$$
\begin{aligned}
C_j C_i &= \exp\left(i\pi \sum_{k<j} N_k\right) a_j \exp\left(i\pi \sum_{k<i} N_k\right) a_i \\
&= \exp\left(i\pi \sum_{k<j} N_k\right) \exp\left(i\pi \sum_{k<i} N_k\right) a_j a_i \,. \qquad (7.11)
\end{aligned}
$$

Notice that by use of (7.6) we can pick up the factor $e^{i\pi}$ in the transposition of a_i and $\exp\left(i\pi \sum_{k<j} N_k\right)$, while that of a_j and $\exp\left(i\pi \sum_{k<i} N_k\right)$ does not bring any extra factors. We can therefore obtain the anticommutation relations for C's:

$$
C_i C_j + C_j C_i = \exp\left(i\pi \sum_{k<i} N_k\right) \exp\left(i\pi \sum_{k<j} N_k\right) (-a_i a_j + a_j a_i) = 0 \,. \quad (7.12)
$$

This shows that we obtain the algebra of fermions from that of bosons by carrying out the transformations of operators indicated in (7.8) and (7.9). These are called the Klein transformations.

Bosonization: an algebraic approach

We now consider the reverse of the above argument. Namely, we begin with fermions in 1+1 dimensions and show the equivalence with bosonic representations. This process is generically called bosonization. There are essentially two approaches for bosonization. This is based on the fact that a quantum theory of any physical system is defined in two ways: one is by a unitary irreducible representation of the algebra of observables; and the other is by correlation functions for the particles of interest. In other words, we can carry out bosonization either by (a) showing the existence of a bosonic realization for the algebra of fermionic observables or by (b) showing equivalence between correlation functions of fermions and those of bosons. Our strategy is to follow the procedure (a).

We first briefly sketch the Dirac theory of free massless fermions in 1+1 dimensions. In this case, the gamma matrices are defined by

$$
\{\gamma_i, \gamma_j\} = 2\eta_{ij}\mathbf{1} \qquad (7.13)
$$

where $i, j = 0, 1$ and η is the Minkowski metric $\eta = (+, -)$ and $\mathbf{1}$ is the (2×2) identity matrix. γ_0 and γ_1 are defined in terms of the (2×2) Pauli matrices

$$
\gamma_0 = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_2 = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \qquad (7.14)
$$

The Dirac equation $i\gamma \cdot \partial\psi = 0$ is then given by

$$
i\sigma_1 \frac{\partial \psi}{\partial x^0} + i(i\sigma_2)\frac{\partial \psi}{\partial x^1} = 0 \qquad (7.15)
$$

where $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ is a two-component Dirac spinor. From $(-i\sigma_1)\times$(7.15), we find

$$\mathbf{1}\frac{\partial\psi}{\partial x^0} - \sigma_3\frac{\partial\psi}{\partial x^1} = 0 \tag{7.16}$$

which can be written down as

$$\frac{\partial\psi_1}{\partial x^0} - \frac{\partial\psi_1}{\partial x^1} = 0\,, \qquad \frac{\partial\psi_2}{\partial x^0} + \frac{\partial\psi_2}{\partial x^1} = 0\,. \tag{7.17}$$

Obviously, the solutions are given by $\psi_1 = \psi_1(x^0 + x^1)$ and $\psi_2 = \psi_2(x^0 - x^1)$. In the momentum-space representation, ψ_1 and ψ_2 correspond to right-moving and left-moving wavefunctions, respectively. Alternatively we can regard these as spinors of right and left chiralties.

In the Dirac theory of fermions, observables are not given by a single field ψ but by biproducts such as $\psi^\dagger\psi$ and $\bar\psi\psi$. In the (1+1)-dimensional case, $\bar\psi$ is defined as

$$\bar\psi = \psi^\dagger\sigma_1\,. \tag{7.18}$$

The fermionic currents are then defined by

$$\begin{aligned} \bar\psi\gamma\psi &= \psi^\dagger\sigma_1\begin{pmatrix} \sigma_1\psi \\ i\sigma_2\psi \end{pmatrix} = \begin{pmatrix} \psi^\dagger\psi \\ -\psi^\dagger\sigma_3\psi \end{pmatrix} \\ &= \begin{pmatrix} \psi_1^\dagger\psi_1 + \psi_2^\dagger\psi_2 \\ -\psi_1^\dagger\psi_1 + \psi_2^\dagger\psi_2 \end{pmatrix} \equiv \begin{pmatrix} J_0 \\ J_1 \end{pmatrix}. \end{aligned} \tag{7.19}$$

We now define $J_\pm$ as

$$\psi_1^\dagger\psi_1 = \frac{J_0 - J_1}{2} \equiv J_-\,, \tag{7.20}$$

$$\psi_2^\dagger\psi_2 = \frac{J_0 + J_1}{2} \equiv J_+ \tag{7.21}$$

and consider an algebra of these observables. As mentioned earlier, our strategy is to obtain a bosonic realization of such an algebra.

Since Dirac spinors are fermions, they obey the equal-time anticommutation relations:

$$\begin{aligned} \psi(x^0,x^1)\psi(x^0,y^1) + \psi(x^0,y^1)\psi(x^0,x^1) &= 0\,, \\ \psi^\dagger(x^0,x^1)\psi^\dagger(x^0,y^1) + \psi^\dagger(x^0,y^1)\psi^\dagger(x^0,x^1) &= 0\,, \\ \psi(x^0,x^1)\psi^\dagger(x^0,y^1) + \psi^\dagger(x^0,y^1)\psi(x^0,x^1) &= \delta(x^1-y^1)\mathbf{1}\,. \end{aligned} \tag{7.22}$$

where the spinor labels are not explicitly shown; the elements of identity $\mathbf{1}$ correspond to these labels. Using (7.22), we can *naively* calculate an equal-time commutation relations for the observables, say J_-'s, as

$$[J_-(x^0,x^1),\, J_-(x^0,y^1)]$$

$$
\begin{aligned}
&= [\psi_1^\dagger(x^0,x^1)\psi_1(x^0,x^1),\ \psi_1^\dagger(x^0,y^1)\psi_1(x^0,y^1)] \\
&= \psi_1^\dagger(x^0,x^1)\left\{\psi_1(x^0,x^1),\ \psi_1^\dagger(x^0,y^1)\right\}\psi_1(x^0,y^1) \\
&\qquad - \psi_1^\dagger(x^0,y^1)\left\{\psi_1(x^0,y^1),\ \psi_1^\dagger(x^0,x^1)\right\}\psi_1(x^0,x^1) \\
&= \psi_1^\dagger(x^0,x^1)\delta(x^1-y^1)\psi_1(x^0,y^1) - \psi_1^\dagger(x^0,y^1)\delta(x^1-y^1)\psi_1(x^0,x^1)\,.
\end{aligned}
\tag{7.23}
$$

Apparently, this quantity vanishes everywhere but, in reality, singularity enters at the same point $x^1 = y^1$. Thus we need to treat it more carefully around this point. We then rewrite J_- as

$$
J_-(x) \;=\; \psi_1^\dagger(x+\epsilon)\,\psi_1(x-\epsilon) \tag{7.24}
$$

where we use composite notations, $x = (x^0, x^1)$ and $x \pm \epsilon = (x^0, x^1 \pm \epsilon)$. We eventually take the limit of $\epsilon \to 0$ at the end of calculation. The commutator (7.23) is now expressed as

$$
\begin{aligned}
[J_-(x),\ J_-(x)] &= \left[\psi_1^\dagger(x+\epsilon)\psi_1(x-\epsilon),\ \psi_1^\dagger(y+\epsilon)\psi_1(y-\epsilon)\right] \\
&= \psi_1^\dagger(x+\epsilon)\psi_1(y-\epsilon)\delta(x-y-2\epsilon) \\
&\qquad - \psi_1^\dagger(y+\epsilon)\psi_1(x-\epsilon)\delta(x-y+2\epsilon)
\end{aligned}
\tag{7.25}
$$

where $\delta(x-y\pm 2\epsilon)$ can be expanded as

$$
\delta(x-y\pm 2\epsilon) = \delta(x-y) \pm 2\epsilon\frac{\partial}{\partial x}\delta(x-y) + \mathcal{O}(\epsilon^2)\,. \tag{7.26}
$$

Thus we need to have

$$
\psi_1^\dagger(x+\epsilon)\,\psi_1(y-\epsilon) \;\sim\; \frac{1}{\epsilon} \quad (\text{as } x \to y) \tag{7.27}
$$

in order to obtain nonzero contributions in (7.25). The relation (7.27) typically occurs in field theories. In the following, we first show this relation and then resume the calculation of (7.25).

Fermion propagators in 1+1 dimensions

There are two ways of evaluating $\psi_1^\dagger(x+\epsilon)\psi_1(y-\epsilon)$. One is by a mode expansion and the other is to use a fermion propagator. Here we take the latter approach. The propagator is defined as

$$
\begin{aligned}
S_-(x,y) &= \langle 0|\psi_1(x^0,x^1)\psi_1^\dagger(y^0,y^1)|0\rangle\theta(x^0-y^0) \\
&\qquad - \langle 0|\psi_1^\dagger(y^0,y^1)\psi_1(x^0,x^1)|0\rangle\theta(y^0-x^0) \\
&= \langle 0|\left[\psi_1(x)\psi_1^\dagger(y)\theta(x^0-y^0) - \psi_1^\dagger(y)\psi_1(x)\theta(y^0-x^0)\right]|0\rangle
\end{aligned}
\tag{7.28}
$$

where $\theta(x^0 - y^0)$ is the Heaviside step function

$$\theta(x^0 - y^0) = \begin{cases} 1 & \text{for } x^0 > y^0\,, \\ 0 & \text{for } x^0 < y^0\,. \end{cases} \tag{7.29}$$

Notice that the minus signs in (7.28) arise from the anticommuting relations at $x^0 = y^0$. As shown in (7.17), ψ_1 obeys the Dirac equation

$$\left(\frac{\partial}{\partial x^0} - \frac{\partial}{\partial x^1}\right)\psi_1 = 0\,. \tag{7.30}$$

Using this equation, we now calculate derivatives of $S_-(x,y)$:

$$\begin{aligned} \frac{\partial}{\partial x^0} S_-(x,y) &= \langle 0|\left[\theta(x^0-y^0)\frac{\partial\psi_1(x)}{\partial x^0}\psi_1^\dagger(y) - \theta(y^0-x^0)\psi_1^\dagger(y)\frac{\partial\psi_1(x)}{\partial x^0}\right]|0\rangle \\ &\quad + \underbrace{\langle 0|\left[\delta(x^0-y^0)\psi_1(x)\psi_1^\dagger(y) + \delta(y^0-x^0)\psi_1^\dagger(y)\psi_1(x)\right]|0\rangle}_{=\delta(x^0-y^0)\langle 0|\{\psi_1(x),\,\psi_1^\dagger(y)\}|0\rangle = \delta(x^0-y^0)\delta(x^1-y^1)} \end{aligned} \tag{7.31}$$

$$\frac{\partial}{\partial x^1} S_-(x,y) = \langle 0|\left[\theta(x^0-y^0)\frac{\partial\psi_1(x)}{\partial x^1}\psi_1^\dagger(y) - \theta(y^0-x^0)\psi_1^\dagger(y)\frac{\partial\psi_1(x)}{\partial x^1}\right]|0\rangle \tag{7.32}$$

where we have used

$$\frac{\partial}{\partial x^0}\theta(x^0-y^0) = \delta(x^0-y^0)\,. \tag{7.33}$$

From (7.31) and (7.32), we find

$$\left(\frac{\partial}{\partial x^0} - \frac{\partial}{\partial x^1}\right) S_-(x,y) = \delta(x^0-y^0)\delta(x^1-y^1)\,. \tag{7.34}$$

The solution is then given by

$$\begin{aligned} S_-(x,y) &= \int \frac{d^2p}{(2\pi)^2}\frac{ie^{-ip_0(x^0-y^0)+ip_1(x^1-y^1)}}{p_0^2-p_1^2}(p_0-p_1) \\ &= i(\partial_0+\partial_1)\underbrace{\int_{-\infty}^{\infty}\frac{dp_0}{2\pi}\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}\frac{ie^{-ip_0(x^0-y^0)+ip_1(x^1-y^1)}}{p_0^2-p_1^2}}_{\equiv G(x,y)}\,. \end{aligned} \tag{7.35}$$

The integral $G(x,y)$ reduces to 2-dimensional Green's function for the Laplace equation if we carry out an analytic continuation from p_1 to $p_E = -ip_1$ as shown in Fig. 7.1.

$$\begin{aligned} G(x,y) &= \int_{-\infty}^{\infty}\frac{dp_0}{2\pi}\int_{-\infty}^{\infty}\frac{dp_E}{2\pi}\frac{e^{-ip_0(x^0-y^0)-ip_E(i(x^1-y^1))}}{p_0^2+p_E^2} \\ &= \frac{1}{4\pi}\log\left((x^0-y^0)^2-(x^1-y^1)^2\right) \end{aligned} \tag{7.36}$$

where the denominator of the integrand should be understood as $(p_E + ip_0 - \eta)(p_E - ip_0 + \eta)$, with η being an infinitesimal positive number (eventually taken to $\eta \to 0$) so that the analytic continuation is properly operated; this kind of rotation from the real axis to the imaginary one is called the Wick rotation.

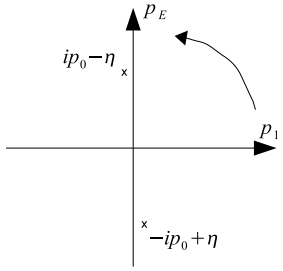

Figure 7.1: Wick rotation from p_1 to p_E

From (7.35) and (7.36), we obtain

$$S_-(x,y) = \frac{i}{2\pi}\frac{1}{(x^0-y^0)+(x^1-y^1)}\,. \tag{7.37}$$

Similarly, we can find

$$S_+(x,y) = \frac{i}{2\pi}\frac{1}{(x^0-y^0)-(x^1-y^1)} \tag{7.38}$$

for J_+'s. The factor $\frac{i}{2\pi}$ can also be checked as follows. Let $S_-(x,y)$ be

$$S_-(x,y) = \frac{c}{(x^0-y^0)+(x^1-y^1)} = c\frac{(x^0-y^0)-(x^1-y^1)}{(x^0-y^0)^2-(x^1-y^1)^2+\epsilon^2} \tag{7.39}$$

where c is a constant to be determined. Since (7.34) means $(\partial_0 - \partial_1)S_-(x,y) = \delta^{(2)}(x-y)$, integrating the equation, we have

$$c(\partial_0-\partial_1)\int_{-\infty}^{\infty}dx^0\int_{-\infty}^{\infty}dx^1\frac{x^0-x^1}{(x^0)^2-(x^1)^2+\epsilon^2} = 1\,. \tag{7.40}$$

The left-hand side can easily calculated as

$$\begin{aligned} & 2c\int_{-\infty}^{\infty}dx^0\int_{-\infty}^{\infty}dx^1\frac{\epsilon^2}{\left[(x^0)^2-(x^1)^2+\epsilon^2\right]^2} \\ = \;& -i2c\int_{-\infty}^{\infty}dx^0\int_{-\infty}^{\infty}dx_E\frac{\epsilon^2}{\left[(x^0)^2+(x_E)^2+\epsilon^2\right]^2} \\ = \;& -i2c\int_0^{\infty}\pi dr^2\frac{\epsilon^2}{(r^2+\epsilon^2)^2} \;=\; -i2\pi c \end{aligned} \tag{7.41}$$

where we use the Wick rotation from x^1 to $x_E = ix^1$. From (7.40) and (7.41), we can confirm the factor of $\frac{i}{2\pi}$ in (7.37) and (7.38).

Evaluating $S_-(x,y)$ on $x^0 > y^0$ and $x^0 \to y^0$, we find

$$\langle 0|\psi_1(x)\psi_1^\dagger(y)|0\rangle = \frac{i}{2\pi}\frac{1}{(x^1-y^1)}\,. \tag{7.42}$$

This is exactly in the form of (7.27).

In general, $\psi_1(x)\psi_1^\dagger(y)$ at $x^0 \to y^0$ can be expressed as

$$\psi_1(x)\psi_1^\dagger(y) = \frac{i}{2\pi}\frac{1}{x^1-y^1} + f(x^1-y^1)\,\mathcal{O} \tag{7.43}$$

where $\langle 0|\mathcal{O}|0\rangle = 0$ and $f(x^1-y^1)$ represents some function. We can claim that $f(x^1-y^1)$ does not give singularities as $x^1 \to y^1$. This can be seen by a simple dimensional analysis. From $\langle 0|\psi_1\psi_1^\dagger|0\rangle \sim \frac{1}{x^1}$ we find that ψ_1 has dimension of $(\text{length})^{-1/2} = L^{-1/2}$. Since the operator $\mathcal{O}$ has zero vacuum expectation value and in a bilinear form in ψ_1 and $\psi_1^\dagger$, this can be given by a linear combination of $\psi_1^\dagger\psi_1$ and $\psi_1^\dagger\nabla\psi_1$. In the former case, the dimension of $\mathcal{O}$ is $\mathcal{O}(\psi_1^\dagger\psi_1) \sim L^{-1}$ and hence we have $f(x^1-y^1) \sim \log(x^1-y^1) \sim L^0$. In the latter case, we find $\mathcal{O}(\psi_1^\dagger\nabla\psi_1) \sim L^{-2}$ and $f(x^1-y^1) \sim L$. Thus, for any choice of $\mathcal{O}$, $f(x^1-y^1)$ does not give singularities at $x^1 \to y^1$.

Algebra of fermionic current $J_-(x)$

Now that we have an explicit form of (7.27), let us resume the calculation of (7.25). From the above results, we find that the leading singularities of $\psi_1^\dagger(x+\epsilon)\psi_1(y-\epsilon)$ and $\psi_1^\dagger(y+\epsilon)\psi_1(x-\epsilon)$ are given by

$$\psi_1^\dagger(x+\epsilon)\psi_1(y-\epsilon) = \frac{i}{2\pi}\frac{1}{(x^1+\epsilon)-(y^1-\epsilon)} \longrightarrow \frac{i}{2\pi}\frac{1}{2\epsilon} \quad (x^1 \to y^1)\,, \tag{7.44}$$

$$\psi_1^\dagger(y+\epsilon)\psi_1(x-\epsilon) = \frac{i}{2\pi}\frac{1}{(y^1+\epsilon)-(x^1-\epsilon)} \longrightarrow \frac{i}{2\pi}\frac{1}{2\epsilon} \quad (x^1 \to y^1)\,. \tag{7.45}$$

Therefore the algebra of $J_-(x)$'s is given by

$$[J_-(x^0,x^1),\, J_-(x^0,y^1)] \;=\; \frac{-i}{\pi}\frac{\partial}{\partial x^1}\delta(x^1-y^1)\,. \tag{7.46}$$

Similarly, when we define J_+ by $J_+(x) = \psi_2^\dagger(x+\epsilon)\psi_2(x-\epsilon)$ as in (7.24), we find the algebra of J_+'s is given by

$$[J_+(x^0,x^1),\, J_+(x^0,y^1)] \;=\; \frac{i}{\pi}\frac{\partial}{\partial x^1}\delta(x^1-y^1)\,. \tag{7.47}$$

Thus the algebra (7.46) or (7.47) provides the algebra of fermionic observables in (1+1)-dimensions. This algebra corresponds to an Abelian version of the so-called Kac-Moody algebra. In the next section, we shall consider non-Abelian generalization of this algebra.

Bosonic realization of $J_-(x)$

We now consider a bosonic realization of the observable $J_-(x^0, x^1)$, introducing a bosonic field $\phi(x^0, x^1)$ represented by a real variable. The time derivative of this field $\frac{\partial}{\partial x^0}\phi(x^0, x^1) = \dot{\phi}(x^0, x^1)$ is an independent variable and we have the following equal-time commutation relations:

$$\begin{aligned} \left[\phi(x^0, x^1),\, \phi(x^0, y^1)\right] &= 0\,, \\ \left[\dot{\phi}(x^0, x^1),\, \dot{\phi}(x^0, y^1)\right] &= 0\,, \\ \left[\phi(x^0, x^1),\, \dot{\phi}(x^0, y^1)\right] &= i\delta(x^1 - y^1)\,. \end{aligned} \tag{7.48}$$

These are equivalent to the commutation rules (7.1) for the bosonic operators a_k and $a_k^\dagger$ if we use the mode expansion

$$\phi(x) = \sum_k \frac{1}{\sqrt{2k_0 L}} \left(a_k e^{-ikx} + a_k^\dagger e^{ikx}\right). \tag{7.49}$$

This can be easily checked with $ikx = ik_0x^0 - ik_1x^1$ and the replacement

$$\sum_k \to L \int \frac{dk_1}{2\pi} \tag{7.50}$$

in the limit of $L \to \infty$.

Now the bosonic realization of J_- is given by

$$J_- = \alpha \left(\frac{\partial \phi}{\partial x^0} - \frac{\partial \phi}{\partial x^1}\right) \tag{7.51}$$

where α is some constant. Indeed, we can calculate the commutation relation of J_-'s as

$$\begin{aligned} & [J_-(x^0, x^1),\, J_-(x^0, y^1)] \\ = \;& \alpha^2 \left[\dot{\phi}(x) - \frac{\partial}{\partial x^1}\phi(x),\, \dot{\phi}(y) - \frac{\partial}{\partial y^1}\phi(y)\right] \\ = \;& -\alpha^2 \frac{\partial}{\partial y^1}\left[\dot{\phi}(x),\, \phi(y)\right] - \alpha^2 \frac{\partial}{\partial x^1}\left[\phi(x),\, \dot{\phi}(y)\right] \\ = \;& i\alpha^2 \frac{\partial}{\partial y^1}\delta(y^1 - x^1) - i\alpha^2 \frac{\partial}{\partial x^1}\delta(x^1 - y^1) \\ = \;& -i\alpha^2 \frac{\partial}{\partial x^1}\left(\delta(y^1 - x^1) + \delta(x^1 - y^1)\right) \\ = \;& -i2\alpha^2 \frac{\partial}{\partial x^1}\delta(x^1 - y^1) \end{aligned} \tag{7.52}$$

where we use the relation $\frac{\partial}{\partial x^1}\delta(x^1 - y^1) = -\frac{\partial}{\partial y^1}\delta(x^1 - y^1)$ which is in consistent with the definition of the delta function

$$\delta(x^1 - y^1) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ik(x^1 - y^1)}\,. \tag{7.53}$$

From (7.46) and (7.52), we can determine α by $\alpha^2 = \frac{1}{2\pi}$. Therefore, one of the bosonic realizations of the fermionic observable (or the current) $J_- = \psi_1^\dagger \psi_1$ is given by

$$J_-(x) = \psi_1^\dagger(x)\psi_1(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{\partial}{\partial x^0} - \frac{\partial}{\partial x^1}\right)\phi(x)\,. \tag{7.54}$$

Similarly, we can also express the bosonization of $J_+ = \psi_2^\dagger \psi_2$ as

$$J_+(x) = \psi_2^\dagger(x)\psi_2(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{\partial}{\partial x^0} + \frac{\partial}{\partial x^1}\right)\phi(x)\,. \tag{7.55}$$

Mapping of fermionic and bosonic fields

The expression (7.54) shows a mapping between the fermionic current and the bosonic scalar field, $J_-(x) \leftrightarrow \phi(x)$. For the fermionic field $\psi_1(x)$ can we construct such a mapping $\psi_1(x) \leftrightarrow \phi(x)$. In the Abelian case it is possible; there exists an exact mapping (by Mandelstam),

$$\psi_1(x) \;=\; A\,\exp\left(i\phi(x) + i\pi\int_{x^1}^{\infty}\dot{\phi}(x^0,\tilde{x}^1)d\tilde{x}^1\right) \equiv e^{i\Phi_1(x)} \tag{7.56}$$

where A is a normalization constant. Using this expression, we can check the anti-commutation relations of $\psi_1^\dagger(x)$ and $\psi_1(y)$:

$$\begin{aligned}
\psi_1^\dagger(x)\psi_1(y) &= A^2 e^{-i\Psi_1(x)} e^{i\Phi_1(y)} \\
&= A^2 e^{-i\Phi_1(x)+i\Phi_1(y)} e^{\frac{1}{2}[\Phi_1(x),\Phi_1(y)]} \\
&= -\psi_1(y)\psi_1^\dagger(x)
\end{aligned} \tag{7.57}$$

where we use the commutation relations of ϕ's in (7.48). We also use the relation

$$e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]} \tag{7.58}$$

for arbitrary operators A and B, with $[A,B]$ being a c-number. Notice that the factor $\exp\left(i\pi\int_{x^1}^{\infty}\dot{\phi}(x^0,\tilde{x}^1)d\tilde{x}^1\right)$ plays an analogous role as the factor of $\exp\left(i\pi\sum_{k<i}N_k\right)$ in (7.8) of the Klein transformations.

One can also define the mapping as

$$\psi_1(x) \;=\; A\,\exp\left(i\sqrt{\pi}\phi(x) + i\sqrt{\pi}\int_{x^1}^{\infty}\dot{\phi}(x^0,\tilde{x}^1)d\tilde{x}^1\right) \tag{7.59}$$

which does not change the computation (7.57). From the chirality relation, we can define the other spinor $\psi_2(x)$ as

$$\psi_2(x) \;=\; A\,\exp\left(-i\sqrt{\pi}\phi(x) + i\sqrt{\pi}\int_{x^1}^{\infty}\dot{\phi}(x^0,\tilde{x}^1)d\tilde{x}^1\right)\,. \tag{7.60}$$

One way to determine the normalization factor A is to use the relation (7.44), leading to $A = \frac{1}{\sqrt{2\pi(2\epsilon)}}$. With this normalization, we can explicitly calculate $\psi_1^\dagger(x+\epsilon)\psi_1(x-\epsilon)$ as

$$\begin{aligned}
\psi_1^\dagger(x+\epsilon)\psi_1(x-\epsilon) &= \frac{1}{2\pi}\frac{1}{2\epsilon}e^{-i\Phi_1(x+\epsilon)}e^{i\Phi_1(x-\epsilon)} \\
&= \frac{i}{2\pi}\frac{1}{2\epsilon}e^{-i(\Phi_1(x+\epsilon)-\Phi_2(x-\epsilon))} \\
&= \frac{i}{2\pi}\frac{1}{2\epsilon}+\frac{1}{2\pi}\frac{\partial}{\partial x^1}\Phi_1(x^0,x^1)+\mathcal{O}(\epsilon) \\
&= \frac{i}{2\pi}\frac{1}{2\epsilon}-\frac{1}{\sqrt{2}}J_-(x)+\mathcal{O}(\epsilon) \qquad (7.61)
\end{aligned}$$

where $\Phi_1(x)$ is now given by

$$\Phi_1(x) = \sqrt{\pi}\phi(x) + \sqrt{\pi}\int_{x^1}^{\infty}\dot{\phi}(x^0,\tilde{x}^1)d\tilde{x}^1 \qquad (7.62)$$

and we use the relation (7.54) to obtain

$$\frac{\partial}{\partial x^1}\Phi_1 = \sqrt{\pi}\left(\frac{\partial}{\partial x^1}\phi - \frac{\partial}{\partial x^0}\phi\right) = -\sqrt{2}\pi J_-(x)\,. \qquad (7.63)$$

Similarly, we find the relation

$$\psi_1(x+\epsilon)\psi_1^\dagger(x-\epsilon) = \frac{i}{2\pi}\frac{1}{2\epsilon}+\frac{1}{\sqrt{2}}\psi_1^\dagger(x+\epsilon)\psi_1(x-\epsilon)+\mathcal{O}(\epsilon)\,. \qquad (7.64)$$

This shows a specific expression of (7.43) in our choice of ψ_1 in (7.59).

Mode expansion

We now consider the mode expansion (7.49) of the bosonic field $\Phi_1(x)$.

$$\Phi_1(x) = \sum_{k\geq 1}\frac{1}{\sqrt{2k_0 L}}\left(a_k e^{-ik_0x^0-ik_1x^1}+a_k^\dagger e^{ik_0x^0+ik_1x^1}\right) \qquad (7.65)$$

where we use the Euclidean metric. For the massless case ($m=0$), $k_0 = \sqrt{k_1^2+m^2} = |k_1|$ and Φ_1 is scale invariant. $\Phi_1(x)$ is then written as

$$\Phi_1(x) = \frac{1}{\sqrt{2L}}\sum_{k\geq 1}\frac{1}{\sqrt{k}}\left(a_k e^{-ik(x^0+x^1)}+a_k^\dagger e^{ik(x^0+x^1)}\right). \qquad (7.66)$$

We now compactify the spacial direction $0 \leq x^1 \leq 2\pi$ ($L = 2\pi$) and impose periodicity in Φ_1. The mode expansion can then be expanded in terms of the integer $n \in \mathbf{Z}$.

$$\Phi_1(x) = \frac{1}{\sqrt{4\pi}}\sum_{n<0}\frac{1}{\sqrt{|n|}}\left(a_n e^{-in(x^0+x^1)}+a_n^\dagger e^{in(x^0+x^1)}\right)$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{4\pi}}\sum_{n>0}\frac{1}{\sqrt{n}}\left(a_n e^{-in(x^0+x^1)}+a_n^\dagger e^{in(x^0+x^1)}\right)+\phi_0\\
=\;&\frac{1}{\sqrt{\pi}}\sum_{n>0}\frac{1}{\sqrt{n}}\left(a_n e^{-in(x^0+x^1)}+a_n^\dagger e^{in(x^0+x^1)}\right)+\phi_0 \qquad (7.67)
\end{aligned}
$$

where we use $a_{-m}=a_m^\dagger$, $a_{-m}^\dagger=a_m$ for $m>0$. We also include the zero mode ϕ_0 for the free bosonic field Φ_1. The canonical conjugate of the zero mode ϕ_0 is defined by $\pi_0=\dot{\phi}_0$, satisfying the canonical commutation relation

$$[\phi_0,\pi_0]=i\,. \qquad (7.68)$$

Time-evolution of the zero mode is then given by $\phi_0+\pi_0 x^0$. Notice that $\Phi_1(x)$ is a function of (x^0+x^1) since we have the relation $\psi_1(x)=\psi_1(x^0+x^1)\sim\exp(i\Phi_1(x^0+x^1))$. The proper contribution to the zero-mode is then given by $\phi_0+\pi_0(x^0+x^1)$. These types of bosons are sometimes called chiral bosons. The mode expansion of Φ_1 can then be redefined as

$$\Phi_1(x^0+x^1)=\sum_{n>0}\frac{1}{\sqrt{n}}\left(a_n e^{-in(x^0+x^1)}+a_n^\dagger e^{in(x^0+x^1)}\right)+\phi_0+\pi_0(x^0+x^1)\,. \qquad (7.69)$$

As mentioned earlier, one can calculate the propagator $\langle 0|\psi_1\psi_1^\dagger|0\rangle$ by use of the above mode expansion. What is usually done is to introduce the complex variable

$$z=e^{i(x^0+x^1)} \qquad (7.70)$$

where x^1 is spacially compactified. This change of variables in (1+1) dimensions is referred to as the radial quantization. In terms of z the mode expansion can be expressed as

$$\Phi_1(z)=\sum_{n>0}\frac{1}{\sqrt{n}}\left(a_n z^{-n}+a_n^\dagger z^n\right)+\phi_0-i\pi_0\log z\,. \qquad (7.71)$$

From (7.59) we may define $\psi_1(z)$ as $\psi_1(z)=\exp(i\Phi_1(z))$, absorbing the scale factor into Φ_1. We can then calculate $\psi_1^\dagger(z)|0\rangle$ as

$$
\begin{aligned}
\psi_1^\dagger(z)|0\rangle &= \exp\left(-i\sum_{n\geq 1}\frac{1}{\sqrt{n}}\left(a_n z^{-n}+a_n^\dagger z^n\right)\right)e^{-i\phi_0-\pi_0\log z}|0\rangle\\
&= \exp\left(-i\sum_{n\geq 1}\frac{a_n^\dagger z^n}{\sqrt{n}}\right)\exp\left(\frac{1}{2}\sum_{n\geq 1}\frac{1}{n}\right)e^{-i\phi_0}z^{\frac{1}{2}}|0\rangle \qquad (7.72)
\end{aligned}
$$

where we have used (7.1), (7.58), (7.68) and the definition of the vacuum state $a_n|0\rangle=\pi_0|0\rangle=0$. Obviously, the factor of $\sum_{n\geq 1}\frac{1}{n}$ diverges. To obtain finite physical quantities out of $\psi_1(z)$, we need to redefine $\psi_1(z)$ using a normal ordered product of $e^{i\Phi_1(z)}$. Namely, we have

$$\psi_1(z) = :e^{i\Phi_1(z)}:$$

$$\equiv \quad \exp\left(i\sum_{n\geq 1}\frac{a_n^\dagger z^n}{\sqrt{n}}\right)\exp\left(i\sum_{n\geq 1}\frac{a_n z^{-n}}{\sqrt{n}}\right)e^{i\phi_0}e^{\pi_0\log z}\,. \quad (7.73)$$

This operator corresponds to the vertex operator of chiral bosons in conformal field theory. The computation of the propagator $\langle 0|\psi_1\psi_1^\dagger|0\rangle$ is then carried out as

$$\begin{aligned}
\langle 0|\psi_1(z)\psi_1^\dagger(w)|0\rangle &= \langle 0|\exp\left(i\sum_{n\geq 1}\frac{a_n z^{-n}}{\sqrt{n}}\right)\exp\left(-i\sum_{m\geq 1}\frac{a_m^\dagger z^w}{\sqrt{m}}\right)e^{\pi_0\log z}e^{-i\phi_0}|0\rangle \\
&= \langle 0|\exp\left(\sum_{n,m}\frac{z^{-n}w^m}{\sqrt{nm}}[a_n,a_m^\dagger]\right)e^{-i[\pi_0,\phi_0]\log z}|0\rangle \\
&= \exp\left(\sum_{n\geq 1}\frac{1}{n}\left(\frac{w}{z}\right)^n\right)z^{-1} \\
&= e^{-\log\left(1-\frac{w}{z}\right)}z^{-1} = \frac{1}{z-w} \qquad (7.74)
\end{aligned}$$

where we use the relation

$$e^A e^B = e^B e^A e^{[A,B]}\,. \quad (7.75)$$

Substituting $z = e^{i(x^0+x^1)}$ and $w = e^{i(y^0+y^1)}$, we find

$$\frac{1}{z-w} = \frac{1}{e^{i(x^0+x^1)} - e^{i(y^0+y^1)}} \sim \frac{1}{(x^0-y^0)+(x^1-y^1)} \qquad (z\to w) \quad (7.76)$$

This is in agreement with the previous result on the propagator $S_-(x,y)$ in (7.37).

7.2 Non-Abelian bosonization

In this section, we derive a non-Abelian version of the fermionic current algebra or the Kac-Moody algebra. This can be carried out by considering N free fermions. In the non-Abelian case the spinors are expressed as

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \to \begin{pmatrix}\psi_1^a\\ \psi_2^a\end{pmatrix} \qquad (a = 1,2,\cdots,N) \quad (7.77)$$

The currents of interest are then written as $J_-^{ab} \sim \psi_1^{\dagger a}\cdot\psi_1^b$. We can consider these as N^2 hermitian operators. We now introduce $SU(N)$ symmetry and classify the N^2 currents as

$$\begin{cases} \psi_1^{\dagger a}\psi_1^a & \text{trace-like current} \\ \psi_1^{\dagger a}(T^\alpha)_{ab}\psi_1^b & SU(N)\text{ currents} \end{cases} \quad (7.78)$$

where T^α ($\alpha = 1, 2, \cdots, N^2-1$) are the generators of the $SU(N)$ group, represented by traceless $N \times N$ hermitian matrices. The $SU(N)$ algebra is then given by

$$[T^\alpha,\, T^\beta] = i f^{\alpha\beta\gamma} T^\gamma \,. \tag{7.79}$$

As usual, the normalization of T^α is chosen by

$$\mathrm{Tr}(T^\alpha T^\beta) = \frac{1}{2}\delta^{\alpha\beta} \,. \tag{7.80}$$

In analogy with the Abelian case (7.24), we can define the non-Abelian currents as

$$J^\alpha_-(x) \;=\; \psi_1^\dagger(x+\epsilon)\, T^\alpha\, \psi_1(x-\epsilon) \,. \tag{7.81}$$

The equal-time commutation relation can be calculated as

$$\begin{aligned}
& [J^\alpha_-(x^0,x^1),\, J^\beta_-(x^0,y^1)] \\
= \;& \Big[\psi_1^\dagger(x+\epsilon)T^\alpha\psi_1(x-\epsilon),\, \psi_1^\dagger(y+\epsilon)T^\beta\psi_1(y-\epsilon)\Big] \\
= \;& \psi_1^{\dagger\,a}(x+\epsilon)(T^\alpha)_{ab}\left\{\psi_1^b(x-\epsilon),\, \psi_1^{\dagger\,c}(y+\epsilon)\right\}(T^\beta)_{cd}\,\psi_1^d(y-\epsilon) \\
& \quad - \psi_1^{\dagger\,a}(y+\epsilon)(T^\beta)_{ab}\left\{\psi_1^b(y-\epsilon),\, \psi_1^{\dagger\,c}(x+\epsilon)\right\}(T^\alpha)_{cd}\,\psi_1^d(x-\epsilon) \\
= \;& \psi_1^\dagger(x+\epsilon)T^\alpha T^\beta\psi_1(y-\epsilon)\delta(x-y-2\epsilon) \\
& \quad - \psi_1^\dagger(y+\epsilon)T^\beta T^\alpha\psi_1(x-\epsilon)\delta(x-y+2\epsilon)
\end{aligned} \tag{7.82}$$

where we have used

$$\left\{\psi_1^a(x^0,x^1),\, \psi_1^{\dagger\,b}(x^0,y^1)\right\} = \delta^{ab}\delta(x^1-y^1) \,. \tag{7.83}$$

The expression (7.82) is an non-Abelian version of (7.25) and can similarly be simplified as

$$[J^\alpha_-(x),\, J^\beta_-(y)] = i f^{\alpha\beta\gamma} J^\gamma_-(x)\delta(x-y) - \frac{i}{2\pi}\delta^{\alpha\beta}\frac{\partial}{\partial x}\delta(x-y) \tag{7.84}$$

where we use (7.79) and (7.80). Note that we have also used non-Abelian version of (7.44) and (7.45) to obtain leading singularities of $\psi_1^{\dagger\,a}(x+\epsilon)\psi_1^b(y-\epsilon)$ and $\psi_1^{\dagger\,a}(y+\epsilon)\psi_1^b(x-\epsilon)$, *i.e.*,

$$\psi_1^{\dagger\,a}(x+\epsilon)\psi_1^b(y-\epsilon) \;\longrightarrow\; \frac{i}{2\pi}\frac{1}{2\epsilon}\delta^{ab} \quad (x^1 \to y^1) \,, \tag{7.85}$$

$$\psi_1^{\dagger\,a}(y+\epsilon)\psi_1^b(x-\epsilon) \;\longrightarrow\; \frac{i}{2\pi}\frac{1}{2\epsilon}\delta^{ab} \quad (x^1 \to y^1) \,. \tag{7.86}$$

The commutation relation (7.84) gives a non-Abelian version of the Kac-Moody algebra. This is also known as the current algebra. Explicitly in the coordinates, the current algebra (7.84) can be written as

$$[J^\alpha_-(x^0,x^1),\, J^\beta_-(x^0,y^1)] = i f^{\alpha\beta\gamma} J^\gamma_-(x^0,x^1)\delta(x^1-y^1) - \frac{i}{2\pi}\delta^{\alpha\beta}\frac{\partial}{\partial x^1}\delta(x^1-y^1) \,. \tag{7.87}$$

For the completion of our argument, we can similarly calculate the algebra of $J^\alpha_+(x) = \psi^\dagger_2(x+\epsilon)T^\alpha\psi_2(x-\epsilon)$, leading to the expression

$$[J^\alpha_+(x^0,x^1),\, J^\beta_+(x^0,y^1)] = if^{\alpha\beta\gamma}J^\gamma_+(x^0,x^1)\delta(x^1-y^1) + \frac{i}{2\pi}\delta^{\alpha\beta}\frac{\partial}{\partial x^1}\delta(x^1-y^1)\,. \tag{7.88}$$

We now introduce a new current J^α_n in terms of $J^\alpha_-(x)$ by parametrizing as

$$J^\alpha_n = J^\alpha_n(x^0) = \int_0^{2\pi R} e^{in\frac{x^1}{R}}J^\alpha_-(x^0,x^1)dx^1 \tag{7.89}$$

where n is an integer. Using the new current $J^\alpha_n(x^0)$, we can obtain another realization of the current algebra as follows. Multiplying the algebra (7.87) by $\exp\left(in\frac{x^1}{R}\right)\exp\left(im\frac{y^1}{R}\right)$ and integrating over x^1 and y^1 around the circles of radius R, we find

$$\begin{aligned}
[J^\alpha_n,\, J^\beta_m] &= if^{\alpha\beta\gamma}\int_{x^1}\int_{y^1} e^{in\frac{x^1}{R}+im\frac{y^1}{R}}J^\gamma(x)\delta(x^1-y^1)dx^1dy^1 \\
&\quad -\frac{i}{2\pi}\delta^{\alpha\beta}\int_{x^1}\int_{y^1} e^{in\frac{x^1}{R}}\frac{\partial}{\partial x^1}\delta(x^1-y^1)e^{im\frac{y^1}{R}}dx^1dy^1 \\
&= if^{\alpha\beta\gamma}J^\gamma_{n+m} + m\delta^{\alpha\beta}\delta_{n+m,0}
\end{aligned} \tag{7.90}$$

where we use

$$\int_0^{2\pi R} e^{i(n+m)\frac{x^1}{R}}dx^1 \;=\; 2\pi R\,\delta_{n+m,0}\,. \tag{7.91}$$

Notice that for $m=n=0$ the algebra reduces to the ordinary $SU(N)$ algebra. We can then regard the non-Abelian Kac-Moody algebra as an extended version of the $SU(N)$ algebra.

In mathematical literature there is the so-called level number k associated with the current algebra. In this case, the algebra is expressed as

$$[J^\alpha_n,\, J^\beta_m] = if^{\alpha\beta\gamma}J^\gamma_{n+m} + km\delta^{\alpha\beta}\delta_{n+m,0}\,. \tag{7.92}$$

It is known that the level number k of the current algebra needs to be a positive integer due to the requirement of unitarity. The level number k can be embedded into our framework with replacement of the range of $x^1 \in [0,2\pi R]$ by $x^1 \in [0,2\pi kR]$. The above result (7.90) corresponds to the case of $k=1$. In fact, there is a theorem (by Kac) that states, "*The Kac-Moody algebra of $k=1$ has only one unitary irreducible representation.*" This guarantees the uniqueness of the unitary irreducible representation of the algebra (7.90). In this sense Kac's theorem is analogous to the Stone-von Neumann theorem for the Heisenberg algebra which we have discussed in chapters 1, 4 and 6.

Bosonic realization of $J^\alpha_-(x)$

The above derivation of the non-Abelian Kac-Moody algebra is presented in terms of the fermionic currents $J_-^\alpha = \psi_1^\dagger T^\alpha \psi_1$. As indicated in (7.54), for the abelian case the bosonic realization of $J_- = \psi_1^\dagger \psi_1$ is expressed as

$$\psi_1^\dagger \psi_1 \;=\; J_-(x) \;=\; \frac{1}{\sqrt{2\pi}} \partial_- \phi(x) \;=\; -\frac{i}{\sqrt{2\pi}} \partial_- g\, g^{-1}\,, \tag{7.93}$$

$$g \;=\; g(x) \;=\; e^{i\phi(x)} \tag{7.94}$$

where $\partial_- = \frac{\partial}{\partial x^0} - \frac{\partial}{\partial x^1}$. Here $g(x)$ can be considered as an element of the Abelian group $U(1)$. In this parametrization, a non-Abelian extension is straightforward, which can be expressed as

$$\psi^\dagger T^\alpha \psi \;=\; J_-^\alpha(x) \;=\; \frac{1}{\sqrt{2\pi}} \partial_- \phi^\alpha(x) \;=\; -\frac{i}{\sqrt{2\pi}} \partial_- g\, g^{-1}\,, \tag{7.95}$$

$$g \;=\; g(x) \;=\; e^{iT^\alpha \phi^\alpha(x)} \tag{7.96}$$

where $g(x)$ is now an element of the $SU(N)$ group. The relation (7.95) represents the bosonization of the non-Abelican current. There exists a bosonic action for $g(x)$ which describes a two-dimensional physical system, being equivalent to the fermionic scheme in terms of $J_-^\alpha(x)$. This action is called the Wess-Zumino-Witten (WZW) action. In the next section, we shall briefly discuss dynamics of this action.

The expression (7.95) for the non-Abelian bosonization shows a mapping between a fermionic current and a bosonic scalar field, $J_- \leftrightarrow \phi$. As we have seen in (7.59) and (7.60), we can construct an exact mapping between fermionic and bosonic fields $\psi \leftrightarrow \phi$ in the Abelian case. For the non-Abelian case, however, existence of such a function is not known.

7.3 Wess-Zumino-Witten action

In this section, we introduce the Wess-Zumino-Witten (WZW) action and present a simple derivation for the equation of motion. The WZW action is defined by

$$S \;=\; \frac{k}{8\pi} \int d^2x\, \mathrm{Tr}\left(\partial_\mu g\, \partial_\mu g^{-1}\right) + \Gamma_{WZ} \tag{7.97}$$

$$\Gamma_{WZ} \;=\; \frac{k}{4\pi} \int_0^1 d\lambda \int d^2x\, \mathrm{Tr}\left(g^{-1}\partial_\lambda g\, g^{-1}\partial_\mu g\, g^{-1}\partial_\nu g\right) \epsilon^{\mu\nu} \tag{7.98}$$

where we use the Minkowski metric ($\mu, \nu = 0, 1$). As in the previous section $g = g(x) = \exp\left(iT^\alpha \phi^\alpha(x)\right)$ denotes an element of the $SU(N)$ group and $k \in \mathbf{N}$ is the level number. The second term Γ_{WZ} is called the Wess-Zumino term. In the WZ term $g = g(x)$ also depends on the extra variable $\lambda \in [0, 1]$, *i.e.*, $g = g(x, \lambda)$, with the boundary condition:

$$g(x, 0) = 1\,, \quad g(x, 1) = g(x) = e^{iT^\alpha \phi^\alpha(x)}\,. \tag{7.99}$$

We can interpret λ as the third coordinate of the (2+1)-dimensional compact spacetime, $(x,\lambda)\to(x^0,x^1,x^2)$ with $\lambda=x^2$, satisfying $0\le x^2\le 1$. The original (1+1)-dimensional spacetime then corresponds to the surface of 3-dimensional sphere S^3 by stereographic-type projection. The condition (7.99) guarantees $g=1$ at $|\vec{x}|\to\infty$. In this case $g(x,\lambda)$ gives rise to a map from S^3 to $G=SU(N)$, *i.e.*, $g(x,\lambda):S^3\to G$ and Γ_{WZ} can be expressed as

$$\begin{aligned}\Gamma_{WZ} &= \frac{k}{12\pi}\int_{S^3} d^3x\,\mathrm{Tr}\left(g^{-1}\partial_\lambda g\,g^{-1}\partial_\mu g\,g^{-1}\partial_\nu g\right)\epsilon^{\mu\nu\lambda} \\ &= -2\pi k Q[g] \qquad (7.100)\\ Q[g] &= -\frac{1}{24\pi^2}\int_{S^3} d^3x\,\mathrm{Tr}\left(g^{-1}\partial_\lambda g\,g^{-1}\partial_\mu g\,g^{-1}\partial_\nu g\right)\epsilon^{\mu\nu\lambda} \qquad (7.101)\end{aligned}$$

where the indices now denote (2+1)-dimensional coordinates ($\mu,\nu,\lambda=0,1,2$). The integral value of $Q[g]$ proves to be an integer which is known as the winding number of the map $g:S^3\to G$. The WZ term Γ_{WZ} is then given by 2π times integer for $k\in\mathbf{N}$. Since the theory is defined by using the weight $\exp(-iS)$ to construct functional integrals, it is well-defined on the original (1+1)-dimensional configuration, being independent of how we interpolate $g(x,\lambda)$ in terms of λ as long as (7.99) is satisfied.

To obtain the equation of motion we first vary $g(x)$. From $g+\delta g = e^{iT^\alpha(\phi^\alpha+\delta\phi^\alpha)}$, we find

$$\delta g = g\xi\,,\qquad \delta g^{-1} = -\xi g^{-1} \qquad (7.102)$$

where we define ξ as

$$\xi = iT^\alpha\delta\phi^\alpha = g^{-1}\delta g\,. \qquad (7.103)$$

The variation of the first term in (7.97) is then expressed as

$$\begin{aligned}\delta S^{(1)} &= \frac{k}{8\pi}\int d^2x\,\mathrm{Tr}\,\left[\partial_\mu(g\xi)\partial_\mu g^{-1}-\partial_\mu g\partial_\mu(\xi g^{-1})\right] \\ &= \frac{k}{8\pi}\int d^2x\,\mathrm{Tr}\,(\partial_\mu\xi)\left[g\partial_\mu g^{-1}-\partial_\mu g\,g^{-1}\right] \\ &= \frac{k}{4\pi}\int d^2x\,\mathrm{Tr}\,\xi\partial_\mu(\partial_\mu g\,g^{-1}) \\ &= \frac{k}{4\pi}\int d^2x\,\mathrm{Tr}\,\xi\left[\partial_0(\partial_0 g\,g^{-1})-\partial_1(\partial_1 g\,g^{-1})\right]\,. \qquad (7.104)\end{aligned}$$

On the other hand, the variation of the second term becomes

$$\begin{aligned}\delta\Gamma_{WZ} &= \frac{k}{4\pi}\int_0^1 d\lambda\int d^2x\,\mathrm{Tr}\,\partial_\lambda\left(\xi\,g^{-1}\partial_\mu g\,g^{-1}\partial_\nu g\right)\epsilon^{\mu\nu} \\ &= \frac{k}{4\pi}\int d^2x\,\mathrm{Tr}\,\xi\left(g^{-1}\partial_0 g\,g^{-1}\partial_1 g-g^{-1}\partial_1 g\,g^{-1}\partial_0 g\right) \\ &= \frac{k}{4\pi}\int d^2x\,\mathrm{Tr}\,\xi\left[\partial_0(\partial_1 g\,g^{-1})-\partial_1(\partial_0 g\,g^{-1})\right]\,. \qquad (7.105)\end{aligned}$$

From these we find that the equation of motion $\delta S = \delta S^{(1)} + \delta\Gamma_{WZ} = 0$ becomes

$$\partial_0(\partial_0 g g^{-1} + \partial_1 g g^{-1}) - \partial_1(\partial_0 g g^{-1} + \partial_1 g g^{-1}) = 0\,. \tag{7.106}$$

This can also be written as

$$\partial_-(\partial_+ g\, g^{-1}) = 0 \tag{7.107}$$

where $\partial_\pm = \partial_0 \pm \partial_1$. In terms of ϕ^α this can be expressed as $\partial_-\partial_+\phi^\alpha = 0$. Thus (7.107) is considered as a matrix generalization of the free-field equation of motion. We can easily find that the general solution of (7.107) is given by

$$g(x) \;=\; V(x_+)U(x_-) \tag{7.108}$$

where $V(x_+)$ and $U(x_-)$ are arbitrary functions of $x_+ = x^0 + x^1$ and $x_- = x^0 - x^1$, respectively. Note that for the Abelian case $g = e^{i\phi}$, the equation (7.107) reduces to $\partial_-\partial_+\phi = 0$ and the general solution becomes $\phi(x) = \phi_1(x_+) + \phi_2(x_-)$. The solution (7.108) corresponds to the non-Abelian version of this Abelian solution. As we have seen in (7.95), we can define the fermionic current J^α_- as $J^\alpha_-(x) = \frac{-i}{\sqrt{2\pi}}\partial_- g\, g^{-1}$. By construction, J^α_-'s obviously obey the non-Abelian Kac-Moody algebra (7.87).

Part II

Geometric Methods

Chapter 8

Curved manifolds, metric and Riemannian manifolds

8.1 Curved manifolds

We first consider an n-dimensional curved manifold $\mathcal{M}$. Globally $\mathcal{M}$ can be considered as an n-dimensional topological space. Locally $\mathcal{M}$ approximates to an n-dimensional flat space $\mathbf{R}^n$. In a neighbourhood of such a flat space, we can define n local coordinates parametrized by n real variables. There are a number of such neighbourhoods or so-called patches in the manifold $\mathcal{M}$. In an overlap of two different patches, there is a transition function that relates the corresponding two sets of local coordinates. This function gives a one-to-one invertible map between the two patches. This is also a smooth function, *i.e.*, it is an infinitely differentiable function, a C^∞-function. Thus in a local description the curved manifold can be considered as a differentiable manifold. A rough sketch of $\mathcal{M}$ is shown in Fig. 8.1.

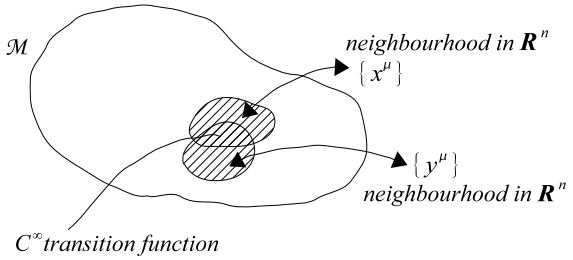

Figure 8.1: A sketch of a curved manifold

A general function on a curved manifold $\mathcal{M}$ is then defined as a mapping from the neighbourhood of $\mathcal{M}$ in $\mathbf{R}^n$ to a real number $\mathbf{R}$;

$$f(x) : \{\text{neighbourhood of } \mathcal{M} \text{ in } \mathbf{R}^n\} \longrightarrow \mathbf{R} \tag{8.1}$$

where x denotes the local coordinate of the neighbourhood, $x \in \mathbf{R}^n$. Similarly we can construct vectors and higher rank tensors. Denoting a coordinate basis by dx^μ ($\mu = 1, 2, \cdots, n$), we can construct a vector as

$$A = \sum_\mu A_\mu(x) dx^\mu \tag{8.2}$$

where $A_\mu(x)$ denote components of the vector. Mathematically A is called a covariant one-form. Since this should be independent of the choice of local basis, we can require the relation $\sum_\mu A_\mu(x)dx^\mu = \sum_\nu \widetilde{A}_\nu(y)dy^\nu$, *i.e.*,

$$\widetilde{A}_\nu(y) \;=\; A_\mu(x)\frac{\partial x^\mu}{\partial y^\nu} \tag{8.3}$$

where the sum over the same indices is contracted.

We now consider a derivative of $A_\mu(x)$ with respect to x. For a function $f(x)$ we can define the derivative $\frac{\partial f}{\partial x^\mu}$ without any problems. For vectors, however, the quantity $\frac{\partial A_\mu}{\partial x^\nu}$ does not have an invariant meaning. This can easily be understood by writing the derivative as

$$\frac{\partial}{\partial y^\alpha}\widetilde{A}_\nu(y) = \left[\frac{\partial}{\partial x^\beta}A_\mu(x)\right]\frac{\partial x^\beta}{\partial y^\alpha}\frac{\partial x^\mu}{\partial y^\nu} + A_\mu(x)\frac{\partial^2 x^\mu}{\partial y^\alpha \partial y^\nu}\,. \tag{8.4}$$

The first term is tensorially well-defined while the second term causes a problem in defining the derivative as a rank-2 tensor. We can eliminate the latter term by antisymmetrizing ∂y^α and ∂y^ν:

$$\begin{aligned}\frac{\partial}{\partial y^\alpha}\widetilde{A}_\nu - \frac{\partial}{\partial y^\nu}\widetilde{A}_\alpha &= \frac{\partial}{\partial x^\beta}A_\mu\left(\frac{\partial x^\beta}{\partial y^\alpha}\frac{\partial x^\mu}{\partial y^\nu} - \frac{\partial x^\beta}{\partial y^\nu}\frac{\partial x^\mu}{\partial y^\alpha}\right) \\ &= \left[\frac{\partial}{\partial x^\beta}A_\mu - \frac{\partial}{\partial x^\mu}A_\beta\right]\frac{\partial x^\beta}{\partial y^\alpha}\frac{\partial x^\mu}{\partial y^\nu}\,.\end{aligned} \tag{8.5}$$

This means that the antisymmetrized derivative in the left-hand side transforms homogeneously. Thus, in order to properly define vectors and tensors, which are necessary to describe most of physical quantities, *we always need to deal with the antisymmetrized derivatives.* We therefore need to use antisymmetric covariant tensors and antisymmetrized derivatives. Mathematically these are equivalent to differential forms and exterior derivatives, respectively.

From a physical point of view, this rule is a natural requirement of general covariance, *i.e.*, invariance under the coordinate transformations or the diffeomorphism. As we shall discuss in Chapter 9, this is related to a weak form of the equivalence principle.

So far, we simply consider the topology of manifolds. To discuss geometry of manifolds, we need to add geometric structures to the manifolds such as metric, complex structure, Kähler structure, etc. In the next section, we first introduce the metric of a manifold.

8.2 Metric

Let x^μ $(\mu = 1, 2, \cdots, n)$ be the local coordinates of an n-dimensional manifold. Then the metric is defined by

$$ds^2 \;=\; g_{\mu\nu}(x) dx^\mu dx^\nu \tag{8.6}$$

where $g_{\mu\nu}(x)$ is the metric tensor. For the three-dimensional flat space, the metric can be written as

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \end{aligned} \tag{8.7}$$

where we use the spherical coordinates $(x^1, x^2, x^3) = (r, \theta, \phi)$ in the second line. In these coordinates, nonzero components of the metric tensor are given by $g_{11} = 1$, $g_{22} = r^2$ and $g_{33} = r^2 \sin^2\theta$.

Let y^μ be another set of local coordinates. Then, from (8.6), we find $ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu = \tilde{g}_{\mu\nu}(y) dy^\mu dy^\nu$, *i.e.*,

$$\tilde{g}_{\mu\nu}(y) \;=\; g_{\alpha\beta}(x) \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} \,. \tag{8.8}$$

The metric tensor $g_{\mu\nu}$ is therefore a symmetric rank-2 tensor.

In the n-dimensional case, $g_{\mu\nu}$ can then be expressed by a symmetric non-degenerate $n \times n$ matrix. Such a matrix is always diagonalized by orthogonal matrices and has n real eigenvalues $\lambda_i \in \mathbf{R}$ $(i = 1, 2, \cdots, n)$. For the non-degenerate matrix, the eigenvalues are also nonsingular $\lambda_i \neq 0$. Notice that we may have coordinate singularity. For example, in the case of (8.7) such a singularity appears at $r = 0$, $\theta = 0, \pi$. But this arises only for particular coordinates and non-degeneracy of $g_{\mu\nu}$ holds in general. The eigenvalues λ_i are therefore positive or negative. In many cases, we would want all of them to be positive. In special relativity, however, light propagation is invariant under Lorentz transformations and we need to require the relation

$$dt^2 - dx^2 - dy^2 - dz^2 \;=\; dt'^2 - dx'^2 - dy'^2 - dz'^2 \tag{8.9}$$

where the primed variables denote the Loretnz transformed variables. In this case, the time component counts the sign which is different from the spatial coordinates. Special relativity takes an indispensable part in most of the modern theoretical frameworks. Thus, in physics, the signs of these eigenvalues are

important and they are labeled by the *signature* of the metric, *e.g.*, $(++\cdots+)$, $(-++\cdots+)$ and so on. In high energy physics, the conventional choice is given by the Minkowski signature $(+---\cdots-)$.

In a matrix representation, the diagonalization of the metric tensor can be expressed as

$$g \;=\; S^T g_{\rm diag} S \;=\; e^T e \tag{8.10}$$

where $g_{\rm diag} = {\rm diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ and $e = {\rm diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})S$. As mentioned above, S is an orthogonal matrix $S^T = S^{-1}$. Notice that S and $g_{\rm diag}$ depend on x in general, and so dose e. In terms of matrix elements, this can be represented by

$$\begin{aligned} g_{\mu\nu} &= e^a_\mu e^a_\nu \;=\; e^a_\mu \eta^{ab} e^b_\nu & (8.11)\\ \eta^{ab} &= (+---\cdots-) & (8.12) \end{aligned}$$

where $a, b = 1, 2, \cdots, n$. e^a_μ is called the frame field. In the case of $n = 4$, this is also called the vierbein or tetrad. η^{ab} is chosen to the Minkowski signature. In terms of these, the metric (8.6) is written as

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \;=\; e^a_\mu e^a_\nu dx^\mu dx^\nu \\ &= \xi^a \xi^a \;=\; \xi^a \xi^b \eta^{ab} \end{aligned} \tag{8.13}$$

where $\xi^a = e^a_\mu dx^\mu$ is a local cartesian basis for infinitesimal length.

Frame fields, covariant derivatives and spin connections

We now consider local transformations of the frame fields

$$e^a_\mu(x) \;\longrightarrow\; e'^a_\mu(x) \;=\; R^{ab}(x) e^b_\mu(x)\,. \tag{8.14}$$

Accordingly, the metric tensor transforms as

$$g_{\mu\nu} = e^a_\mu e^a_\nu \;\longrightarrow\; g'_{\mu\nu} = R^{ab} e^b_\mu R^{ac} e^c_\nu\,. \tag{8.15}$$

Since the choice of frame fields should be independent of physical measurements, we need to require $g_{\mu\nu} = g'_{\mu\nu}$, *i.e.*,

$$(R^T R)^{bc} \;=\; \delta^{bc} \tag{8.16}$$

which means that $R^{ab}(x)$ is an orthogonal matrix. Thus (8.14) can be regarded as local rotation transformations. With this condition, we can guarantee that the frame fields do not appear in physical measurements. Alternatively, we can state that physics should not depend on the local rotation of the frame fields. Since the signature η in (8.12) is diagonal, we have

$$R^T \eta R \;=\; \eta\,. \tag{8.17}$$

This means that the Minkowski signature η is invariant under the local rotation transformations. We can then interpret that this local rotation defines the local Lorentz transformations.

In order to build physical models, one needs to define derivatives at least to the second order. We first consider the derivative of the frame field and its local Lorentz transformation. Using (8.14), we find $\partial_\mu e'^a_\nu = (\partial_\mu R^{ab}) e^b_\nu + R^{ab}\partial_\mu e^b_\nu$. Obviously, this derivative does not behave covariantly. To circumvent this problem, we now introduce the following derivative

$$D_\mu e^a_\nu = \partial_\mu e^a_\nu + \omega^{ab}_\mu e^b_\nu \,. \tag{8.18}$$

Suppressing the roman indices, this can also be written as $D_\mu e_\nu = \partial_\mu e_\nu + \omega_\mu e_\nu$. Interpreting ω_μ and e_ν as a matrix and a vector, respectively, we can easily make the roman indices explicit at any point. Similarly, $\partial_\mu e'^a_\nu$ can also be written as $\partial_\mu e'_\nu = (\partial_\mu R) e_\nu + R\partial_\mu e_\nu$. In what follows, we shall often use these notations for simplicity.

We now consider the transformations of (8.18) along with the local Lorentz transformations (8.14). These can be expressed as

$$\begin{aligned} D'_\mu e'_\nu &= \partial_\mu e'_\nu + \omega'_\mu e'_\nu \\ &= (\partial_\mu R) e_\nu + R\partial_\mu e_\nu + \omega'_\mu R e_\nu \\ &= R(\partial_\mu e_\nu + \omega_\mu e_\nu) \;=\; R(D_\mu e_\nu) \end{aligned} \tag{8.19}$$

where we chose the local transformations of ω_μ such that

$$\omega'_\mu \;=\; R\omega_\mu R^{-1} - \partial_\mu R\, R^{-1} \,. \tag{8.20}$$

Notice that $\partial_\mu R\, R^{-1}$ is an antisymmetric matrix. This can be checked as follows. From the identity $\partial_\mu(R^T R) = 0$, we find the relation ${R^{-1}}^T \partial_\mu R^T + \partial_\mu R\, R^{-1} = 0$. In terms of matrix elements, this can be written as $(R^{-1}\partial_\mu R)^{ba} + (\partial_\mu R\, R^{-1})^{ab} = 0$, leading to $(\partial_\mu R\, R^{-1})^T = -\partial_\mu R\, R^{-1}$. Thus, to make sense of the transformation (8.20), it is sufficient to consider ω^{ab}_μ being antisymmetric in a and b. Such an ω^{ab}_μ is called the *spin connection*. By use of the spin connection, the derivative (8.18) behaves covariantly as shown in (8.19). This derivative is called the *covariant derivative*. These results are derived from a conventional method of gauge theory in particle physics.

Diffeomorphism, torsion and Riemann curvature

The above analysis suggests that we should use only the covariant derivative as a derivative acting on the frame fields. Furthermore, as discussed in (8.5), we need to antisymmetrize the derivative for the diffeomorphism (or the general covariance) in relation to the coordinate transformations of the frame fields,

$$e^a_\mu(x) \;\longrightarrow\; e'^a_\mu(x') \;=\; e^a_\nu(x)\frac{\partial x^\nu}{\partial x'^\mu} \,. \tag{8.21}$$

It is then natural to define the antisymmetrized covariant derivative

$$T^a_{\mu\nu} \;=\; (D_\mu e_\nu)^a - (D_\nu e_\mu)^a \tag{8.22}$$

which transforms homogeneously under the local Lorentz transformations (8.14) and the coordinate transformations (8.21):

$$(D'_\mu e'_\nu - D'_\nu e'_\mu)^a \;=\; R^{ab}(D_\alpha e_\beta - D_\beta e_\alpha)^b \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\,. \tag{8.23}$$

The quantity $T^a_{\mu\nu}$ in (8.22) is called the *torsion tensor*.

Let ϕ^a be a scalar function and consider its local transformations $\phi^a \to \phi'^a = R^{ab}\phi^b$. The covariant derivative is given by $(D_\mu\phi)^a = \partial_\mu\phi^a + \omega^{ab}_\mu\phi^b$. From (8.20) we find $(D'_\mu\phi')^a = R^{ab}(D_\mu\phi)^b$. We now consider the second order derivative acting on ϕ^a and its local transformations. Obviously, we have $D_\mu(D_\nu\phi)^a = \partial_\mu(D_\nu\phi)^a + \omega^{ab}_\mu(D_\nu\phi)^b$. Thus $(D_\nu\phi)^a$ behaves like e^a_ν in (8.18) and, by construction, we find

$$(D'_\mu D'_\nu\phi')^a \;=\; R^{ab}(D_\mu D_\nu\phi)^b\,. \tag{8.24}$$

Therefore, as in the case of torsion, the antisymmetrized combination $(D_\mu D_\nu\phi - D_\nu D_\mu\phi)^a$ transforms homogenously. Separating out the symmetric part of $D_\mu D_\nu\phi^a$, we can explicitly write down the combination as

$$\begin{aligned}(D_\mu D_\nu\phi - D_\nu D_\mu\phi)^a &= \mathcal{R}^{ab}_{\mu\nu}\phi^b\,, &(8.25)\\ \mathcal{R}^{ab}_{\mu\nu} &= \partial_\mu\omega^{ab}_\nu - \partial_\nu\omega^{ab}_\mu + \omega^{ac}_\mu\omega^{cb}_\nu - \omega^{ac}_\nu\omega^{cb}_\mu\,. &(8.26)\end{aligned}$$

The transformations of (8.25) lead to the relation $R^{ab}\mathcal{R}^{bc}_{\mu\nu}\phi^c = \mathcal{R}'^{ab}_{\mu\nu}\phi'^b$ where $\phi'^b = R^{bc}\phi^c$, *i.e.*,

$$\mathcal{R}'^{ad}_{\mu\nu} \;=\; R^{ab}\mathcal{R}^{bc}_{\mu\nu}(R^{-1})^{cd}\,. \tag{8.27}$$

Thus $\mathcal{R}^{ab}_{\mu\nu}$ is a nice tensorial object whose local transformations are given by (8.27). This tensor is called the *Riemann curvature tensor*.

Notice that both the torsion $T^a_{\mu\nu}$ and the Riemann curvature $\mathcal{R}^{ab}_{\mu\nu}$ are derived from the frame fields e^a_μ and the spin connection ω^{ab}_μ or, more fundamentally, from the metric $g_{\mu\nu}$ that is defined in a general curved manifold. From its definition (8.22), we find that $T^a_{\mu\nu}$ detects the effects of frame rotation or Lorentz transformation. In physical measurements, we observe no evidence of nonzero torsion. The torsion-free condition is thus physically important. In fact, as we shall see in a moment, this condition leads the general curved manifold to be a Riemannian manifold and provides a crucial foundation for the construction of Einstein's theory of general relativity. In this respect, the Riemann curvature is more relevant to physics than the torsion; it determines the local shape of manifolds and is responsible for physical quantities such as tidal forces due to gravity.

8.3 Riemannian manifolds

As mentioned above a Riemannian manifold is defined by imposing the torsion-free condition

$$T^a_{\mu\nu} \,=\, (D_\mu e_\nu)^a - (D_\nu e_\mu)^a \,=\, 0 \tag{8.28}$$

to the general curved manifold. In this section, we consider the consequences of this condition and reveal characteristic features of the Riemannian manifold. (Note that some theories such as supergravity and string theory may contain non-vanishing torsion.)

From (8.28) we can easily find that the $(D_\mu e_\nu)^a$ is symmetric in μ and ν. Let us denote this quantity by $\Gamma^a_{\mu\nu}$. Using the fact that the frame field e^a_μ is invertible, we can then express $\Gamma^a_{\mu\nu}$ as $\Gamma^\lambda_{\mu\nu} e^a_\lambda$, *i.e.*,

$$(D_\mu e_\nu)^a \,=\, \partial_\mu e^a_\nu + \omega^{ab}_\mu e^b_\nu \,\equiv\, \Gamma^\lambda_{\mu\nu} e^a_\lambda \tag{8.29}$$

where $\Gamma^\lambda_{\mu\nu}$ is symmetric in μ and ν. We now see how this affects the metric by calculating the derivative of $g_{\alpha\beta}$ as follows

$$\begin{aligned}
\partial_\mu g_{\alpha\beta} &= (\partial_\mu e^a_\alpha) e^a_\beta + e^a_\alpha (\partial_\mu e^a_\beta) \\
&= (\partial_\mu e^a_\alpha + \omega^{ab}_\mu e^b_\alpha) e^a_\beta + e^a_\alpha (\partial_\mu e^a_\beta + \omega^{ab}_\mu e^b_\beta) - \underbrace{\omega^{ab}_\mu (e^b_\alpha e^a_\beta + e^a_\alpha e^b_\beta)}_{=0} \\
&= \Gamma^\lambda_{\mu\alpha} e^a_\lambda e^a_\beta + \Gamma^\lambda_{\mu\beta} e^a_\alpha e^a_\lambda \\
&= \Gamma^\lambda_{\mu\alpha}\, g_{\lambda\beta} + \Gamma^\lambda_{\mu\beta}\, g_{\alpha\lambda}
\end{aligned} \tag{8.30}$$

where we use the fact that ω^{ab}_μ is antysimmetric in a and b. Utilizing symmetric properties of indices in (8.30), we can obtain the relation

$$\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\alpha\mu} - \partial_\mu g_{\alpha\beta} \,=\, 2\Gamma^\lambda_{\alpha\beta}\, g_{\mu\lambda}\,. \tag{8.31}$$

We now denote the inverse of the metric with upper indices, $g_{\alpha\beta} g^{\beta\lambda} = \delta^\lambda_\alpha$. Namely, in this notation, we have $(g^{-1})^{\beta\lambda} = g^{\beta\lambda}$. Then, in terms of the metric, $\Gamma^\lambda_{\alpha\beta}$ can be expressed as

$$\Gamma^\lambda_{\alpha\beta} \,=\, \frac{1}{2} g^{\lambda\mu} \left(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\alpha\mu} - \partial_\mu g_{\alpha\beta}\right)\,. \tag{8.32}$$

The quantity $\Gamma^\lambda_{\alpha\beta}$ is called the *Christoffel symbol*. Let us denote local coordinate transformations of (8.29) as $\widetilde{D}_\mu \widetilde{e}_\nu = \widetilde{\Gamma}^\lambda_{\mu\nu} \widetilde{e}_\lambda$. Using (8.3) and (8.4), the right-hand-side can be calculated as

$$\begin{aligned}
\widetilde{D}_\mu \widetilde{e}_\nu &= \left(\frac{\partial}{\partial y^\mu} + \omega_\alpha(x) \frac{\partial x^\alpha}{\partial y^\mu}\right) \left[e_\beta(x) \frac{\partial x^\beta}{\partial y^\nu}\right] \\
&= (D_\alpha e_\beta) \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} + e_\beta \frac{\partial^2 x^\beta}{\partial y^\mu \partial y^\nu}
\end{aligned}$$

$$= \left(\Gamma^\lambda_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} + \frac{\partial^2 x^\lambda}{\partial y^\mu \partial y^\nu} \right) e_\lambda \,. \tag{8.33}$$

The local coordinate transformations of $\Gamma^\lambda_{\mu\nu}$ is then expressed as

$$\widetilde{\Gamma}^\lambda_{\mu\nu} = \Gamma^\sigma_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} \frac{\partial y^\lambda}{\partial x^\sigma} + \frac{\partial^2 x^\lambda}{\partial y^\mu \partial y^\nu} \frac{\partial y^\lambda}{\partial x^\sigma} \,. \tag{8.34}$$

This explicitly shows that the Christoffel symbol does not behave as a tensor.

We can also express the spin connection ω^{ab}_μ in terms of e^a_μ and $\Gamma^\lambda_{\mu\nu}$. From (8.29), we find

$$\partial_\mu e^a_\nu (e^{-1})^{b\nu} + \omega^{ab}_\mu = \Gamma^\lambda_{\mu\nu} e^a_\lambda (e^{-1})^{b\nu} \tag{8.35}$$

where we denote the inverse of e^b_ν as $(e^{-1})^{b\nu}$ with upper indices. We now consider $\Gamma^\lambda_{\mu\nu} = (\Gamma_\mu)^\lambda_\nu$ as a matrix element of (Γ_μ). Similarly, we consider $e^a_\lambda = (e^a)_\lambda$ and $(e^{-1})^{b\nu}$ as elements of the column vector (e^a) and the row vector $(e^{-1})^b$, respectively. The spin connection is then expressed as

$$\begin{aligned} \omega^{ab}_\mu &= e^a_\lambda \Gamma^\lambda_{\mu\nu} (e^{-1})^{b\nu} - \partial_\mu e^a_\nu (e^{-1})^{b\nu} \\ &= e^a \Gamma_\mu (e^{-1})^b - \partial_\mu e^a (e^{-1})^b \end{aligned} \tag{8.36}$$

where we use the matrix notation in the second line. Notice that (8.36) is analogous to the general result for the curved manifolds $\omega'_\mu = R\omega_\mu R^{-1} - \partial_\mu R R^{-1}$ in (8.20). The expression (8.36) shows that on the Riemannian manifolds (with $T^a_{\mu\nu} = 0$) the spin connection is completely determined by the metric $g_{\mu\nu} = e^a_\mu e^a_\nu$.

We can similarly write the Riemann curvature tensor (8.26). In the matrix notation, we have

$$\mathcal{R}_{\mu\nu} \,=\, \partial_\mu \omega_\nu - \partial_\nu \omega_\mu + [\omega_\mu, \omega_\nu] \,. \tag{8.37}$$

We now write the spin connection in a general form

$$\omega_\mu \,=\, M \Omega_\mu M^{-1} - \partial_\mu M \, M^{-1} \tag{8.38}$$

where M and Ω_μ are arbitrary $n \times n$ matrices, with M^{-1} being the inverse of M. Substituting (8.38) into $\partial_\mu \omega_\nu - \partial_\nu \omega_\mu$, we find

$$\begin{aligned} \partial_\mu \omega_\nu - \partial_\nu \omega_\mu &= M(\partial_\mu \Omega_\nu - \partial_\nu \Omega_\mu) M^{-1} + [\partial_\mu M \, M^{-1}, M \Omega_\nu M^{-1}] \\ &\quad + [M \Omega_\mu M^{-1}, \partial_\nu M \, M^{-1}] - [\partial_\mu M \, M^{-1}, \partial_\nu M \, M^{-1}] \end{aligned} \tag{8.39}$$

where we use $\partial_\mu M = -M^{-1} \partial_\mu M M^{-1}$. The commutator $[\omega_\mu, \omega_\nu]$ is expanded as

$$\begin{aligned} [\omega_\mu, \omega_\nu] &= M[\omega_\mu, \omega_\nu] M^{-1} - [M \Omega_\mu M^{-1}, \partial_\nu M \, M^{-1}] \\ &\quad - [\partial_\mu M \, M^{-1}, M \Omega_\nu M^{-1}] + [\partial_\mu M \, M^{-1}, \partial_\nu M \, M^{-1}] \,. \end{aligned} \tag{8.40}$$

Thus, for the general spin connection (8.38), the Riemann curvature (8.37) is expressed as

$$\mathcal{R}_{\mu\nu} \,=\, M \left(\partial_\mu \Omega_\nu - \partial_\nu \Omega_\mu + [\Omega_\mu, \Omega_\nu] \right) M^{-1} \,. \tag{8.41}$$

By construction, this relation holds in general, not necessarily for the Riemannian manifold. For the choice of $(M, \Omega_\mu) = (R, \omega_\mu)$, this relation proves the Lorentz covariance of the Riemann curvature $\mathcal{R}'^{ad}_{\mu\nu} = R^{ab}\mathcal{R}^{bc}_{\mu\nu}(R^{-1})^{cd}$ in (8.27) for curved manifolds in general where R denotes the local rotational matrix as before.

From the expression (8.36) we find that the choice of $(M, \Omega_\mu) = (e, \Gamma_\mu)$ corresponds to the Riemannian manifold. Thus, directly from (8.41), we find that the Riemann curvature on the Riemannian manifold is expressed as

$$\mathcal{R}_{\mu\nu} = e\left(\partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu]\right) e^{-1} \tag{8.42}$$

in the matrix notation. The rest of the indices can be recovered automatically by explicitly writing down the matrix elements:

$$\mathcal{R}^{ab}_{\mu\nu} = e^a_\lambda \, \mathcal{R}^\lambda_{\mu\nu\alpha} \, (e^{-1})^{b\alpha} , \tag{8.43}$$

$$\mathcal{R}^\lambda_{\mu\nu\alpha} = \partial_\mu \Gamma^\lambda_{\nu\alpha} - \partial_\nu \Gamma^\lambda_{\mu\alpha} + \Gamma^\lambda_{\mu\beta}\Gamma^\beta_{\nu\alpha} - \Gamma^\lambda_{\nu\beta}\Gamma^\beta_{\mu\alpha} . \tag{8.44}$$

The above derivation shows the matrix parametrization of ω_μ in (8.38) is powerful and useful. Once we obtain matrix expressions such as the Riemann curvature tensor, we can correctly recover the full tensorial indices in the notation of matrix elements.

Summary

In this section, we consider the general curved manifold and its specialization. We first begin with a differentiable topological manifold. We then add the geometric information to it by defining the metric. From the symmetric and non-degenerate properties of the metric tensor $g_{\mu\nu}$, we can represent it as a product of the frame fields e^a_μ. The spin connection ω^{ab}_μ emerges from the definition of the covariant derivative D_μ acting on the frame fields. The torsion tensor $T^a_{\mu\nu}$ and the Riemann curvature tensor $R^{ab}_{\mu\nu}$ are then obtained from $(e^a_\mu, \omega^{ab}_\mu)$.

$$g_{\mu\nu} = e^a_\mu e^a_\nu \tag{8.45}$$

$$(D_\mu e_\nu)^a = \partial_\mu e^a_\nu + \omega^{ab}_\mu e^b_\nu \tag{8.46}$$

$$(D_\mu e_\nu - D_\nu e_\mu)^a = T^a_{\mu\nu} \tag{8.47}$$

$$[(D_\mu D_\nu - D_\nu D_\mu)\phi]^a = R^{ab}_{\mu\nu}\phi^b \tag{8.48}$$

The Riemannian manifold is obtained by the imposition of the torsion-free condition $T^a_{\mu\nu} = 0$ on top of these relations. We can then solve for the spin connection $\omega^{ab}_{\mu\nu}$ in terms of derivatives of the metric. This can be carried out efficiently by introducing the Christoffel symbol $\Gamma^\lambda_{\mu\nu}$. The Riemann curvature can also be written in terms of $(e^a_\mu, \Gamma^\lambda_{\mu\nu})$.

$$(D_\mu e_\nu)^a = \Gamma^\lambda_{\mu\nu} e^a_\lambda \quad \longleftarrow \quad T^a_{\mu\nu} = 0 \ \ \text{(symmetric in } \mu, \nu) \tag{8.49}$$

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha} \left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu} \right) \tag{8.50}$$

$$\omega^{ab}_{\mu} = e^{a}_{\lambda} \Gamma^{\lambda}_{\mu\nu} (e^{-1})^{\nu b} - \partial_\mu e^{a}_{\lambda} \, (e^{-1})^{\lambda b} \tag{8.51}$$

$$\mathcal{R}^{ab}_{\mu\nu} = e^{a}_{\lambda} \mathcal{R}^{\lambda}_{\mu\nu\alpha} (e^{-1})^{\alpha b} \tag{8.52}$$

$$\mathcal{R}^{\lambda}_{\mu\nu\alpha} = \partial_\mu \Gamma^{\lambda}_{\nu\alpha} - \partial_\nu \Gamma^{\lambda}_{\mu\alpha} + \Gamma^{\lambda}_{\mu\beta} \Gamma^{\beta}_{\nu\alpha} - \Gamma^{\lambda}_{\nu\beta} \Gamma^{\beta}_{\mu\alpha} \tag{8.53}$$

All physical quantities are derived from the metric tensor $g_{\mu\nu}$. In this sense, physics on the curved manifold which includes the Riemannian manifold can be described by a “metric theory”.

Chapter 9

Modern derivation of Einstein equations

In physics, the Riemannian manifold is very important. In fact, we can postulate the followings.

1. The world is described by a Riemannian manifold with signature $(+---)$.
2. The metric $g_{\mu\nu}$ is determined dynamically by the matter content.

The latter assertion leads to the equations of motion for $g_{\mu\nu}$, the Einstein equations. One of the main goals of this chapter is to derive the Einstein equations by use of Riemannian geometry and desired symmetries.

These postulates are closely related to the equivalence principle. Depending on the level of equivalence, there are historically two versions of the principle. These are stated as below:

1. The inertial mass is the same as the gravitational mass (weak statement).
2. Coupling to gravity can be implemented by replacing ordinary derivatives with covariant derivatives (strong statement).

The strong statement means that an accelerating frame is equivalent to a gravitational field. This is true for point particles with no spins. For particles with spins, the strong statement is not generally true.

9.1 Motion of point particles

To illustrate the usefulness of symmetry, in this section we obtain the equations of motion for point particles in the Lagrangian formalism. The point particle is

the simplest example of the matter content, *i.e.*, things except gravity such as electromagnetic fields, mesons, baryons, fluids and gases.

As discussed in (8.13), the metric is given by

$$ds^2 \;=\; g_{\mu\nu}dx^\mu dx^\nu \;=\; dt^2 - dx^2 - dy^2 - dz^2 \tag{9.1}$$

for the flat Minkowski space. By definition this quantity is Lorentz invariant. We can then specify a relativistically invariant action by

$$\mathcal{S} \;=\; -m\int\sqrt{ds^2} \;=\; -m\int\sqrt{g_{\mu\nu}dx^\mu dx^\nu}\,. \tag{9.2}$$

The coefficient $-m$ can be understood if we consider the non-relativistic limit:

$$\begin{aligned} S &= -m\int dt\sqrt{1-v^2} \\ &\approx \int dt\left(-m+\frac{mv^2}{2}\right) \end{aligned} \tag{9.3}$$

for small v where v denotes the speed of the particle in the unit of $c = 1$. The action (9.2) depends only on $g_{\mu\nu}$. This is a consequence of the Lorentz invariance imposed on the action. Since the Riemann curvature tensor $\mathcal{R}^{ab}_{\mu\nu}$ is also Lorentz covariant, we can in principle construct Lorentz invariant terms out of $\mathcal{R}^{ab}_{\mu\nu}$. For example, as $\mathcal{R}^{ab}_{\mu\nu}$ has spin information, we can add terms like $\mathcal{R}^{ab}_{\mu\nu}S_{\mu\nu}$ for particle with spins where $S_{\mu\nu}$ denotes the spin contribution to the Lorentz generators. Such an additional term, however, violates the strong statement of the equivalence principle. For spin-0 particles, on the other hand, the strong statement is valid and we can use the action (9.2) to analyze the motion of point particles. Notice that this action is interpreted as an integral of curved distance. An interesting modification is then given by an action of an integral over surface area. This is known as the Nambu-Goto action and leads to the foundation of string theory.

Introducing a time-like parameter λ, we can express the action (9.2) as

$$\mathcal{S} \;=\; -m\int d\lambda\;\mathcal{L}\,, \tag{9.4}$$

$$\mathcal{L} \;=\; \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\,. \tag{9.5}$$

The Euler-Lagrange equation for $\mathcal{L}$ becomes

$$\frac{\partial\mathcal{L}}{\partial x^\mu} - \frac{d}{d\lambda}\frac{\partial\mathcal{L}}{\partial\left(\frac{dx^\mu}{d\lambda}\right)} \;=\; 0\,. \tag{9.6}$$

The first term on the left-hand side can be simplified as

$$\frac{\partial\mathcal{L}}{\partial x^\mu} \;=\; \frac{1}{\sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}}\frac{1}{2}(\partial_\mu g_{\beta\tau})\frac{dx^\beta}{d\lambda}\frac{dx^\tau}{d\lambda}$$

$$= \frac{1}{2}\partial_\mu g_{\beta\tau}\frac{dx^\beta}{d\lambda}\frac{dx^\tau}{ds} = \left(\frac{1}{2}\partial_\mu g_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}\right)\frac{ds}{d\lambda} \tag{9.7}$$

where we use

$$\sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} = \frac{ds}{d\lambda}\,. \tag{9.8}$$

Similarly, we can calculate the second term as

$$\begin{aligned}\frac{d}{d\lambda}\frac{\partial\mathcal{L}}{\partial\left(\frac{dx^\mu}{d\lambda}\right)} &= \frac{d}{d\lambda}\left(\frac{g_{\mu\beta}\frac{dx^\beta}{d\lambda}}{\sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}}\right)\\ &= \frac{d}{d\lambda}\left(g_{\mu\beta}\frac{dx^\beta}{ds}\right) = \partial_\alpha g_{\mu\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{ds} + g_{\mu\beta}\frac{d^2x^\beta}{ds^2}\frac{ds}{d\lambda}\\ &= \left(\partial_\alpha g_{\mu\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} + g_{\mu\beta}\frac{d^2x^\beta}{ds^2}\right)\frac{ds}{d\lambda}\,.\end{aligned} \tag{9.9}$$

Putting together (9.7) and (9.9), the Euler-Lagrange (9.6) becomes

$$g_{\mu\beta}\frac{d^2x^\beta}{ds^2} + \left(\partial_\alpha g_{\mu\beta} - \frac{1}{2}\partial_\mu g_{\alpha\beta}\right)\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0\,. \tag{9.10}$$

Multiplying $g^{\lambda\mu}$, the inverse of the metric $g_{\lambda\mu}$, this equation can be simplified as

$$\begin{aligned}&\frac{d^2x^\lambda}{ds^2} + g^{\lambda\mu}\left(\partial_\alpha g_{\mu\beta} - \frac{1}{2}\partial_\mu g_{\alpha\beta}\right)\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}\\ =\ &\frac{d^2x^\lambda}{ds^2} + \frac{1}{2}g^{\lambda\mu}\left(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta}\right)\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}\\ =\ &0\end{aligned} \tag{9.11}$$

where we use that the indices (α, β) are symmetric, and so are (λ, μ). From the definition of the Christoffel symbol in (8.32), we then find that the Euler-Lagrange equation eventually reduces to the equation

$$\frac{d^2x^\lambda}{ds^2} + \Gamma^\lambda_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0\,. \tag{9.12}$$

This is called the geodesic equation. The Lagrangian (9.5) that we start with is given by the proper distance. Thus the geodesic equation describes a *straight line* which minimizes the proper distance. Notice that the derivation of (9.12) is independent of the choice of the metric. Thus such a straight line represents the propagation of light (or the geodesic) on a curved spacetime in general.

9.2 Modern derivation of Einstein equations

In the previous section, we see that the equations of motion for point particles can be derived from the invariance under the Lorentz transformations. In this

section, we consider the equations of motion for the metric, or the Einstein equations, in a similar fashion. As mentioned earlier, the metric is to be determined dynamically by the matter content. The resultant equations should serve as a mathematical realization of this assertion and provide a criterion to fix a metric in a given matter distribution. Most of the dynamics of anything is determined by symmetries. Thus our strategy to derive the Einstein equations is first to write down an action $\mathcal{S}$ such that it is invariant under all desired symmetries. In a general Riemannian manifold, such an action can be written as

$$\mathcal{S} \,=\, \int d^4x\sqrt{-\det g}\,\Big(\text{invariant terms}\Big) \tag{9.13}$$

where the integral measure $\int d^4x\sqrt{-\det g}$ is invariant under the coordinate transformations; one can check that the factor $\sqrt{-\det g}$ cancels the Jacobian arising from the coordinate transformations in the Minkowski space.

Apart from a constant factor, the invariant factor in (9.13) is expected to be written in terms of the Riemann curvature tensor $\mathcal{R}^{\lambda}_{\mu\nu\alpha}$. As defined in (8.44), it is given by

$$\mathcal{R}^{\lambda}_{\mu\nu\alpha} \,=\, \partial_\mu\Gamma^{\lambda}_{\nu\alpha} - \partial_\nu\Gamma^{\lambda}_{\mu\alpha} + \Gamma^{\lambda}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} - \Gamma^{\lambda}_{\nu\beta}\Gamma^{\beta}_{\mu\alpha} \tag{9.14}$$

where the Christoffel symbol is defined by

$$\Gamma^{\lambda}_{\alpha\beta} \,=\, \frac{1}{2}g^{\lambda\mu}\left(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta}\right) . \tag{9.15}$$

In order to obtain an invariant quantity out of $\mathcal{R}^{\lambda}_{\mu\nu\alpha}$, it is convenient to consider the contraction of indices:

$$\begin{aligned} \mathcal{R}^{\lambda}_{\mu\nu\alpha} \;\longrightarrow\; \mathcal{R}^{\lambda}_{\lambda\nu\alpha} \;&\equiv\; \mathcal{R}_{\nu\alpha}\,, \\ \mathcal{R}_{\nu\alpha} \;&=\; \partial_\lambda\Gamma^{\lambda}_{\nu\alpha} - \partial_\nu\Gamma^{\lambda}_{\lambda\alpha} + \Gamma^{\lambda}_{\lambda\beta}\Gamma^{\beta}_{\nu\alpha} - \Gamma^{\lambda}_{\nu\beta}\Gamma^{\beta}_{\lambda\alpha}\,. \end{aligned} \tag{9.16}$$

$\mathcal{R}_{\nu\alpha}$ is called the *Ricci tensor*. Notice that $\Gamma^{\lambda}_{\nu\alpha}$ is symmetric in ν and α. Thus, from (9.15) and (9.16), one can easily check that $\mathcal{R}_{\nu\alpha}$ is symmetric in ν and α as well. Since the metric tensor $g_{\mu\nu}$ is also symmetric in the indices, we can introduce a nontrivial scalar

$$\mathcal{R} \,=\, g^{\nu\alpha}\mathcal{R}_{\nu\alpha}\,. \tag{9.17}$$

This is called the *Ricci scalar*. For a spherical manifold with radius r, the Ricci scalar is proportional to r^{-2}.

From the symmetry principle, the invariant action (9.13) can tentatively be written as

$$\begin{aligned} \mathcal{S} \;&=\; \int d^4x\sqrt{-\det g}\,\Big[(\text{const}) + \#\mathcal{R} + \big(\mathcal{R}^2,\, \mathcal{R}_{\nu\alpha}\mathcal{R}^{\nu\alpha},\, \mathcal{R}^{\lambda}_{\mu\nu\alpha}\mathcal{R}^{\mu\nu\alpha}_{\lambda},\, \cdots\big)\Big] \\ &=\; \int d^4x\sqrt{-\det g}\,\left[\frac{1}{16\pi G}\mathcal{R} - \Lambda\right] + \cdots \end{aligned} \tag{9.18}$$

where # is a certain constant. For later convenience we fix it to $\frac{1}{16\pi G}$, with G being Newton's constant. Λ is also a constant and is called the *cosmological constant*. We shall discuss its physical meaning later in this section.

We now consider variation of the action (9.18). The variation of the Ricci scalar $\mathcal{R} = \mathcal{R}_{\alpha\beta} g^{\alpha\beta}$ is given by

$$\delta\mathcal{R} \;=\; \delta\mathcal{R}_{\alpha\beta}\, g^{\alpha\beta} \;+\; \mathcal{R}_{\alpha\beta}\delta g^{\alpha\beta}\,. \tag{9.19}$$

The variation of $\sqrt{-\det g}$ can be obtained as follows. Let M be an invertible matrix. Using the identity

$$\log\det M \;=\; \mathrm{Tr}\log M\,, \tag{9.20}$$

we find the relation

$$\delta(\log\det M) \;=\; \frac{\delta(\det M)}{\det M} \;=\; \mathrm{Tr}M^{-1}\delta M \;=\; -\mathrm{Tr}M\delta M^{-1}\,. \tag{9.21}$$

Thus we can easily calculate the variation of $\sqrt{-\det g}$ as

$$\delta\sqrt{-\det g} \;=\; -\frac{1}{2}\sqrt{-\det g}\, g_{\alpha\beta}\delta g^{\alpha\beta} \tag{9.22}$$

From (9.18), (9.19) and (9.22), we find

$$\begin{aligned}\delta\mathcal{S} \;\;=\;\; & \int d^4x\sqrt{-\det g}\left[-\frac{1}{2}g_{\alpha\beta}\left(\frac{1}{16\pi G}\mathcal{R}-\Lambda\right)+\frac{1}{16\pi G}\mathcal{R}_{\alpha\beta}\right]\delta g^{\alpha\beta} \\ & + \int d^4x\sqrt{-\det g}\,\frac{1}{16\pi G}\delta\mathcal{R}_{\alpha\beta}g^{\alpha\beta}\,.\end{aligned} \tag{9.23}$$

The integrand of the last term is a total divergence. Thus, by use of the variational principle, we obtain the following equation of motion:

$$\mathcal{R}_{\alpha\beta} \;-\; \frac{1}{2}g_{\alpha\beta}\mathcal{R} \;+\; 8\pi G\Lambda\, g_{\alpha\beta} \;=\; 0\,. \tag{9.24}$$

This is called the vacuum Einstein equation.

We now consider the incorporation of the matter source into the free field action (9.18). Let $\mathcal{S}_{\mathrm{matter}}$ be the matter action

$$\mathcal{S}_{\mathrm{matter}} \;=\; \int d^4x\sqrt{-\det g}\,\mathcal{L}_{\mathrm{matter}}\,. \tag{9.25}$$

By definition, the matter action satisfies the equation

$$\frac{\delta\mathcal{S}_{\mathrm{matter}}}{\delta g^{\mu\nu}} \;=\; \frac{1}{2}T_{\mu\nu} \tag{9.26}$$

where $T_{\mu\nu}$ is the energy-momentum tensor which is symmetric in μ and ν. This equation can be rewritten as

$$\delta\mathcal{S}_{\mathrm{matter}} \;=\; \int d^4x\sqrt{-\det g}\,\left(\frac{1}{2}T_{\mu\nu}\delta g^{\mu\nu}\right)\,. \tag{9.27}$$

Combined with (9.23), we then obtain the full Einstein equation

$$\mathcal{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta}\mathcal{R} + 8\pi G\Lambda\, g_{\alpha\beta} = -8\pi G\, T_{\mu\nu}\,. \tag{9.28}$$

The identification of G with Newton's constant can be explained as follows. Consider a point particle of mass m under the gravitational potential $\Phi = -\frac{GM}{r}$. As discussed in (9.3), the motion of a point particle is given by the action

$$\begin{aligned} \mathcal{S} &= -m\int \sqrt{g_{\mu\nu}dx^\mu dx^\nu} \\ &= -m\int \sqrt{g_{00}-v^2}dt \\ &\approx -m + \frac{1}{2}mv^2 - m\Phi \end{aligned} \tag{9.29}$$

where $g_{ij} = -\delta_{ij}$ and the $(0,0)$-component of the metric approximates to

$$g_{00} \approx 1 + 2\Phi = 1 - \frac{2GM}{r} \tag{9.30}$$

in the non-relativistic limit. Linearlizing the Einstein equation, with the term 2Φ regarded as perturbative, we find that the Einstein equation reduces to Poission's equation for the gravitational potential:

$$\nabla^2\Phi = 4\pi G\rho \tag{9.31}$$

where ρ is the mass density of the matter source, $T_{00} \approx \rho = M\delta^{(0)}(x)$. From this linearlization, we can then reduce the Einstein equation to the classical Newton theory by identifying G as Newtons's constant.

On the other hand, the cosmological constant Λ does not have analogs to the Newton theory. From the definition (9.18), we find that Λ serves as the energy density of the vacuum. According to observational data, the universe is expanding with a positive acceleration. In terms of Λ, this fact means that $\Lambda > 0$. Theoretically, however, we cannot determine the sign of Λ. There are possibilities of $\Lambda = 0$ (constant expansion of universe) and of $\Lambda < 0$ (shrinking universe).

Lastly, we consider all the remaining terms in the action (9.18) such as $\mathcal{R}^2$ and $\mathcal{R}_{\nu\alpha}\mathcal{R}^{\nu\alpha}$:

$$\mathcal{S} = \int d^4x\sqrt{-\det g}\,\frac{1}{16\pi G}\Big[\mathcal{R} - 16\pi G\Lambda + 16\pi G\big(c_1\mathcal{R}^2 + c_2\mathcal{R}_{\nu\alpha}\mathcal{R}^{\nu\alpha} + \cdots\big)\Big] + \mathcal{S}_{\text{matter}} \tag{9.32}$$

where c_1, c_2 are constants. In the non-relativistic limit (9.30), we can calculate $\mathcal{R}$ as $\mathcal{R} = \frac{GM}{r^3}$. This can be derived from the definition of the Ricci scalar; note that $\mathcal{R}$ is roughly given by the second order derivative of the metric, $\mathcal{R} \sim \nabla^2 g_{00}$. We can then evaluate the factor $G\mathcal{R}$ as

$$G\mathcal{R} = G\frac{GM}{r^3} \sim G^2(\text{density}) \ll 1 \tag{9.33}$$

for any element in the periodic table. The largest density of all elements is given by the osmium (Os). Even for the osmium, the value can be evaluated as $G^2(\text{density})_{\text{Os}} \leq 10^{-90}$ in a suitable unit. Thus in most of the cases the higher derivative terms such as $\mathcal{R}^2$ and $\mathcal{R}_{\nu\alpha}\mathcal{R}^{\nu\alpha}$ are negligible in the study of general relativity. At high-energy and short-distance limit, these derivative terms may become important but at low-energy and long-distance limit we can minimize the number of derivatives in obtaining an effective action.

9.3 Schwarzschild metric

For a point matter source $\rho = M\delta^{(3)}(x)$ (*e.g.*, a matter distribution around the sun) a solution to the Einstein equation (9.28) is given by the Schwarzschild metric

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2GM}{r}} - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2 \,. \tag{9.34}$$

Using this metric, we can specify an action for the motion of a planet:

$$\begin{aligned}
\mathcal{S} &= -m \int dt \sqrt{\left(1 - \frac{2GM}{r}\right) - \frac{\dot{r}^2}{1 - \frac{2GM}{r}} - r^2\dot{\theta}^2 - r^2 \sin^2\theta \dot{\phi}^2} \\
&\approx -m \int dt \left[\sqrt{1 - \frac{2GM}{r}} - \frac{1}{2} \frac{1}{1 - \frac{2GM}{r}} \left(\frac{\dot{r}^2}{1 - \frac{2GM}{r}} + r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2\right)\right] \\
&\approx -m \int dt \left[\left(1 - \frac{GM}{r}\right) - \frac{1}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2\right)\right]
\end{aligned} \tag{9.35}$$

where we use $\frac{2GM}{r} \ll 1$. The action is then written as

$$S = \int dt \left[-m + \frac{1}{2}mv^2 - \left(-\frac{GMm}{r}\right)\right] + \cdots \tag{9.36}$$

where $v = \dot{r}$ in the unit of $c = 1$. The terms in the square brackets denote the rest mass, the kinetic energy and the gravitational potential, respectively. These are in accord with the Newtonian theory except the explicit rest mass term. The terms denoted by the ellipsis denote calculable correction to the Newtonian theory. For example, the perihelion precession of Mercury can be computed from these terms.

The Schwarzschild metric has a singularity at $r_{\text{sc}} = 2GM$. This is called the Schwarzschild radius. As discussed earlier, the propagation of light is described by the geodesic equation (9.12). Applying the Schwarzschild metric, we observe that there are no geodesics that start inside $r < r_{\text{sc}}$ and end outside $r > r_{\text{sc}}$. The region inside $r < r_{\text{sc}}$ is called a black hole.

There is an important theorem, called Birkhoff's theorem, regarding the Schwarzschild metric. The theorem states that

> The only non-rotating and uncharged solution of the Einstein equation (with $\Lambda = 0$) with spherical symmetry is the Schwarzschild solution.

If rotating solutions are further allowed, the solution of the Einstein equation is uniquely given by the so-called Kerr solution (found by Roy Kerr around 1964). This is known as the black hole uniqueness theorem.

In terms of the frame field e^a_μ, the Schwarzschild metric is described by

$$\begin{aligned} e^0_0 &= \sqrt{1 - \frac{2GM}{r}}\,, \qquad & e^2_2 &= r\,, \\ e^1_1 &= \frac{1}{\sqrt{1 - \frac{2GM}{r}}}\,, \qquad & e^3_3 &= r\sin\theta \end{aligned} \tag{9.37}$$

where the set of indices $(0, 1, 2, 3)$ correspond to the polar coordinates (t, r, θ, ϕ). The remaining components of the frame field are all zero. It is a good exercise to check that this is indeed a solution of the Einstein equation. (The best approach would be to find the spin connection ω^{ab}_μ first and use it to obtain the Ricci tensor $\mathcal{R}_{\mu\nu}$ and the Ricci scalar $\mathcal{R}$.)

Chapter 10

Isometries and cosmological solutions

10.1 Covariant derivatives, isometries and Killing equations

Covariant derivatives on the Riemannian manifold

Let ϕ^a be a scalar function. As defined in (8.24), a general covariant derivative acting on ϕ^a is expressed as

$$(D_\mu \phi)^a \;=\; \partial_\mu \phi^a + \omega_\mu^{ab} \phi^b \,. \tag{10.1}$$

where ω_μ^{ab} is the spin connection. We have obtained this expression by requiring the covariance of the derivative $(D'_\mu \phi')^a = L^{ab}(D_\mu \phi)^b$ under the local Lorentz transformation of ϕ^a, $\phi^a \to \phi'^a = L^{ab}\phi^b$; see (8.18), (8.19) for details.

We now consider the expansion of ϕ^a by the frame fields:

$$\phi^a \;=\; e_\alpha^a \phi^\alpha \,. \tag{10.2}$$

Using this notation, we can rewrite (10.1) as

$$\begin{aligned} (D_\mu \phi)^a &= (\partial_\mu e_\alpha^a + \omega_\mu^{ab} e_\alpha^b)\phi^\alpha + e_\alpha^a \partial_\mu \phi^\alpha \\ &= e_\lambda^a \left(\partial_\mu \phi^\lambda + \Gamma_{\mu\alpha}^{la} \phi^\alpha\right) \end{aligned} \tag{10.3}$$

where we use the basic relation of the Riemannian manifold $\partial_\mu e_\alpha^a + \omega_\mu^{ab} e_\alpha^b = \Gamma_{\mu\alpha}^\lambda e_\lambda^a$; see (8.29). From (10.3) we find that the quantity $\partial_\mu \phi^\lambda + \Gamma_{\mu\alpha}^{la}\phi^\alpha$ is a homogeneous tensor. We can then define the covariant derivative ∇_μ on the Riemannian manifold as

$$\nabla_\mu \phi^\alpha \;=\; \partial_\mu \phi^\alpha + \Gamma_{\mu\beta}^\alpha \phi^\beta \,. \tag{10.4}$$

Notice that we do not impose antisymmetrization of the derivative here. So we may call ∇ a symmetric covariant derivative.

Consider the contraction of the indices by $\phi^\alpha \chi_\alpha$ where χ_α denotes another set of scalar function. An analog of (10.2) for χ_α is expressed as

$$\chi^a \,=\, (e^{-1})^{\alpha a}\chi_\alpha \,. \tag{10.5}$$

The covariant derivative acting on χ_α is then defined by

$$\nabla_\mu \chi_\alpha \,=\, \partial_\mu \chi_\alpha - \Gamma^\beta_{\mu\alpha}\chi_\beta \,. \tag{10.6}$$

In general, the action of the covariant derivative ∇_μ on a tensor of the form $T^{\alpha_1\alpha_2\cdots\alpha_p}_{\beta_1\beta_2\cdots\beta_q}$ is given by

$$\begin{aligned}\nabla_\mu T^{\alpha_1\alpha_2\cdots\alpha_p}_{\beta_1\beta_2\cdots\beta_q} \;&=\; \partial_\mu T^{\alpha_1\cdots\alpha_p}_{\beta_1\cdots\beta_q} + \Gamma^{\alpha_1}_{\mu\alpha}T^{\alpha\alpha_2\cdots\alpha_p}_{\beta_1\cdots\beta_q} + \cdots + \Gamma^{\alpha_p}_{\mu\alpha}T^{\alpha_1\cdots\alpha_{p-1}\alpha}_{\beta_1\cdots\beta_q} \\ &\qquad - \Gamma^\beta_{\mu\beta_1}T^{\alpha_1\cdots\alpha_p}_{\beta\beta_2\cdots\beta_q} - \cdots - \Gamma^\beta_{\mu\beta_q}T^{\alpha_1\cdots\alpha_p}_{\beta_1\cdots\beta_{q-1}\beta} \,.\end{aligned} \tag{10.7}$$

In particular, action of ∇_α on the metric $g_{\mu\nu}$ vanishes:

$$\nabla_\alpha g_{\mu\nu} \,=\, \partial_\alpha g_{\mu\nu} - \Gamma^\lambda_{\mu\alpha}g_{\nu\lambda} - \Gamma^\lambda_{\nu\alpha}g_{\beta\lambda} \,=\, 0 \tag{10.8}$$

where we use the identity (8.30). This means that the covariant derivative is compatible with the metric. In other words, $g_{\mu\nu}$ is a covariant metric. This fact is crucial for the importance of isometries in physics.

Isometries and Killing equations

Isometries refer to the coordinate transformations

$$x^\mu \,\to\, x^\mu + \xi^\mu(x) \tag{10.9}$$

which preserve the metric $ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu$. The change in the metric can be calculated straightforwardly:

$$\begin{aligned}(ds^2)_{x+\xi} - (ds^2)_x \;&=\; \xi^\alpha\partial_\alpha g_{\mu\nu}dx^\mu dx^\nu + g_{\mu\nu}\frac{\partial\xi^\mu}{\partial x^\alpha}dx^\alpha dx^\nu + g_{\mu\nu}dx^\mu\frac{\partial x^\nu}{\partial x^\alpha}dx^\alpha \\ &=\; \left[\xi^\alpha\partial_\alpha g_{\mu\nu} + g_{\alpha\nu}\frac{\partial\xi^\alpha}{\partial x^\mu} + g_{\mu\alpha}\frac{\partial\xi^\alpha}{\partial x^\nu}\right]dx^\mu dx^\nu \,.\end{aligned} \tag{10.10}$$

Thus ξ^μ leads to an isometry if it obeys the condition

$$\xi^\alpha\partial_\alpha g_{\mu\nu} + g_{\alpha\nu}\frac{\partial\xi^\alpha}{\partial x^\mu} + g_{\mu\alpha}\frac{\partial\xi^\alpha}{\partial x^\nu} \,=\, 0 \,. \tag{10.11}$$

This is called the Killing equation. The ξ^μ is called the Killing vector. For a given $g_{\mu\nu}$ we can write down the Killing equation and solve for ξ^α to obtain the isometries.

Using the relation (8.31), *i.e.*,

$$\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu} \;=\; 2\,\Gamma^\lambda_{\mu\nu}\, g_{\alpha\lambda}\,, \tag{10.12}$$

we can simplify the Killing equation as

$$\begin{aligned} & \frac{\partial}{\partial x^\mu}\,(\xi^\alpha g_{\alpha\nu}) + \frac{\partial}{\partial x^\nu}\,(\xi^\alpha g_{\mu\alpha}) - \xi^\alpha \partial_\mu g_{\alpha\nu} - \xi^\alpha \partial_\nu g_{\mu\alpha} + \xi^\alpha \partial_\alpha g_{\mu\nu} \\ = \;& \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \xi_\beta g^{\alpha\beta}(\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu}) \\ = \;& \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - 2\xi_\beta \Gamma^\beta_{\mu\nu} \\ = \;& (\partial_\mu \xi_\nu - \Gamma^\beta_{\mu\nu}\xi_\beta) + (\partial_\nu \xi_\mu - \Gamma^\beta_{\nu\mu}\xi_\beta) \\ = \;& \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu \;=\; 0 \end{aligned} \tag{10.13}$$

where ∇'s are the symmetric covariant derivative defined in (10.6). The Killing equation can therefore be simplified as $\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0$.

10.2 Isometries and operator algebra

Isometries are given by the solutions of the Killing equation. As discussed in the previous chapter, the motion of point particles is governed by the action $\mathcal{S} = -m \int ds$. Thus we can state that

> The symmetries of point particle dynamics are equivalent to the isometries of underlying spacetime.

Another example is given by a relativistic scalar particle which obeys the Klein-Gordon equation

$$(\Box + m^2)\phi \;=\; 0 \tag{10.14}$$

where $\Box$ denotes the d'Alembert operator. In terms of the metric this can be written as

$$\Box \;=\; \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\mu} g^{\mu\nu}\sqrt{-g}\frac{\partial}{\partial x^\nu} \tag{10.15}$$

where $g = \det(g_{\mu\nu})$. A curved-space action which derives the Klein-Gordon equation is given by

$$\mathcal{S} \;=\; \frac{1}{2}\int d^4x\sqrt{-g}\,\left(\frac{\partial\phi}{\partial x^\mu}\frac{\partial\phi}{\partial x^\mu} g^{\mu\nu} - m^2\phi^2\right). \tag{10.16}$$

Symmetries of this action are also isometries of the underlying curved space.

We now claim that *isometries give a complete set of operators in quantum theory on a curves space.* In quantum theory, one interprets the scalar function ϕ as an operator which acts on a Hilbert space of the physical system. In fact, *quantum theory can be defined by one unitary irreducible representation of the*

complete set of operators for the physical system. This definition is noting to do with the choice of coordinates.

The simplest example that illustrates our claim is given by the Minkowski space $ds^2 = dx_0^2 - dx_0^1 - dx_0^2 - dx_0^3 = \eta_{\mu\nu} dx^\mu dx^\nu$ with $\eta_{\mu\nu} = (+---)$. In this case, the Killing equation becomes

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu \;=\; 0\,. \tag{10.17}$$

The most general solution is given by

$$\xi_\mu \;=\; a_\mu \;+\; \omega_{\mu\nu} x^\nu \tag{10.18}$$

where a_μ is a constant four-vector and $\omega_{\mu\nu}$ satisfies the antisymmetric property

$$\omega_{\mu\nu} + \omega_{\nu\mu} \;=\; 0\,. \tag{10.19}$$

In terms of these, the coordinate transformations are written as

$$x^\mu \to x^\mu + \xi^\mu = x^\mu + a^\mu + \omega^{\mu\nu} x^\nu\,. \tag{10.20}$$

From this expression, we find that a^μ denotes translations and $\omega^{\mu\nu}$ denotes spatial rotation plus Loretnz transformations. Combination of these transformations form the Poincaré group. The transformations of the field operator are expressed as

$$\begin{aligned} \phi(x) \;\longrightarrow\; \phi(x+\xi) \;&=\; \phi(x) + \xi^\mu \frac{\partial \phi}{\partial x^\mu} \\ &\equiv\; \phi(x) + i\mathcal{O}\cdot\phi\,. \end{aligned} \tag{10.21}$$

The operator $\mathcal{O}$ can be read off from (10.18):

$$\begin{aligned} \mathcal{O} \;&=\; a^\mu \left(-i\frac{\partial}{\partial x^\mu}\right) + \omega^{\mu\nu} x_\nu \left(-i\frac{\partial}{\partial x^\mu}\right) \\ &=\; a^\mu P_\mu - \frac{\omega^{\mu\nu}}{2} M_{\mu\nu} \end{aligned} \tag{10.22}$$

where $P_\mu = -i\frac{\partial}{\partial x^\mu}$ and $M_{\mu\nu} = x_\mu P_\nu - x_\nu P_\mu$. These are the generators of the translations and the spacetime rotations, respectively, for a (spinless) scalar field. The operator algebra of these forms the Poincaré algebra:

$$\begin{aligned} [P_\mu, P_\nu] \;&=\; 0\,, \\ [M_{\mu\nu}, P_\alpha] \;&=\; i\left(\eta_{\mu\alpha} P_\nu - \eta_{\nu\alpha} P_\mu\right)\,, \\ [M_{\mu\nu}, M_{\alpha\beta}] \;&=\; i(\eta_{\mu\alpha} M_{\nu\beta} - \eta_{\nu\alpha} M_{\mu\beta} - \eta_{\mu\beta} M_{\nu\alpha} + \eta_{\nu\beta} M_{\mu\alpha})\,. \end{aligned} \tag{10.23}$$

In other curved spaces, we can similarly obtain a corresponding operator algebra. For example, consider the Schwarzschild metric

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2GM}{r}} - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2\,. \tag{10.24}$$

From this metric, we see that the isometries are given by (a) time translation ($t \to t+a$) and (b) translations of spatial angles with fixed r, t. The corresponding operators are $P_0 = H$ (the Hamiltonian) and $L_i = \frac{1}{2}\epsilon^{ijk}M_{jk}$ (the angular momentum) where $i = 1, 2, 3..$ The operator algebra of the system is then given by

$$[H, L_i] = 0\,, \quad [L_i, L_j] = i\epsilon^{ijk}L_k\,. \tag{10.25}$$

If we further impose a symmetry, *i.e.*, a condition of no precession for the planetary motion, then the system reduces to the Kepler problem. As discussed in Chapter 2, we need to consider the Lunge-Lenz vector in addition so as to construct the full operator algebra of the Kepler problem; the relevant algebra is shown in (2.29).

10.3 FRW metric and cosmological solutions

In this section, we consider the most probable metric for the universe and its solutions. Firstly, we cannot impose time-translation invariance otherwise history does not exist. We then impose two principle on our universe:

1. Homogeneous universe: no preferred origin (translational invariance);

2. Isotropic universe: no preferred direction (rotational invariance).

We can therefore infer that the metric of universe has spatially translational and rotational Killing vectors. An ansatz for such a metric is given by the Friedmann-Robertson-Walker metric (FRW metric):

$$ds^2 \;=\; dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right] \tag{10.26}$$

where $a(t)$ denotes a time-dependent scale factor. Regarding the value of k, we only need to consider three cases, $k = 0, \pm 1$ due to the normalization for r. For $k = 0$, the FRW metric reduces to the flat space metric. For $k = 1$, the spatial part of the FRW metric becomes

$$\frac{dr^2}{1-r^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \;=\; d\alpha^2\sin^2\alpha(d\theta^2 + \sin^2 d\phi^2) \tag{10.27}$$

where we parametrize r by $r = \sin\alpha$. This gives the metric of S^3. For $k = -1$, we can use a similar parametrization $r = \sinh\alpha$ to obtain the hyperbolic (or Lobachevskian) geometry. The space geometry is then classified as follows.

$$k \;=\; \begin{cases} 0 & \text{flat} \\ 1 & S^3 \\ -1 & \text{hyperbolic} \end{cases} \tag{10.28}$$

Einstein equations for the FRW metric can be obtained straightforwardly. The right-hand side of the Einstein equation is essentially given by the fluid energy-momentum tensor

$$T_{\mu\nu} \;=\; (\rho + P)u_\mu u_\nu - P g_{\mu\nu} \tag{10.29}$$

where u_μ is fluid's four-velocity, P denotes the pressure and ρ denotes the density of the fluid. In the isotropic and homogeneous universe, non-vanishing components of $T_{\mu\nu}$ are given by

$$T_{00} = \rho\,, \quad T_{11} = T_{22} = T_{33} = P\,. \tag{10.30}$$

Then the Einstein equations reduce to

$$3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) \;=\; 8\pi G(\rho + \Lambda)\,, \tag{10.31}$$

$$-\frac{2a\ddot{a} + \dot{a}^2 + k}{a^2} \;=\; 8\pi G(P - \Lambda)\,. \tag{10.32}$$

The time derivative of (10.31) becomes

$$3\left[\frac{2\dot{a}\ddot{a}}{a^2} - \frac{2(\dot{a}^2 + k)}{a^3}\dot{a}\right] \;=\; 8\pi G\frac{d\rho}{dt}\,. \tag{10.33}$$

On the other hand, the sum of (10.31) and (10.32) can be expressed as

$$\left[\frac{2(\dot{a}^2 + k)}{a^2} - \frac{2a\ddot{a}}{a^2}\right] \;=\; 8\pi G(\rho + P)\,. \tag{10.34}$$

Thus, from (10.33) and (10.34), we find

$$\dot{\rho} + \frac{3\dot{a}}{a}(\rho + P) \;=\; 0\,. \tag{10.35}$$

This is called the Friedmann equation. Notice that this can also be derived from the energy-momentum conservation law $\partial_\mu T_{\mu\nu} = 0$.

The two basic equations of the system — the isotropic and homogeneous universe — are given by (10.31) and (10.35). Cosmological solutions are then obtained by the solutions these. Based on this idea, we can categorize the universe into three different eras.

1. Λ dominated era

We first consider the case in which Λ is dominant, $\Lambda \gg \rho$. Since the cosmological constant is one of the main candidates for dark energy. This era is usually referred to as the dark-energy dominated era. The equation (10.31) becomes

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} \;=\; \frac{8\pi G}{3}\Lambda\,. \tag{10.36}$$

For the flat space $k = 0$, we then find

$$\frac{\dot{a}}{a} \,=\, \pm\sqrt{\frac{8\pi G\Lambda}{3}}\,. \tag{10.37}$$

Observation shows that the universe is expanding. Thus we can take the plus sign in the above and obtain the solution

$$a(t) \,=\, a(0)\exp\left(\sqrt{\frac{8\pi G\Lambda}{3}}t\right)\,. \tag{10.38}$$

This shows that the universe is inflationary at this era. Since we neglect the k-term in (10.36), the spatial geometry at the inflationary epoch is essentially flat. In the current universe, the cosmological constant Λ is mostly canceled out by some other effects which no one knows. Thus we can consider that the current universe had already exited from the inflationary epoch.

2. Radiation dominated era

This era corresponds to the relativistic limit where the universe is dominated by photons. In the isotropic universe under adiabatic expansion, we have the relation

$$P \,=\, \frac{1}{3}\rho\,. \tag{10.39}$$

Then the equation (10.35) becomes

$$\frac{\dot{\rho}}{\rho} + \frac{4\dot{a}}{a} \,=\, 0 \quad\longrightarrow\quad \rho \,=\, \frac{\sigma}{a^4} \tag{10.40}$$

where σ is some constant. With $k = 0$, the equation (10.31) reduces to

$$a\dot{a} \,=\, \sqrt{\frac{8\pi G\sigma}{3}}\,. \tag{10.41}$$

The general solution is then given by

$$a^2 \,=\, 2\sqrt{\frac{8\pi G\sigma}{3}}\,t \quad\longrightarrow\quad a \sim \sqrt{t}\,. \tag{10.42}$$

The temperature of the universe $T \sim 1/a$ comes down as time goes on. This suggests that at certain point, recombination of charged particles (mostly electrons and protons), happens and neutral gas (mostly the neutral hydrogen) emerges. Consequently, radiation decouples from matter in the universe. The temperature further decreases due to the expansion after neutrization of the universe. This picture of thermal history is observationally confirmed by the 3K cosmic microwave background radiation and the helium abundance in the current universe.

3. Matter dominated era

Once radiation has decoupled from matter, the density ρ dominates and we can neglect P and Λ. We can also assume $k = 0$ for the inflationary universe. Then the basic equations (10.35), (10.31) become

$$\frac{\dot{\rho}}{\rho} + \frac{3\dot{a}}{a} = 0 \quad \longrightarrow \quad \rho = \frac{\sigma'}{a^3}\,, \tag{10.43}$$

$$\frac{3\dot{a}^2}{a^2} = \frac{8\pi G\sigma'}{a^3} \quad \longrightarrow \quad \sqrt{a}\dot{a} = \sqrt{\frac{8\pi G\sigma'}{3}} \tag{10.44}$$

where σ' is a constant. From (10.44) we find

$$\frac{d}{dt}\left(a^{\frac{3}{2}}\right) \;=\; \frac{3}{2}\sqrt{\frac{8\pi G\sigma'}{3}} \quad \longrightarrow \quad a \sim t^{2/3}\,. \tag{10.45}$$

We also find $\rho \sim \frac{\sigma'}{t^2}$. Notice that the t-dependence of ρ remains the same as the previous case in (10.40). From (10.42) and (10.45), we also see that the expansion rate of the universe is larger in the matter dominated era than the radiation dominated era.

Observations: Hubble constant

The Hubble constant is defined by

$$\left[\frac{\dot{a}}{a}\right]_{\text{today}} \;=\; H_0 \;\simeq\; h_0\left[100\,\frac{\text{km}}{\text{sec}}\frac{1}{\text{Mpc}}\right] \tag{10.46}$$

where H_0 refers to today's value of the Hubble constant $H(t)$. h_0 denotes the conventional normalization and is roughly estimated by $h_0 \sim 0.70$.

Using H_0, we can express the Einstein equation (10.31) as

$$1 + \frac{k}{a_0^2 H_0^2} \;=\; \frac{8\pi G\rho}{3H_0^2} + \frac{8\pi G\Lambda}{3H_0^2}\,. \tag{10.47}$$

The right-hand side terms represent density parameters for matter and cosmological constant. Respectively, these are denoted by $\Omega_{\text{matter}} = \frac{8\pi G\rho}{3H_0^2}$ and $\Omega_\Lambda = \frac{8\pi G\Lambda}{3H_0^2}$. Observational data shows that

$$\Omega_{\text{matter}} \approx 0.3\,, \qquad \Omega_\Lambda \approx 0.7\,. \tag{10.48}$$

Thus, from (10.47) and (10.48), we find that the spatial geometry is expected to be flat ($k = 0$) at present. There are no good theoretical reasons for this simple fact. There are actually a few big puzzles in cosmology regarding this matter. They are as follows.

1. We need inflation to explain naturally why the spatial geometry is flat. (Inflation is also necessary to get rid of the so-called horizon problem.)

2. It is not clear how the universe exited from the inflationary era.

3. Ω_{matter} would be great in an early universe. How is this compatible with the value of $\Omega_\Lambda (\approx 0.7)$ today?

Chapter 11

Conformal isometries

In this chapter, we further consider isometries. In physics there are many things which are not sensible to scale. For example, the massless photons are insensible to scale. As well-known, the cosmic microwave background radiation at $T \simeq 3K$ can be explained by the Planckian distribution of the form $\frac{1}{e^{\omega/T}-1}$ where ω denotes the angular frequency of the photon. This form, for instance, is unchanged under the scaling if we properly redefine the temperature T. We can see this fact as follows. Consider the flat metric $ds^2 = dt^2 - a^2(dx^2 + dy^2 + dz^2)$, with a being the scale factor. The wave equation for the photon is then given by $\left(\frac{\partial}{\partial tr} - \frac{1}{a^2}\right) A = 0$. For the plain wave solution $A \sim e^{\pm ipx}$, we find $\omega^2 \sim \frac{k^2}{a^2}$ ($k^2 = |\vec{k}|^2$) where $\vec{k}$ denotes the momentum vector. In other words, ω gets scaled as long as the photon is massless. As discussed in the previous chapter, temperature goes down as the universe expands, *i.e.*, we can naturally assume $T(t) \sim \frac{1}{a(t)}$. Thus in such a case the distribution function $\frac{1}{e^{|\vec{k}|/aT}-1}$ remains the same under the scaling.

The scale factor a is not exactly a symmetry of the metric in the above example. But the scale invariance of a theory is intimately related to a conformal symmetry or isometry, *i.e.*, scale invariance of the metric, in general. In fact, as we shall see later, for a two-dimensional theory that contains local interactions, the scale invariance algebraically implies the conformal invariance. In this chapter, we consider some general properties of the conformal symmetry.

11.1 Conformal symmetry and conformal Killing equation

The conformal symmetry (or isometry) is a symmetry under the scale invariance of the metric, $g_{\mu\nu} \to e^{\Omega} g_{\mu\nu}$ where Ω is some constant. In other words, the

variation of the metric is given by

$$\delta g_{\mu\nu} \;=\; \lambda g_{\mu\nu} \tag{11.1}$$

where λ is a constant. This symmetry has profound consequences. From (10.9)-(10.13), we find that the conformal isometry is defined by

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu \;=\; \lambda g_{\mu\nu} \tag{11.2}$$

with the coordinate transformations $x^\mu \to x^\mu + \xi^\mu(x)$. Contracting with $g^{\mu\nu}$, we find that (11.2) leads to the relation

$$2\nabla_\mu \xi^\mu \;=\; 4\lambda \quad\longrightarrow\quad \lambda \;=\; \frac{1}{2}\nabla\cdot\xi \tag{11.3}$$

where we use the fact that ∇_μ is a symmetric covariant derivative. For the definition of ∇_μ, see (10.4).

We can then rewrite the equation (11.2) as

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu - \frac{1}{2}(\nabla\cdot\xi) \;=\; 0\,. \tag{11.4}$$

This is called the conformal Killing equation. Any solution ξ of (11.4) gives a conformal transformation of the given metric $g_{\mu\nu}$.

Now, for simplicity, consider the flat Minkowski space, $g_{\mu\nu} = \eta_{\mu\nu}$ with $\eta_{\mu\nu} = (+---)$. Since we have $\Gamma = 0$ in this case, the conformal Killing equation becomes

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \frac{1}{2}(\partial\cdot\xi)\eta_{\mu\nu} \;=\; 0\,. \tag{11.5}$$

The solutions are given by

$$\xi_\mu \;=\; \begin{cases} x_\mu + \omega_{\mu\alpha}x^\alpha & : \text{Poincaré transformations,} \\ \epsilon x_\mu & : \text{Dilatation transformations,} \\ b^\alpha(x^2\eta_{\mu\alpha} - 2x_\mu x_\alpha) & : \text{Special conformal transformations} \end{cases} \tag{11.6}$$

where $\omega_{\mu\alpha}$ is antisymmetric in μ and α, *i.e.*, it is traceless, $\omega_{\mu\mu} = 0$. ϵ and b^α are an arbitrary constant and a vector, respectively. The solutions (11.6) show that the conformal isometry includes not only the ordinary isometry, represented by the Poincaré transformations. The dilatation solution $\xi_\mu = \epsilon x_\mu$ can easily be checked by the relations $\partial_\mu\xi_\nu + \partial_\nu\xi_\mu = 2\epsilon\eta_{\mu\nu}$ and $(\partial\cdot\xi) = \epsilon\partial_\mu x^\mu = 4\epsilon$. For the special conformal transformation $\xi_\mu = b^\alpha(x^2\eta_{\mu\alpha} - 2x_\mu x_\alpha)$, we have

$$\begin{aligned} \partial_\mu\xi_\nu &= b^\alpha(2x_\mu\eta_{\nu\alpha} - 2\eta_{\mu\nu}x_\alpha - 2\eta_{\mu\alpha}x_\nu)\,, \\ \partial_\nu\xi_\mu &= b^\alpha(2x_\nu\eta_{\mu\alpha} - 2\eta_{\mu\nu}x_\alpha - 2\eta_{\nu\alpha}x_\mu)\,, \\ \partial\cdot\xi &= b^\alpha(2x_\alpha - 8x_\alpha - 2x_\alpha) \;=\; -8b^\alpha x_\alpha\,. \end{aligned} \tag{11.7}$$

From these, we can easily check that the special conformal transformation is also a solution of the conformal Killing equation (11.5).

We now define the inversion of the coordinate $y^\mu = \frac{x^\mu}{x^2}$. A variation of $x^\mu = \frac{y^\mu}{y^2}$ under an infinitesimal transformation of $y^\mu \to y^\mu + b^\mu$ is evaluated as

$$x'^\mu = \frac{(y+b)^\mu}{(y+b)^2} \simeq \frac{y^\mu + b^\mu}{y^2 + 2y\cdot b} \simeq \frac{y^\mu}{y^2} + \frac{b^\mu}{y^2} - \frac{y^\mu}{y^2}\frac{2y\cdot b}{y^2} = x^\mu + b^\alpha(x^2\eta_{\alpha\mu} - 2x^\mu x^\alpha)\,. \quad (11.8)$$

Thus the special conformal transformation can be understood from a linear transformation of the inverse coordinate.

As discussed in the previous chapter, isometries of the Minkowski space are given by the Poincaré group. In other words, any theory built on the Minkowski space must be some representations of Poincaré group. Similarly, we can define conformal theories by choosing representations for the full set of conformal transformations. We shall consider examples of such theories in the rest of this chapter.

11.2 Conformal theories

A typical example is given by a theory of photons, *i.e.*, Maxwell's theory of electromagnetism.

Action for Maxwell's theory

It is well known that on a flat space the action for Maxwell's theory is given by

$$\mathcal{S} = -\frac{1}{4}\int d^4x\, F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}\int d^4x\,\left(E^2 - B^2\right) \quad (11.9)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor, with $A_\mu = (A_0, A_i) = (\phi, \vec{A})$ $(i = 1, 2, 3)$. Here ϕ is the electric scalar potential and $\vec{A}$ denotes the magnetic vector potential. In terms of $F_{\mu\nu}$, the electric and magnetic fields are defined by $E_i = F_{0i}$ and $F_{ij} = \epsilon_{ijk}B_k$, respectively.

On a curved space, the action can be written as

$$\mathcal{S} = -\frac{1}{4}\int \sqrt{-g}\, d^4x\, F_{\mu\nu}F_{\alpha\beta}g^{\mu\alpha}g^{\nu\beta}\,. \quad (11.10)$$

As discussed in the beginning of Chapter 9, according to the strong statement of the equivalence principle, a curved-space action can be obtained from the flat-space counterpart simply by replacing the ordinary derivative with the covariant derivative. Notice that the field strength is invariant under this prescription; $\nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}$ as the symmetric covariant derivative is defined by $\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\beta_{\mu\nu}A_\beta$. With the contraction of indices by $g_{\mu\nu}$ and the insertion of $\sqrt{-g} = \sqrt{-\det g}$, the above action therefore satisfies the strong equivalence principle.

The strong equivalence principle is generally wrong for particles with spins. The weak equivalence principle, on the other hand, allows us to add terms like $\int \sqrt{-g} d^4x \, \mathcal{R}^{\alpha}_{\mu\nu\beta} F^{\mu\nu} F^{\gamma\beta} g_{\alpha\gamma}$ to the classical action (11.10). As mentioned earlier, $\mathcal{R}$-involved terms are generally small and we can neglect them in practice. (Remember that the Riemannian curvature is relevant to a tidal force which is generally small in nature.) Now, neglecting such terms, we find that the action (11.10) is conformally invariant for any spacetime. This can easily be checked with the transformations

$$\begin{aligned} & g_{\mu\nu} \longrightarrow e^{\Omega} g_{\mu\nu} \,, \quad g^{\mu\nu} \longrightarrow e^{-\Omega} g^{\mu\nu} \,, \\ & g = \det g \longrightarrow e^{4\Omega} g \,, \quad \sqrt{-g} \longrightarrow e^{2\Omega} \sqrt{-g} \,. \end{aligned} \tag{11.11}$$

The conformal invariance of Maxwell's equations was shown back in 1909 by Bateman and Cunningham.

Scalar particle

Another example of conformal theories can be given by a theory of scalar particles. On a curved space, an action for a massive scalar particle can be written as

$$\mathcal{S} = \frac{1}{2} \int \nabla_\mu \phi \, \nabla_\mu \phi \, g^{\mu\nu} \sqrt{-g} \, d^4x - \frac{m^2}{2} \int \phi^2 \sqrt{-g} \, d^4x \tag{11.12}$$

where ϕ denotes the scalar wavefunction. This action is *not* conformally invariant in general spacetime due to the ∇_μ derivative.

For the Minkowski space, the action reduces to

$$\mathcal{S} = \frac{1}{2} \int \partial_\mu \phi \, \partial_\nu \phi \, \eta^{\mu\nu} \sqrt{-\eta} \, d^4x - \frac{m^2}{2} \int \phi^2 \sqrt{-\eta} \, d^4x \,. \tag{11.13}$$

We now consider the transformation of metric (11.11). Under the choice of

$$\phi \longrightarrow e^{-\frac{1}{2}\Omega} \phi \,, \tag{11.14}$$

the action is conformally invariant for $m = 0$. This means that *massless scalar theory in four dimensions is conformally invariant in the Minkowski space.* Notice that there is a possibility to put terms like $-\int \frac{\mathcal{R}}{6} \phi^2 \sqrt{-g} d^4x$ in the action if we take account of fluctuations from the flat space.

Four-dimensional conformal algebra

As in the case of Poincaré algebra (10.23), the algebra of conformal symmetry is given in terms of the generators of conformal transformations (11.6). These generators are given by

$$\begin{array}{ll} P_\mu = -i\partial_\mu & : \text{translation}, \\ M_{\mu\nu} = x_\mu P_\nu - x_\nu P_\mu & : \text{Lorentz transformation}, \\ D = -ix^\mu \partial_\mu & : \text{dilatation}, \\ K_\mu = -i(2x_\mu x^\nu \partial_\nu - x^2 \partial_\mu) & : \text{special conformal transformation}. \end{array} \tag{11.15}$$

The conformal algebra is then formed by the combination of Poincaré algebra (10.23) and the following relations

$$\begin{aligned} [D, P_\mu] &= iP_\mu \,, \\ [D, K_\mu] &= -iK_\mu \,, \\ [K_\mu, P_\nu] &= i2(\eta_{\mu\nu} D + M_{\mu\nu}) \,, \\ [M_{\mu\nu}, K_\alpha] &= i(\eta_{\mu\alpha} K_\nu - \eta_{\mu\alpha} K_\mu) \,. \end{aligned} \tag{11.16}$$

The number of independent generators is 15. This is the same as the dimension of the $SO(6)$ algebra. In fact, it is known that the four-dimensional conformal algebra can be regarded as the six-dimensional Lorentz algebra, or the $SO(1,5)$ algebra.

For the free theory of massless scalar particles in four dimensions, the two-point correlation function is given by

$$\langle \phi(x)\phi(y) \rangle \sim \frac{1}{(x-y)^2} \,. \tag{11.17}$$

This is a long-range correlation. It is known that the second order phase transition is characterized by a long-range correlation. Thus this observation illustrates the statement:

> Physics at the second order transitions (or at the critical points) are described by conformally invariant theories.

In other words, types of critical points are classified by representations of conformal transformations (plus some other extra operators). This scheme works exactly in two dimensions. All critical points in two dimensions have been classified in 1980s. Two-dimensional conformal theory has also been developed along with the attentive study of string theory. In the next section we shall consider the two-dimensional case in some detail.

11.3 Two-dimensional conformal transformations

For the two-dimensional plane, the metric is given by $g_{\mu\nu} = \delta_{\mu\nu}$ $(\mu, \nu = 1, 2)$. The conformal Killing equation becomes

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu - (\partial \cdot \xi)\delta_{\mu\nu} = 0 \tag{11.18}$$

which reduces to the set of equations:

$$\begin{aligned} \partial_1 \xi_1 - \partial_2 \xi_2 &= 0 \,, \\ \partial_1 \xi_2 + \partial_2 \xi_1 &= 0 \,. \end{aligned} \tag{11.19}$$

An easy way to solve this is to use complex variables. We now set

$$\xi_1 + i\xi_2 = f \,, \quad \xi_1 - i\xi_2 = \bar{f} \,,$$

$$x_1 + ix_2 = z\,, \quad x_1 - ix_2 = \bar{z}\,,$$
$$\partial_{\bar{z}} = \frac{\partial_1 + i\partial_2}{2}\,, \quad \partial_z = \frac{\partial_1 - i\partial_2}{2}\,. \tag{11.20}$$

Then the equations (11.19) are expressed by

$$\partial_{\bar{z}} f \;=\; 0\,. \tag{11.21}$$

The general solution is given by $f = f(z)$. Namely, f is an analytic function of z. Taking singularities at $z = 0, \infty$, we can write down the Laurent expansion

$$f(z) \;=\; \sum_{n=-\infty}^{\infty} -\epsilon_n\, z^{n+1} \tag{11.22}$$

where ϵ_n are expansion coefficients. Notice that if f has no singularity, the solution is given by $f =$ constant.

The conformal transformation $x_i \to x_i + \xi_i$ is realized by

$$z \;\to\; z + f(z)\,. \tag{11.23}$$

The generator of the conformal transformation corresponding to the parameter ϵ_n is then given by

$$l_n \;=\; -z^{n+1}\partial_z\,. \tag{11.24}$$

The generator satisfies the commutation relation

$$[l_m, l_n] \;=\; (m-n)l_{m+n}\,. \tag{11.25}$$

This is called the classical Virasoro algebra or the Witt algebra. The central extension of the Witt algebra is generally expressed as

$$[L_m, L_n] \;=\; (m-n)L_{m+n} + c_{m,n} \tag{11.26}$$

where $c_{m,n}$ denotes a **c**-number that commutes with all the operators. From the definition, we find $c_{m,n} = -c_{n,m}$. From the Jacobi identity $[[L_l, L_m], L_n] + [[L_m, L_n], L_l] + [[L_n, L_l], L_m] = 0$, we also find

$$(l-m)c_{l+m,n} + (m-n)c_{m+n,l} + (n-l)c_{n+l,m} \;=\; 0\,. \tag{11.27}$$

There is arbitrariness in the choice of $c_{m,n}$. Namely, the central extension (11.26) allows the transformations

$$\left\{ \begin{array}{rcl} L_m & \longrightarrow & L'_m = L_m + c_m\,, \\ c_{m,n} & \longrightarrow & c'_{m,n} = c_{m,n} - (m-n)c_{m+n} \end{array} \right. \tag{11.28}$$

where c_m is a **c**-number. In other words, we can define the values of $c_{m,n}$ in an arbitrary way such that they are absorbed into the transformations (11.28). We can thus fix certain values of $c_{m,n}$ without losing generality. Conventionally, these are chosen by $c_{m,0} = 0$ $(m \neq 0)$ and $c_{1,-1} = 0$. (This choice is suitable to

express a set of quantum corrections to the conformal generators.) With $l = 0$, (11.27) becomes $(m+n)c_{m,n} = (n-m)c_{m+n,0}$, meaning that $c_{m,n}$ vanishes unless $m+n=0$. Further, with $l = -n+1$ and $m = -1$, (11.27) becomes

$$c_{m,-m} = \frac{(m+1)}{(m-2)} c_{m-1,-(m-1)} \tag{11.29}$$

where we use $c_{1,-1} = 0$ and $c_{-m,m} = -c_{m,-m}$. Then we find

$$c_{m,n} = \frac{(m+1)m(m-1)}{3!} c_{2,-2}\delta_{m+n,0}\,. \tag{11.30}$$

An explicit form of the central extension (11.27) is given by

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m+n,0} \tag{11.31}$$

where we define $c = \frac{c_{2,-2}}{2}$. c is called the central charge or the conformal anomaly. The commutation relation (11.31) defines the algebra of conformal transformations in two dimensions, which is called the Virasoro algebra.

Unitary irreducible representations of the Virasoro algebra

Unitary irreducible representations of the Virasoro algebra define conformal theories in two dimensions. In other words, two-dimensional critical phenomena can be categorized by unitary irreducible representations of the Virasoro algebra, characterized by the value of central charge. For example, the Ising model corresponds to the case of $c = \frac{1}{2}$, the so-called XXZ spin chain corresponds to $c = 1$ and the bosonic string theory corresponds to $c = 26$.

Obviously, the study of two-dimensional conformal field theory is very rich and profound. In the following, we shall briefly sketch how to construct unitary irreducible representations of the Virasoro algebra.

Setting $m = 0, \pm 1$ in (11.31), we find

$$[L_1, L_{-1}] = 2L_0\,, \quad [L_0, L_1] = -L_1\,, \quad [L_0, L_{-1}] = L_{-1}\,. \tag{11.32}$$

These commutation relations form the $SL(2, \mathbf{C})$ algebra. This is a closed algebra. But, for higher m's, we no longer obtain a closed algebra. For example, in the case of $m = 0, \pm 1, \pm 2$, we have

$$[L_2, L_{-2}] = 4L_0 + \frac{c}{2}\,, \quad [L_0, L_2] = -2L_2\,, \quad [L_1, L_2] = -L_3\,, \quad \cdots \tag{11.33}$$

Thus for higher m's there is no closed subalgebra and we need to include the whole algebra.

Setting $m = 0$ in (11.31), we find

$$[L_0, L_n] = -nL_n\,. \tag{11.34}$$

This means that the action of L_n lowers the eigenvalue of L_0 by n. We can then construct a representation of the Virasoro algebra in terms of the highest weight state $|h\rangle$ that satisfies the conditions

$$L_0|h\rangle = h|h\rangle\,, \qquad L_n|h\rangle = 0 \;\; (n \geq 1) \tag{11.35}$$

with normalization $\langle h|h\rangle = 1$. All the other states can be obtained by actions of L_{-m}'s on $|h\rangle$:

$$\begin{array}{c} L_{-1}|h\rangle\,, \\ L_{-2}|h\rangle\,,\; L_{-1}^2|h\rangle\,, \\ L_{-3}|h\rangle\,,\; L_{-2}L_{-1}|h\rangle\,,\; L_{-1}^3|h\rangle\,, \\ \vdots \end{array} \tag{11.36}$$

These form an infinite dimensional representation of the Virasoro algebra for an arbitrary choice of m.

Notice that a quantum theory can be defined by a unitary irreducible representation of the operator algebra, with each operator acting on a physical Hilbert space. If the theory has scale invariance, the Hilbert space carries an unitary irreducible representation of the L_0 operator in terms of the Virasoro algebra. This means that for a local field theory, which contains local interactions, the scale invariance algebraically leads to the full conformal invariance.

Since $L_{-n}^\dagger = L_n$, unitarity of the representation can be imposed by

$$\begin{aligned} \langle h|L_nL_{-n}|h\rangle &= \langle h|\left(2nL_0 + \frac{c}{12}(n^3-n)\right)|h\rangle \\ &= 2nh + \frac{c}{12}n(n^2-1) \geq 0 \end{aligned} \tag{11.37}$$

where we use $\langle h|h\rangle = 1$. For $n = 1$, this means $h \geq 0$. Furthermore, for sufficiently large n, we need to have $c \geq 0$. The unitarity condition is therefore given by

$$h \geq 0\,, \quad c \geq 0\,. \tag{11.38}$$

We now consider the states $|\alpha\rangle = L_{-1}^2|h\rangle$ and $|\beta\rangle = L_{-2}|h\rangle$. These states form a subspace in which the eigenvalue of L_0 is $h+2$. From the Virasoro algebra, we find

$$M = \begin{pmatrix} \langle\alpha|\alpha\rangle & \langle\alpha|\beta\rangle \\ \langle\beta|\alpha\rangle & \langle\beta|\beta\rangle \end{pmatrix} = \begin{pmatrix} 4h(1+2h) & 6h \\ 6h & 4h+\frac{c}{2} \end{pmatrix}. \tag{11.39}$$

The eigenvalues of this matrix correspond to the norms of orthogonalized states, *i.e.*, orthogonal linear combinations of the two states $|\alpha\rangle$, $|\beta\rangle$. For irreducible representations, these norms should be non-negative. If there exists a zero-norm state, a so-called null state, we have degenerate representations. In the degenerate case, we can construct a *finite*-dimensional unitary irreducible representation and a corresponding conformal theory is called a minimal model. For irreducible representations, we thus need to require

$$\det M \;=\; 4h\left(8h^2 + (c-5)h + \frac{c}{2}\right) \geq 0\,. \tag{11.40}$$

For the minimal model, $\det M = 0$, we can specify h by the value of c:

$$h = \frac{5 - c \pm \sqrt{(c-1)(c-25)}}{16} . \tag{11.41}$$

The condition (11.40) restricts the value of h for a given c. Similar analyses can be made for other cases where the eigenvalue of L_0 is greater than $h+2$. For $c \geq 1$, any value of $h \geq 0$ is allowed to obtain unitary irreducible representations. For $c < 1$, however, it is known from these analyses that the unitary irreducible representations appear only at a discrete set of c:

$$c = 1 - \frac{6}{m(m+1)} \tag{11.42}$$

where $m = 2, 3, 4, \cdots$. In this case, h is parametrized by

$$h_{p,q} = \frac{((m+1)p - mq)^2 - 1}{4m(m+1)} \tag{11.43}$$

where $1 \leq p \leq m-1$ and $1 \leq q \leq p$. The values of c and $h_{p,q}$ characterize the unitary irreducible representations of the Virasoro algebra and, hence, the corresponding conformal theories in two dimensions.

Chapter 12

Geometric facets of Lie groups

There are two points of view to the Lie group. One is algebraic and the other is geometric. In this chapter, we shall focus on the latter perspective, mainly dealing with compact and simply-connected Lie groups.

12.1 Elements of Lie groups

Definition of the group

Let $\{a_i\}$ be a set of elements for a group G in general with $i = 1, 2, \cdots, \dim G$. The group G is then defined by the following axioms.

1. Closure under a composition law: $a_i \cdot a_j \in G$
2. Existence of the identity element: $a_i \cdot \mathbf{1} = \mathbf{1} \cdot a_i = a_i$
3. Associativity: $a_i \cdot (a_j \cdot a_k) = (a_i \cdot a_j) \cdot a_k$
4. Existence of the inverse element: $a_i \cdot (a_i^{-1}) = \mathbf{1} = (a_i^{-1}) \cdot a_i$

G is called a finite group if the number of elements is finite, while it is called an infinite group if it has ∞-number of elements.

Groups are generally split into two types; discrete and continuous. A typical example of discrete groups is a set of integers under addition. The continuous groups, on the other hand, are characterized by a continuous set of parameters which label the elements of the groups. If such parameters further have differentiable properties, the continuous groups become the Lie groups.

Definition of the Lie group

Let $g(\theta)$ be an element of a Lie group G, $g(\theta) \in G$. $g(\theta)$ depends on a continuous set of parameters, θ's. The number of such parameters defines the dimension of the Lie group, dim$G$. The composition law is realized by

$$g(\theta) \cdot g(\theta') \;=\; g\left(\beta(\theta, \theta')\right) \,. \tag{12.1}$$

In terms of this expression, the Lie groups are defined by

1. $\beta(\theta, \theta')$ is an analytic function of θ and θ'.
2. There exists a particular parameter α such that $g(\theta) \cdot g(\alpha) = \mathbf{1}$. The parameter α is also an analytic function of θ, $\alpha = \alpha(\theta)$.

We now precede further to incorporate the notion of derivative. Let $f(g)$ be a function of the group elements $g = g(\theta)$. The derivative of f with respect to θ is written as $\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial \theta}$. Thus, as a requirement of analyticity, we need to define $g(\theta + d\theta)$. For this purpose, let us consider $g(\theta)g(d\theta) = g\left(\beta(\theta, d\theta)\right)$. Expanding $\beta(\theta, d\theta)$ in powers of $d\theta$, we have

$$\beta(\theta, d\theta) \;\simeq\; \beta(\theta, 0) + \frac{\partial \beta(\theta, 0)}{\partial \theta} d\theta \;=\; \theta + \frac{\partial \beta}{\partial \theta} d\theta \tag{12.2}$$

where we set $g(\theta = 0) = \mathbf{1}$. Suppose we have N parameters. Namely, the parameters are labeled by θ^i with $i = 1, 2, \cdots, N$. The expansion form (12.2) can then be written as

$$\beta^i \;\simeq\; \theta^i + \frac{\partial \beta^i}{\partial \theta^k} d\theta^k \,. \tag{12.3}$$

Thus a proper shift of the parameter is given by

$$E^i_k \;\equiv\; \frac{\partial \beta^i}{\partial \theta^k} \,. \tag{12.4}$$

From these analyses, we find that a differential operator

$$X_i \;=\; (E^{-1})^k_i \frac{\partial}{\partial \theta^k} \tag{12.5}$$

generates an infinitesimal group composition on any function of the group elements. This is a key concept of Lie groups. The commutation relations of X_i's are calculated as

$$\begin{aligned}
[X_i, X_j] \;&=\; \left[(E^{-1})^k_i \frac{\partial}{\partial \theta^k},\, (E^{-1})^l_j \frac{\partial}{\partial \theta^l}\right] \\
&=\; \left[(E^{-1})^k_i \frac{\partial (E^{-1})^l_j}{\partial \theta^k} - (E^{-1})^j_k \frac{\partial (E^{-1})^l_i}{\partial \theta^k}\right] \frac{\partial}{\partial \theta^l} \\
&=\; \left[(E^{-1})^k_i \frac{\partial (E^{-1})^l_j}{\partial \theta^k} - (E^{-1})^j_k \frac{\partial (E^{-1})^l_i}{\partial \theta^k}\right] E^m_l \underbrace{(E^{-1})^n_m \frac{\partial}{\partial \theta^n}}_{X_m}
\end{aligned}$$

$$\equiv \quad C_{ij}^{m}\, X_m \tag{12.6}$$

where

$$C_{ij}^{m} \;=\; E_l^m \left((E^{-1})_i^k \frac{\partial (E^{-1})_j^l}{\partial \theta^k} - (E^{-1})_j^k \frac{\partial (E^{-1})_i^l}{\partial \theta^k} \right). \tag{12.7}$$

C_{ij}^{m} is in general some function of θ's.

Lie's first theorem

Lie's first theorem states that

> For a Lie group, C_{ij}^{m} is constant, being independent of θ's.

This means that one can evaluate C_{ij}^{m} by considering a small neighbourhood around $\theta = 0$. In other words, a global property of Lie groups can be obtained from a local analysis. In this sense, Lie's first theorem is analogous to Cauchy's theorem in complex analysis. The constant C_{ij}^{m} is called the structure constant. A lot of analyses in Lie groups can be done by the expansion around the origin at $\theta = 0$.

Lie algebra

Algebras in general are defined by a set of elements $\{t_a\}$ which form a vector space, *i.e.*, $\{t_a\} \in V$, $\alpha t_a + \beta t_b \in V$ (α, β are some constants). For such elements, we then consider a bracket operator, *e.g.*, a Poisson bracket, $\{t_a, t_b\} = C_{ab}^{c} t_c$ where C_{ab}^{c} is a constant. The bracket gives a mapping $V \times V \to V$. Lie algebras are defined by two properties on this bracket: (i) antisymmetry $\{t_a, t_b\} = -\{t_b, t_a\}$ and (ii) the Jacobi identity $\{t_a, \{t_b, t_c\}\} + \{t_b, \{t_c, t_a\}\} + \{t_c, \{t_a, t_b\}\} = 0$. In terms of C_{ab}^{c} the Lie algebras are then defined by

$$C_{ab}^{c} + C_{ba}^{c} \;=\; 0\,, \tag{12.8}$$

$$C_{ab}^{d} C_{cd}^{e} + C_{bc}^{d} C_{ad}^{e} + C_{ca}^{d} C_{bd}^{e} \;=\; 0\,. \tag{12.9}$$

Lie's second theorem

Lie's second theorem states that

> The differential operators $X_i = (E^{-1})_i^k \frac{\partial}{\partial \theta^k}$ form (basis) elements of a Lie algebra. For every Lie group G, there is an associated Lie algebra $\underline{G}$.

In other words, there exists a correspondence between the generators X_i and the elements t_a. The inverse of this statement is as follows.

> Given any Lie algebra $\underline{G}$, we can construct a Lie group $\widetilde{G}$, with its elements being $\widetilde{g} = \exp(i t_a \theta_a)$, but $\widetilde{G}$ is not unique. It is a simply-connected version of G known as a covering group of G.

Classification of Lie algebras

Based on the conditions (12.8) and (12.9), classification of all semisimple Lie algebras has been made by Cartan and Killing. A list of classification symbols and corresponding groups is shown below. All these groups are compact if the parameters θ's are taken to be real.

1. A_N: Special unitary groups $SU(N+1)$
2. B_N: Special orthogonal groups $SO(2N+1)$
3. C_N: Symplectic groups $Sp(2N)$
4. D_N: Special orthogonal groups $SO(2N)$
5. G_2, F_4, E_6, E_7, E_8: Exceptional groups

where $N = 1, 2, \cdots$. The dimension of each group (or associated algebra) is given by the following table.

Type	A_N	B_N, C_N	D_N	G_2	F_4	E_6	E_7	E_8
dim	$N(N+2)$	$N(2N+1)$	$N(2N-1)$	14	52	78	133	248

12.2 Geometric facets of Lie groups

In this section, we start geometric analyses of Lie groups. We first study the case of the $SU(2)$ group and then consider its generalization. Towards the end of this section, we shall show that Lie groups in general can be interpreted as Riemannian manifolds.

The $SU(2)$ group

Elements of the $SU(2)$ group are parametrized by 2×2 special unitary matrices

$$u = e^{iH}, \quad \det u = 1 \tag{12.10}$$

where H are 2×2 traceless hermitian matrices. Generally, H can be parametrized by the Pauli matrices:

$$H = \frac{\sigma_i}{2}\theta^i \tag{12.11}$$

for $i = 1, 2, 3$. The $SU(2)$ elements are then expressed as

$$g(\theta) = u = \exp\left(i\frac{\sigma_i}{2}\theta^i\right). \tag{12.12}$$

This is the same as the expression in (1.38). A variation of u (in linear order) can be calculated as

$$u + du \quad = \quad \exp\left(i\frac{\sigma_k}{2}(\theta^k + d\theta^k)\right)$$

$$
\begin{aligned}
&= 1 + i\frac{\sigma_k}{2}(\theta^k + d\theta^k) + \frac{i^2}{2!}\frac{\sigma_k}{2}\frac{\sigma_l}{2}(\theta^k + d\theta^k)(\theta^l + d\theta^l) + \cdots \\
&= u + i\frac{\sigma_k}{2}d\theta^k + \frac{i^2}{2!}\left(\frac{\sigma_k}{2}\frac{\sigma_l}{2} + \frac{\sigma_l}{2}\frac{\sigma_k}{2}\right)\theta^k d\theta^l + \cdots \\
&= u + i\frac{\sigma_k}{2}\theta^k\, i\frac{\sigma_l}{2}d\theta^l + \frac{i^2}{2}\underbrace{\left[\frac{\sigma_l}{2}, \frac{\sigma_k}{2}\right]}_{i\epsilon_{lkm}\frac{\sigma_m}{2}}\theta^k d\theta^l + \cdots \\
&= u + \left(1 + i\frac{\sigma_k}{2}\theta^k\right)\left[i\frac{\sigma_l}{2}d\theta^l - \frac{i}{2}\epsilon_{lkm}\frac{\sigma_m}{2}\theta^k d\theta^l\right] + \cdots \\
&\equiv u + iu\frac{\sigma_m}{2}E_l^m(\theta)d\theta^l \qquad (12.13)
\end{aligned}
$$

where

$$E_l^m(\theta) \simeq \delta_l^m - \frac{1}{2}\epsilon_{lkm}\theta^k\,. \qquad (12.14)$$

Alternatively, we have

$$u^{-1}du = i\frac{\sigma_m}{2}E_l^m(\theta)d\theta^l\,. \qquad (12.15)$$

Notice that Lie's theorem guarantees the expansion of $\exp\left(i\frac{\sigma_k}{2}(\theta^k + d\theta^k)\right)$ and its convergence.

The quantity $E_l^m(\theta)$ is necessary for the definition of derivatives. In fact, from $E_l^m(\theta)$ we can obtain the corresponding Lie algebra; as discussed earlier, this can be carried out by use of the differential operators $X_i = (E^{-1})_i^k\frac{\partial}{\partial\theta^k}$. Since u is represented by a 2×2 matrix, we can parametrize it as

$$u = a\mathbf{1} + b_i\sigma_i = \begin{pmatrix} a + ib_3 & ib_1 + b_2 \\ ib_1 - b_2 & a - ib_3 \end{pmatrix} \qquad (12.16)$$

where a, b_i $(i = 1, 2, 3)$ are real. The condition $\det u = 1$, or equivalently the normalization $u^\dagger u = 1$ (with $u^\dagger = a^* - ib_i\sigma_i$) leads to

$$a^2 + b_1^2 + b_2^2 + b_3^2 = 1\,. \qquad (12.17)$$

This means that the $SU(2)$ group can be identified as the 3-dimensional sphere S^3. (In fact, as we shall see later, any Lie groups can be interpreted as Riemannian manifolds. We can thus regard the relation $SU(2) \simeq S^3$ as a simple example of this interpretation.)

Using the parametrization (12.16), we can directly calculate the quantity $E_l^m(\theta)$ in (12.14). Noticing $a = \sqrt{1 - b\cdot b}$, we find

$$du = da + idb\cdot\sigma = -\frac{b\cdot db}{a} + idb\cdot\sigma\,. \qquad (12.18)$$

In terms of these parameters, the quantity $u^{-1}du$ can then be calculated as

$$u^{-1}du = (a - ib\cdot\sigma)\left[-\frac{b\cdot db}{a} + idb\cdot\sigma\right]$$

$$
\begin{aligned}
&= -b_i db_i + iadb_i\sigma_i + i\frac{b_i b_j}{a}db_j\sigma_i + db_i b_j \sigma_i\sigma_j \\
&= i\sigma_i\left[adb_i + \frac{b_i b_j}{a}db_j + \epsilon_{ijk}db_j b_k\right] \\
&\equiv i\frac{\sigma_i}{2}E^i_k(a,b)db_k\,, \qquad (12.19)
\end{aligned}
$$

$$
E^i_k(a,b) = 2\left(a\delta^i_k + \frac{b^i b_k}{a} + \epsilon_{ijk}b_k\right) \qquad (12.20)
$$

where we use the relation $\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$. The metric of the $SU(2)$ group is defined by the so-called *Cartan-Killin metric*. This is given by

$$
ds^2 = -2\mathrm{Tr}(u^{-1}du\,u^{-1}du)\,. \qquad (12.21)
$$

This metric has many isometries. One can show that a set of such isometries form the $SU(2)$ algebra. Using (12.19), we can express the Cartan-Killing metric as

$$
\begin{aligned}
ds^2 &= -2\mathrm{Tr}\left(i\frac{\sigma^a}{2}\right)\left(i\frac{\sigma^b}{2}\right)E^a_\alpha E^b_\beta\, db^\alpha db^\beta \\
&= E^a_\alpha E^b_\beta\, db^\alpha db^\beta\,. \qquad (12.22)
\end{aligned}
$$

This means that we can identify E^a_α's as frame fields for the $SU(2)$ group considered as a Riemannian manifold.

Generalization

We can straightforwardly apply these results to a general case. Let $g(\theta)$ be an element of a Lie group G. Then the generalization of the above results can be expressed as

$$
\begin{aligned}
g^{-1}dg &= it^a E^a_\alpha(\theta)d\theta^\alpha\,, \qquad (12.23)\\
ds^2 &= E^a_\alpha E^b_\beta d\theta^\alpha d\theta^\beta \qquad (12.24)
\end{aligned}
$$

where t^a $(a = 1, 2, \cdots, \mathrm{dim}G)$ are the generators of the Lie group G, satisfying the algebra $\underline{G}$:

$$
\left[t^a, t^b\right] = iC^{abc}t^c \qquad (12.25)
$$

with normalization $\mathrm{Tr}(t^a t^b) = \frac{1}{2}\delta^{ab}$. C^{abc} denote the structure constants of the algebra. From (12.23) we can define the quantity

$$
A_\alpha \equiv g^{-1}\frac{\partial g}{\partial\theta^\alpha} = it^a E^a_\alpha\,. \qquad (12.26)
$$

The derivative of A_β with respect to θ^α is calculated as

$$
\frac{\partial}{\partial\theta^\alpha}A_\beta = \left(-g^{-1}\frac{\partial g}{\partial\theta^\alpha}g^{-1}\right)\frac{\partial g}{\partial\theta^\beta} + g^{-1}\frac{\partial^2 g}{\partial\theta^\alpha\partial\theta^\beta}
$$

$$= \quad -A_\alpha A_\beta + g^{-1}\frac{\partial^2 g}{\partial\theta^\alpha\partial\theta^\beta} \tag{12.27}$$

where we use the relation $\frac{\partial g^{-1}}{\partial\theta^\alpha} = -g^{-1}\frac{\partial g}{\partial\theta^\alpha}g^{-1}$ obtained from $\frac{\partial}{\partial\theta^\alpha}(gg^{-1}) = 0$. Antisymmetrization of the derivative then leads to the identity

$$\partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta] \;=\; 0\,. \tag{12.28}$$

This is called the *Mauer-Cartan identity*. In terms of the frame fields, this identity is written as

$$\partial_\alpha E^a_\beta - \partial_\beta E^a_\alpha - C^{abc}E^b_\alpha E^c_\beta \;=\; 0\,. \tag{12.29}$$

This is analogous to the torsion-free condition in its form. In what follows, we shall further consider the analogy between the torsion-free condition and the Mauer-Cartan identity.

The torsion $T^a_{\mu\nu}$ is defined in (8.22). Its explicit form is given by

$$T^a_{\mu\nu} \;=\; \partial_\mu e^a_\nu - \partial_\nu e^a_\mu + \omega^{ab}_\mu e^b_\nu - \omega^{ab}_\nu e^b_\mu \tag{12.30}$$

where ω^{ab}_μ's are the spin connections. On the other hand, the Mauer-Cartan identity (12.29) can be rewritten as

$$\partial_\alpha E^a_\beta - \partial_\beta E^a_\alpha - \frac{1}{2}C^{abc}\left(E^b_\alpha E^c_\beta - E^b_\beta E^c_\alpha\right) \;=\; 0\,. \tag{12.31}$$

Solutions of this, or interpretations of it in comparison with (12.30), can be given by the following two cases.

Solution 1: Parallelizable manifolds

One solution is given by the choice of $\omega^{ab}_\mu = 0$; namely, $T^a_{\mu\nu} = \partial_\mu e^a_\nu - \partial_\nu e^a_\mu$ from (12.30). In this case, the torsion corresponding to the frame fields E^a_α's becomes $T^a_{\alpha\beta} = \partial_\alpha E^a_\beta - \partial_\beta E^a_\alpha$ and from (12.31) we find

$$T^a_{\alpha\beta} \;=\; C^{abc}E^b_\alpha E^c_\beta\,. \tag{12.32}$$

As introduced in (8.26), the Riemann curvature tensor $\mathcal{R}^{ab}_{\mu\nu}$ is defined by

$$R^{ab}_{\mu\nu} \;=\; \partial_\mu\omega^{ab}_\nu - \partial_\nu\omega^{ab}_\mu + \omega^{ac}_\mu\omega^{cb}_\nu - \omega^{ac}_\nu\omega^{cb}_\mu\,. \tag{12.33}$$

Thus, in this case, the Lie group G can be interpreted as a general differential manifold with *zero curvature but non-zero torsion*. Such manifolds are called parallelizable. It is known that *the only compact manifolds that are parallelizable are compact Lie groups and* S^7.

Solution 2: Riemannian manifolds

The other solution is given by $T^a_{\mu\nu} = 0$. In this case, from (12.31) we find that the spin connection corresponding to the frame fields E^a_α's becomes

$$\omega^{ab}_\alpha \;=\; \frac{1}{2} C^{abc} E^c_\alpha \,. \tag{12.34}$$

The corresponding Riemann curvature tensor is then calculated as

$$\begin{aligned} R^{ab}_{\alpha\beta} &= \partial_\alpha \omega^{ab}_\beta - \partial_\beta \omega^{ab}_\alpha + \omega^{ac}_\alpha \omega^{cb}_\beta - \omega^{ac}_\beta \omega^{cb}_\alpha \\ &= \frac{1}{2} C^{abc} \underbrace{\left(\partial_\alpha E^c_\beta - \partial_\beta E^c_\alpha\right)}_{C^{ckl} E^k_\alpha E^l_\beta} + \frac{1}{4} \left(C^{ack} C^{cbl} - C^{acl} C^{cbk}\right) E^k_\alpha E^l_\beta \\ &= \left[\frac{1}{4} C^{abc} C^{ckl} + \frac{1}{4} \left(C^{abc} C^{ckl} + C^{kac} C^{cbl} + C^{bkc} C^{cal}\right)\right] E^k_\alpha E^l_\beta \\ &= \frac{1}{4} C^{abc} C^{ckl} E^k_\alpha E^l_\beta \end{aligned} \tag{12.35}$$

where we use the Jacobi identity for the structure constants. From these, we find that Lie group G can be interpreted as a differential manifold with *zero torsion but non-zero curvature*. This manifold is nothing but a Riemannian manifold. It is therefore suitable to consider Lie groups as Riemannian manifolds in general. Their geometric quantities can be calculated from the frame field one-form $g^{-1}dg$ in (12.23).

Metric on a coset space $S^2 = SU(2)/U(1)$

Now back to the $SU(2)$ example, one can write a general $SU(2)$ element as

$$g \;=\; \frac{1}{\sqrt{1+z\bar{z}}} \begin{pmatrix} 1 & z \\ -\bar{z} & 1 \end{pmatrix} \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} \tag{12.36}$$

where $z = x + iy$ is a complex variable. For small θ, x and y, the matrix element can be expressed as

$$g \;\approx\; \begin{pmatrix} 1 + i\theta/2 & x + iy \\ -x + iy & 1 - i\theta/2 \end{pmatrix} \;=\; \mathbf{1} + i\frac{\theta}{2}\sigma_3 + ix\sigma_2 + iy\sigma_1 \,. \tag{12.37}$$

We now consider the separation of variables, *i.e.*, $g(z,\theta) = v(z)h(\theta)$ with

$$v(z) = \frac{1}{\sqrt{1+z\bar{z}}} \begin{pmatrix} 1 & z \\ -\bar{z} & 1 \end{pmatrix}, \quad h(\theta) = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}. \tag{12.38}$$

Normalization is chosen such that $v^\dagger v = 1$. The frame field one-form is then calculated as

$$g^{-1}dg \;=\; h^{-1}(v^{-1}dv)h + h^{-1}dh \tag{12.39}$$

where

$$v^{-1}dv \;=\; \frac{1}{\sqrt{1+z\bar{z}}} \begin{pmatrix} 1 & -z \\ \bar{z} & 1 \end{pmatrix}$$

$$\cdot\left[\frac{1}{\sqrt{1+z\bar{z}}}\begin{pmatrix} 0 & dz \\ -d\bar{z} & 0 \end{pmatrix} - \begin{pmatrix} 1 & z \\ -\bar{z} & 1 \end{pmatrix}\frac{\bar{z}dz + zd\bar{z}}{2(1+z\bar{z})^{3/2}}\right]$$

$$= \frac{1}{1+z\bar{z}}\begin{pmatrix} zd\bar{z} & dz \\ -d\bar{z} & \bar{z}dz \end{pmatrix} - \frac{\bar{z}dz + zd\bar{z}}{2(1+z\bar{z})}\mathbf{1}$$

$$= \frac{1}{1+z\bar{z}}\begin{pmatrix} (zd\bar{z} - \bar{z}dz)/2 & dz \\ -d\bar{z} & -(zd\bar{z} - \bar{z}dz)/2 \end{pmatrix}$$

$$= \frac{\sigma_1}{2}\frac{dz - d\bar{z}}{1+z\bar{z}} + i\frac{\sigma_2}{2}\frac{dz + d\bar{z}}{1+z\bar{z}} + \frac{\sigma_3}{2}\frac{zd\bar{z} - \bar{z}dz}{1+z\bar{z}} \tag{12.40}$$

$$h^{-1}(v^{-1}dv)h = \begin{pmatrix} (zd\bar{z} - \bar{z}dz)/2 & e^{-i\theta}dz \\ -e^{i\theta}d\bar{z} & -(zd\bar{z} - \bar{z}dz)/2 \end{pmatrix}\frac{1}{1+z\bar{z}}$$

$$= \frac{\sigma_1}{2}\frac{e^{-i\theta}dz - e^{i\theta}d\bar{z}}{1+z\bar{z}} + i\frac{\sigma_2}{2}\frac{e^{-i\theta}dz + e^{i\theta}d\bar{z}}{1+z\bar{z}} + \frac{\sigma_3}{2}\frac{zd\bar{z} - \bar{z}dz}{1+z\bar{z}} \tag{12.41}$$

$$h^{-1}dh = \begin{pmatrix} \frac{i}{2}d\theta & 0 \\ 0 & -\frac{i}{2}d\theta \end{pmatrix} = i\frac{\sigma_3}{2}d\theta\,. \tag{12.42}$$

The Cartan-Killing metric can similarly be calculated as

$$\begin{aligned} ds^2 &= -2\mathrm{Tr}(g^{-1}dg\,g^{-1}dg) \\ &= -2\mathrm{Tr}\left[(v^{-1}dv)^2 + 2v^{-1}dvdhh^{-1} + (h^{-1}dh)^2\right] \\ &= -\left(\frac{dz - d\bar{z}}{1+z\bar{z}}\right)^2 + \left(\frac{dz + d\bar{z}}{1+z\bar{z}}\right)^2 - \left(\frac{zd\bar{z} - \bar{z}dz}{1+z\bar{z}}\right)^2 - i2\left(\frac{zd\bar{z} - \bar{z}dz}{1+z\bar{z}}\right)d\theta + d\theta^2 \\ &= 4\frac{dzd\bar{z}}{(1+z\bar{z})^2} - \left(\frac{zd\bar{z} - \bar{z}dz}{1+z\bar{z}} + id\theta\right)^2 \end{aligned} \tag{12.43}$$

where the first term corresponds to the standard metric for the two-sphere S^2. Notice that it can be obtained by the stereographic projection:

$$x_1 = \frac{z + \bar{z}}{1+z\bar{z}}\,, \quad x_2 = i\frac{z - \bar{z}}{1+z\bar{z}}\,, \quad x_3 = \frac{1 - z\bar{z}}{1+z\bar{z}} \tag{12.44}$$

for the parametrization of two-sphere, $x_1^2 + x_2^2 + x_3^2 = 1$. The full $SU(2)$ metric (12.43) therefore makes the coset relation $S^2 = SU(2)/U(1)$ explicit. This metric will be useful in the analysis of spontaneous braking of the $SU(2)$ symmetry. In physics, this is relevant to the ferromagnetism and the dynamics of spin waves. We shall discuss this phenomenon in connection with Goldstone's theorem later in Chapter 14.

12.3 Irreducible representations

In this section, we consider representations of the Lie group elements $g \in G$. The group axioms introduced in the beginning of this chapter are satisfied by invertible (or regular) matrices. In fact, there exists homomorphism between elements of the group and a set of matrices.

In (12.12) we have introduced a 2×2 matrix representation of $SU(2)$ denoted by $g(\theta) = \exp\left(i\frac{\sigma^a}{2}\theta^a\right)$. This is called the defining representation of $SU(2)$. One can construct higher dimensional representations out of $g(\theta)$. For example, we can obtain a 4×4 *reducible* representation by use of block diagonalized matrices

$$\begin{pmatrix} g_1(\theta) & 0 \\ 0 & g_1(\theta) \end{pmatrix}\begin{pmatrix} g_2(\theta) & 0 \\ 0 & g_2(\theta) \end{pmatrix} = \begin{pmatrix} g_1 g_2 & 0 \\ 0 & g_1 g_2 \end{pmatrix}. \tag{12.45}$$

An irreducible representations can be defined by a representation which can not be brought to a block-diagonal form by a similarity transformation. The similarity transformation of the expression (12.45) is generally defined by

$$U^\dagger \begin{pmatrix} g_1(\theta) & 0 \\ 0 & g_1(\theta) \end{pmatrix} U\, U^\dagger \begin{pmatrix} g_2(\theta) & 0 \\ 0 & g_2(\theta) \end{pmatrix} U \;=\; U^\dagger \begin{pmatrix} g_1 g_2 & 0 \\ 0 & g_1 g_2 \end{pmatrix} U\,. \tag{12.46}$$

where U denotes some fixed 4×4 unitary matrix U.

Review of the tensor analysis

As discussed in Chapter 5 (for a particular case of the $SU(3)$ algebra), irreducible representations of a given Lie group/algebra can be constructed by means of a tensor analysis. In the following, we briefly recapitulate the tensor analysis for a general case. First of all, $N\times N$ matrices defines linear transformations of a vector space V. Let ϕ_i be a basis of V, with $i = 1, 2, \cdots N$. The transformations of ϕ_i can be expressed as

$$\phi_i' \;=\; g_{ij}\phi_i \tag{12.47}$$

where g_{ij} denotes the matrix representation of the group element of interest. Similarly, can consider a rank 2 tensor $\Psi_{ij} \equiv \phi_i\chi_j$, a set of basis elements for the direct product space $V\otimes V$ (χ_i being the other basis of V), and its transformations

$$\Psi_{ij}' \;=\; g_{ik}g_{jl}\Psi_{kl} \;\equiv\; G_{ij,kl}\Psi_{kl}\,. \tag{12.48}$$

Notice that the group composition rule $g^{(1)}\cdot g^{(2)} = g^{(3)}$ is preserved for $G_{ij,kl}$:

$$G^{(1)}_{ij,kl}G^{(2)}_{kl,mn} \;=\; g^{(1)}_{ik}g^{(2)}_{km}g^{(1)}_{jl}g^{(2)}_{ln} \;=\; g^{(3)}_{im}g^{(3)}_{jn} \;=\; G^{(3)}_{ij,mn}\,. \tag{12.49}$$

This shows that $G_{ij,kl}$ is a product representation of the two starting representations.

In general, the product representation is reducible, containing both symmetric and antisymmetric parts. To obtain an irreducible representation out of the product representation of arbitrary rank tensors, we need to apply the following reduction rules.

1. Separate symmetric and antisymmetric indices. As shown in (5.21) and (5.22), the symmetric and antisymmetric parts transform separately, regardless of the structure of groups.

2. There are invariant tensors for certain groups. One can use these to contract with either symmetric or antisymmetric indices.

As discussed in Section 5.2, the invariant tensors of $SU(3)$ are given by δ^i_j and ϵ_{ijk} where we distinguish the symmetric and antisymmetric indices by upper and lower indices, respectively. For the $SU(2)$ group, there are only one invariant tensor and it is given by ϵ_{ij}. The above tensor analysis for the construction of irreducible representations is applicable to compact groups in general.

In fact, there is a theorem due to Weyl that states

1. For a compact group, all unitary irreducible representations are finite dimensional and can be obtained by taking tensor products of the defining representation and suitably reducing them.

2. For a non-compact group, all finite-dimensional representations are non-unitary and all unitary representations are ∞-dimensional.

As mentioned at the end of Section 12.1, all groups that appear in the Cartan-Killing classification are compact if the parameters are taken to be real. A typical example of a non-compact group is, on the other hand, given by a Lorentz group.

Weyl's theorem on the rank of a Lie algebra

The rank of a Lie algebra is defined by the maximum number of simultaneously diagonalizable matrices in g_{ij} of (12.47). For example, since the Pauli matrices have only one diagonalizable matrix the rank of $SU(2)$ is 1. Similarly, that of $SU(3)$ is 2 as the Gell-Mann matrices (1.49) have two diagonalizable matrices.

Let t^a be a basis of Lie algebra ($a = 1, 2, \cdots, \dim G$) in a matrix representation. In terms of t^a, we can obtain a larger set of matrices $\{t^a, t^a t^b, t^a t^b t^c, \cdots\}$, including $t^2 = \delta^{ab} t^a t^b$, etc. These can be reduced by use of the characteristic equation (or the Cayley-Hamilton theorem) for each of t^a's. But in general these form a different algebra from the original one since the products such as t^2 do not necessarily belong to the elements of the starting Lie algebra. The resulting (larger) algebra is called the enveloping algebra. There are some elements of the enveloping algebra which commute with every element of the Lie algebra These are called the Casimir operators.

Weyl's theorem states that the rank of a Lie algebra is equal to

1. the number of independent Casimir operators of the Lie algebra; and

2. the number of independent invariant tensors in the tensor analysis of the Lie algebra.

Casimir operators for $SU(3)$

The rank of $SU(3)$ is 2. Thus there are two invariant tensors and corresponding two Casimir operators. Such a tensor can be obtained from a trace over a series of t^a's ($a = 1, 2, \cdots, 8$) since the trace is invariant under the transformations $t^a \to h^{-1}t^a h$ with $h \in G = SU(3)$. Notice that the general group element $g = \exp(it^a\theta^a)$ transforms as $g \to h^{-1}gh = \exp(ih^{-1}t^a h\theta^a)$ here. Therefore the trace

$$\mathrm{Tr}(t^a t^b) \longrightarrow \mathrm{Tr}(h^{-1}t^a hh^{-1}t^b h) \,=\, \mathrm{Tr}(t^a t^b) \,=\, \frac{1}{2}\delta^{ab} \tag{12.50}$$

is invariant under the transformations of t^a. The Casimir operator corresponding to the invariant tensor δ^{ab} is then given by $\delta^{ab}t^a t^b = t^2$. The other Casimir operator can similarly be calculated from the symmetric trace $\mathrm{Tr}\frac{1}{2}t^a(t^b t^c + t^c t^b) \equiv d^{abc}$. The Casimir is then given by $d^{abc}t^a t^b t^c$. Notice that for $SU(2)$ the symmetric symbol d^{abc} vanishes since we have $(t^b t^c + t^c t^b) = \frac{1}{2}\delta^{bc}\mathbf{1}$ for $t^a = \frac{\sigma^a}{2}$. The Casimir operators of $SU(N)$ ($N > 3$) can also be obtained in a similar fashion by calculating higher order symmetric traces.

Applications to physical problems

The knowledge of unitary irreducible representation (UIR) is very important in physics. Assume that a physical system has a symmetry governed by a Lie group G. If the symmetry is not broken as time goes on, then the physical states are classified by UIR's of G. Thus the knowledge of UIR's is crucial and algebraic methods are useful in such a system.

On the other hand, if the symmetry is spontaneously broken, a coset space G/H ($H \subset G$) becomes physically relevant. In this case, the knowledge of the coset space, such as Riemannian metric on G/H, is important and geometric methods become more useful than algebraic ones. We shall consider the physics of spontaneous symmetry breaking in Chapter 14.

Chapter 13

Wigner $\mathcal{D}$-functions and related topics

13.1 Wigner $\mathcal{D}$-functions

We first consider the adjoint representation of the $SU(N)$ group. As in the previous chapter, the group elements are written as $g = \exp(it^a\theta^a)$ where t^a are $N \times N$ traceless hermitian matrices ($a = 1, 2, \cdots, N^2 - 1$), satisfying the normalization $\mathrm{Tr}(t^a t^b) = \frac{1}{2}\delta^{ab}$. Any $N \times N$ matrix Φ can be parametrized as $\Phi = \phi^a t^a + \phi^0 \mathbf{1}$ where $\mathbf{1}$ is the $N \times N$ identity matrix.

Adjoint representation

We now consider the matrix $g^{-1}t^a g$. Since $\mathrm{Tr}(g^{-1}t^a g) = \mathrm{Tr}(t^a) = 0$, this can be parametrized as

$$g^{-1}t^a g \;=\; \mathcal{D}^{ab}(g)t^b \tag{13.1}$$

where $\mathcal{D}^{ab}(g)$ denote coefficients of the expansion. From this parametrization, we find

$$\mathcal{D}^{ab}(g_1)\mathcal{D}^{bc}(g_2) \;=\; \mathcal{D}^{ac}(g_3) \tag{13.2}$$

for the composition of the group elements $g_1 g_2 = g_3$. Thus, interpreting $\mathcal{D}^{ab}(g)$ as a matrix element (a, b) of some matrix, we find that $\mathcal{D}^{ab}(g)$ form a representation of the $SU(N)$ group. This representation is called the adjoint representation.

An infinitesimal expansion of g is given by $g = \exp(it^a\theta^a) \approx 1 + it^a\theta^a$ for a small θ^a. Similarly, we can write $\mathcal{D}^{ab}(g) \approx \delta^{ab} + i(T^c)^{ab}\theta^c$ where $(T^c)^{ab}$ denote the elements of Lie algebra in the adjoint representation. Since the infinitesimal expansion of (13.1) can be calculated as

$$(1 - it^b\theta^b)t^a(1 + it^c\theta^c) \quad = \quad t^a + i[t^a, t^c]\theta^c + \cdots$$

$$\approx \quad \delta^{ab} t^b + i\left(-if^{cab}\right)\theta^c t^b\,, \tag{13.3}$$

we find $(T^c)^{ab} = -if^{cab}$. The adjoint representation of Lie algebra is therefore given by the structure constants. Notice that the structure constants satisfy the Jacobi identity, which guarantees that they do form a Lie algebra.

Clebsch-Gordan theorem

Matrix operators act on a set of states which form a vector space. A particular state is then labeled by an irreducible representation R of the group and the vector space α. A product of such sates can be expressed as

$$|R,\alpha\rangle \otimes |R',\beta\rangle \;=\; \sum_{R'',\gamma} C^{RR'R''}_{\alpha\beta\gamma}\, |R'',\gamma\rangle \tag{13.4}$$

where $C^{RR'R''}_{\alpha\beta\gamma}$'s are expansion coefficients, called the Clebsch-Gordan coefficients. The Clebsch-Gordan theorem states that these coefficients are completely determined by group properties.

Wigner $\mathcal{D}$-functions

We now ask the following question. What is the a matrix representation of the group element $g = \exp(it^a\theta^a)$ in some representation R? Let T^a be an element of G in the R-representation. Then such a matrix representation can be expressed as

$$\begin{aligned} \mathcal{D}^{(R)}_{\alpha\beta}(g) &= \left[e^{i(T^a)\theta^a}\right]_{\alpha\beta} \\ &= \langle R,\alpha|e^{i\hat{T}^a\theta^a}|R,\beta\rangle \end{aligned} \tag{13.5}$$

where α, β label particular states in the R-representation. For fixed α and β, $\mathcal{D}^{(R)}_{\alpha\beta}(g)$ takes a value of complex number. We can thus consider (13.5) as a complex valued function on the group G. This interpretation is important because of the following theorem.

Peter-Weyl theorem

> Any function $f(g)$ on a compact Lie group G can be expanded in terms of $\mathcal{D}^{(R)}_{\alpha\beta}(g)$ where R labels a unitary irreducible representation of G. Explicitly, the function is expressed as
>
> $$f(g) \;=\; \sum_R \sum_{\alpha,\beta} b^{(R)}_{\alpha\beta}\, \mathcal{D}^{(R)}_{\alpha\beta}(g) \tag{13.6}$$
>
> where $b^{(R)}_{\alpha\beta}$ is a expansion coefficient.

The simplest example is given by the $U(1)$ group, $g = e^{i\theta}$ with $0 \leq \theta \leq 2\pi$. For arbitrary representations, the element can be parametrized as $g_n = e^{in\theta}$ with $n = 0, \pm 1, \pm 2, \cdots$ since it satisfies $g_n(\theta = 0) = g_n(\theta = 2\pi)$. Thus, in this case, the Peter-Weyl theorem says that any periodic function of θ can be expanded as

$$f(\theta) = \sum_{n=-\infty}^{\infty} b_n e^{in\theta} \,. \tag{13.7}$$

This is nothing but the Fourier expansion of $f(\theta)$. We can then regard the Peter-Weyl expansion (13.6) as a group theoretic generalization of the Fourier expansion (13.7). The existence of inverse Fourier formula then suggests that we need to think of integration over the group element g.

Integration over group

As discussed in Chapter 12, the frame field one-form is defined by

$$g^{-1}dg = it^a E^a_\alpha d\theta^\alpha \,. \tag{13.8}$$

In terms of the frame fields E^a_α the Cartan-Killing metric tensor is written as

$$g_{\alpha\beta} = E^a_\alpha E^a_\beta \,. \tag{13.9}$$

See (12.21)-(12.24) for details on these expressions. Form (13.9), we easily find $\sqrt{\det g} = \det E$. Thus the volume element over the Lie group G is given by

$$dV(g) = \det E \, d\theta^1 d\theta^2 \cdots d\theta^{\dim G} \,. \tag{13.10}$$

Now suppose we use gh instead of g for some *fixed* $h \in G$, *i.e.*,

$$(gh)^{-1}d(gh) = it^a E'^a_\alpha d\theta^\alpha \,, \tag{13.11}$$

$$dV(gh) = \det E' \, d\theta^1 d\theta^2 \cdots d\theta^{\dim G} \,. \tag{13.12}$$

The left-hand side of (13.11) can be calculated as

$$h^{-1}(g^{-1}dg)h = h^{-1}(it^a E^a_\alpha d\theta^\alpha)h = \mathcal{D}^{ab}(h)(it^b E^a_\alpha d\theta^\alpha) \tag{13.13}$$

where we introduce the adjoint representation

$$h^{-1}t^a h = \mathcal{D}^{ab}(h)t^b \,. \tag{13.14}$$

From (13.11) and (13.13), we find

$$E'^b_\alpha = \mathcal{D}^{ab}(h)E^a_\alpha \,. \tag{13.15}$$

Interpreting this as a matrix equation, $E'_{\alpha b} = E_{\alpha a}\mathcal{D}_{ab}(h)$, and taking the determinants, we find

$$\det E' = \det E \det \mathcal{D}(h) \tag{13.16}$$

where $\mathcal{D}(h) = \exp(iT^c\theta^c)$. Remember that the matrix generator $(T^c)_{ab}$ of G in the adjoint representation is given by the structure constant, $(T^c)_{ab} = -if^c_{ab}$. Since it is antisymmetric in the indices, we have $\mathrm{Tr}T^c = 0$. This means that we always have $\det\mathcal{D}(h) = 1$ for the adjoint representation of any (compact and simply-connected) Lie groups. We therefore obtain $\det E' = \det E$ and

$$dV(gh) \,=\, dV(g)\,. \tag{13.17}$$

This is analogous to the translation invariance of the integral measure over real variables, *e.g.*, $d(x+h) = dx$.

If we define another frame field by $dg\,g^{-1} = it^a\widetilde{E}^a_\alpha\theta^\alpha$, then we find $E^a_\alpha = \mathcal{D}^{ba}(g)\widetilde{E}^b_\alpha$. Thus, in a similar manner, we can show that the volume element defined by $dg\,g^{-1}$ is the same as $dV(g)$ in (13.10). This means that the volume element is the same for both right and left actions.

Orthogonality theorem

By use of the volume element $dV(g)$ we can properly define integrals over the group element g. For example, the orthogonality relation of Wigner $\mathcal{D}$-functions is given by

$$\int dV(g)\; \mathcal{D}^{(R)*}_{\alpha\beta}(g)\mathcal{D}^{(R')}_{mn}(g) \,=\, \frac{1}{(\mathrm{dim}R)}\delta^{RR'}\delta_{\alpha m}\delta_{\beta n}\,. \tag{13.18}$$

Also, the inversion formula of the Peter-Weyl theorem

$$f(g) \,=\, \sum_{R,\alpha,\beta} b^{(R)}_{\alpha\beta}\,\mathcal{D}^{(R)}_{\alpha\beta}(g) \tag{13.19}$$

is given by

$$\int dV(g)\; \mathcal{D}^{(R)*}_{\alpha\beta}(g)\, f(g) \,=\, \frac{b^{(R)}_{\alpha\beta}}{(\mathrm{dim}R)}\,. \tag{13.20}$$

13.2 Wigner-Eckart theorem

As an introduction, we consider a tensor operator of the form

$$\langle R',\alpha|Q^A|R,m\rangle \,=\, \langle R',\alpha|UU^{-1}Q^AUU^{-1}|R,m\rangle \tag{13.21}$$

where $U = U(\theta)$ is an element of some symmetry group G (with a group parameter being θ) and Q^A is some physical operator. For example, consider the coordinate operator $Q^A = x^A$ under rotations, $U(\theta) = \exp(iL^A\theta^A)$ where $L^A = \epsilon^{ABC}x^Bp^C$ is the angular momentum operator. The symmetry group is then given by $SO(3)$. In this case, the quantity $U^{-1}Q^AU$ can be calculated as

$$U^{-1}(\theta)x^AU(\theta) \,=\, e^{-iL\cdot\theta}\,x^A\,e^{iL\cdot\theta} \,=\, \mathcal{D}^{AB}(\theta)x^B \tag{13.22}$$

where $\mathcal{D}^{AB}(\theta)$ is an element of $SO(3)$ in the adjoint representation, *i.e.*, $\mathcal{D}^{AB}(\theta) = \exp\left[i(T^C)^{AB}\theta^C\right]$ where $(T^C)^{AB} = -i\epsilon^{ABC}$. One can check this relation for small θ:

$$\begin{aligned} U^{-1}(\theta)x^A U(\theta) &\approx (1 - iL^B\theta^B)x^A(1 + iL^C\theta^C) \\ &\approx x^A - i[L^B, X^A]\theta^B \\ &= x^A + \epsilon^{BAC}x^C\theta^B \\ &= \left(\delta^{AB} + i(-i\epsilon^{ABC})\theta^C\right)x^B \\ &= \left(\mathbf{1} + iT^C\theta^C\right)^{AB}x^B \end{aligned} \tag{13.23}$$

where we use the relation $[L^A, x^B] = i\epsilon^{ABC}x^C$.

This example illustrates that Q^A's form a tensor operator that transforms as a representation of the symmetry group G. Let the representation of interests, not necessarily the adjoint one, be R. Then, using the Wigner $\mathcal{D}$-function in (13.5), we can express the generalization of (13.22) as

$$U^{-1}(\theta)\, Q^A\, U(\theta) \;=\; \mathcal{D}^{(R)}_{AB}(\theta)\, Q^B \tag{13.24}$$

where we write the $\mathcal{D}$-function as $\mathcal{D}^{(R)}_{AB}(\theta)$ rather than $\mathcal{D}^{(R)}_{AB}(g)$ to show the dependence on θ explicitly.

In what follows, we look for an expression of (13.21) in terms of the Wigner $\mathcal{D}$-functions. We first notice the relation

$$U^{-1}(\theta)|R, m\rangle \;=\; \sum_n \mathcal{D}^{(R)}_{mn}(\theta)|R, n\rangle\,. \tag{13.25}$$

Taking the complex conjugate of this, we also find

$$\langle R', \alpha|U(\theta) \;=\; \sum_\beta \langle R', \beta|\mathcal{D}^{(R')*}_{\alpha\beta}(\theta)\,. \tag{13.26}$$

From (13.24)-(13.26), we then find

$$(13.21) \;=\; \mathcal{D}^{(R')*}_{\alpha\beta}(\theta)\mathcal{D}^{(R)}_{mn}(\theta)\mathcal{D}^{(R'')}_{AB}(\theta)\langle R', \beta|Q^B|R, n\rangle\,. \tag{13.27}$$

Since this holds for all θ's, we can rewrite it as

$$(13.21) \;=\; \int dV(g)\frac{\mathcal{D}^{(R')*}_{\alpha\beta}(g)\mathcal{D}^{(R)}_{mn}(g)\mathcal{D}^{(R'')}_{AB}(g)}{(\text{volume of } G)}\langle R', \beta|Q^B|R, n\rangle\,.$$

Remember that the Wigner $\mathcal{D}$-functions are written as

$$\begin{aligned} \mathcal{D}^{(R'')}_{AB}(g) &= \langle R'', A|\hat{g}|R'', B\rangle\,, \\ \mathcal{D}^{(R)}_{mn}(g) &= \langle R, m|\hat{g}|R, n\rangle\,. \end{aligned} \tag{13.28}$$

From the Clebsch-Gordan theorem in (13.4), we also have

$$|R'', B\rangle \otimes |R, n\rangle = \sum_{\widetilde{R},\lambda} C^{R''R\widetilde{R}}_{Bn\lambda} |\widetilde{R}, \lambda\rangle \,, \tag{13.29}$$

$$\langle R'', A| \otimes \langle R, m| = \sum_{\widetilde{R},\sigma} C^{R''R\widetilde{R}^*}_{Am\sigma} \langle \widetilde{R}, \sigma| \tag{13.30}$$

where $C^{R''R\widetilde{R}}_{Bn\lambda}$ and $C^{R''R\widetilde{R}}_{Am\sigma}$ are the corresponding Clebsch-Gordan coefficients; $C^{R''R\widetilde{R}^*}_{Am\sigma}$ is the complex conjugate of $C^{R''R\widetilde{R}}_{Am\sigma}$. Using (13.28)-(13.30), we find

$$\mathcal{D}^{(R'')}_{AB}(g)\mathcal{D}^{(R)}_{mn}(g) = \sum_{\widetilde{R},\lambda,\sigma} C^{R''R\widetilde{R}}_{Bn\lambda} C^{R''R\widetilde{R}^*}_{Am\sigma} \underbrace{\langle \widetilde{R}, \sigma|\hat{g}|\widetilde{R}, \lambda\rangle}_{\mathcal{D}^{(\widetilde{R})}_{\sigma\lambda}(g)} \tag{13.31}$$

where we use the same $\widetilde{R}$ for the irreducible representations of the states because the same operator $\hat{g}$ acts on these states. From the orthogonality relation (13.18), we then find that

$$\begin{aligned} (13.28) &= \sum_{\widetilde{R},\lambda,\sigma} C^{R''R\widetilde{R}}_{Bn\lambda} C^{R''R\widetilde{R}^*}_{Am\sigma} \frac{\delta_{\alpha\sigma}\delta_{\beta\lambda}\delta^{R'\widetilde{R}}}{(\dim R')(\text{volume of } G)} \langle R', \beta|Q^B|R, n\rangle \\ &= C^{R''RR'^*}_{Am\alpha} \underbrace{\left[\frac{C^{R''RR'}_{Bn\beta}\langle R', \beta|Q^B|R, n\rangle}{(\dim R')(\text{volume of } G)}\right]}_{\equiv \langle R'||Q^{R''}||R\rangle} \end{aligned} \tag{13.32}$$

where $\langle R'||Q^{R''}||R\rangle$ is called the reduced matrix. Notice that the reduced matrix is independent of the starting states $|R', \alpha\rangle$, $|R, m\rangle$. In terms of it the resultant form is given by

$$\langle R', \alpha|Q^A|R, m\rangle = C^{R'RR''}_{\alpha mA} \langle R'||Q^{R''}||R\rangle \tag{13.33}$$

where we use $C^{R''RR'^*}_{Am\alpha} = C^{R''R'R}_{A\alpha m}$. This expression is known as Wigner-Eckart theorem. This is useful in physics. In the following section, we shall consider some of its applications.

13.3 Wigner-Eckart type applications

In this section, we consider the following three applications of the Wigner-Eckart theorem.

1. Selection rules for absorption and emission of radiation
2. Hadron spectroscopy
3. Case-Gasiorowicz-Weinberg-Witten theorem

13.3.1 Selection rules for absorption and emission of radiation

As discussed in Chapter 2, the Hamiltonian of hydrogen atoms is given by

$$H_0 \;=\; \frac{p^2}{2m} - \frac{e^2}{r}\,. \tag{13.34}$$

See (2.1) for notations. Since this is spherically symmetric, we have

$$\left[L^A, H\right] \;=\; 0 \tag{13.35}$$

where, as defined earlier, L^A denotes the angular momentum operator. Let E_α be the energy eigenvalue of the eigenstate $|\alpha\rangle$, *i.e.*, $H|\alpha\rangle = E_\alpha|\alpha\rangle$. From (13.35), we find $L^A H|\alpha\rangle = HL^A|\alpha\rangle = E_\alpha L^A|\alpha\rangle$. This means that $L^A|\alpha\rangle$ has the same energy as $|\alpha\rangle$. Moreover, we can easily infer that rotation of the state $e^{iL^A\theta^A}|\alpha\rangle$ has the same energy as $|\alpha\rangle$. In other words, energy levels can be grouped into multiplets which form representations of the angular momentum.

Suppose that $|\alpha\rangle$ belongs to some irreducible representation R and $|\beta\rangle = e^{iL\cdot\theta}|\alpha\rangle$ for some θ. Then we have $\langle\beta|e^{iL\cdot\theta}|\alpha\rangle = \langle\alpha|e^{-iL\cdot\theta}e^{iL\cdot\theta}|\alpha\rangle = 1$. This relation holds for $|\beta\rangle \in R$, otherwise group properties would not hold between the elements $e^{-iL\cdot\theta}$ and $e^{iL\cdot\theta}$. On the other hand, if $|\beta\rangle \notin R$, we have $\langle\beta|e^{iL\cdot\theta}|\alpha\rangle = 0$. Therefore degenerate states are given by irreducible representations of the rotation group. Notice that details of this point are discussed in chapter 1.

We now introduce an interaction term in the Hamiltonian:

$$H \;=\; H_0 + e\vec{E}\cdot\vec{x} \tag{13.36}$$

where $\vec{E}$ denotes the electric field. The interaction term is given in the dipole approximation since x^A transforms as a spin $l = 1$ tensor operator. Physical states are labeled by $|l, m\rangle$ where m is the ordinary magnetic quantum number. By use of the Wigner-Eckart theorem (13.33), a matrix element of interest is then expressed as

$$\langle l, m|x^A|l', m'\rangle \;=\; C^{1ll'}_{Amm'}\langle l||x^{(1)}||l'\rangle \tag{13.37}$$

where $x^{(1)}$ denotes that x^A transforms as a $l = 1$ tensor. The l-part of the Clebsch-Gordan coefficient leads to the selection rules for $\Delta l = l - l'$:

$$\Delta l = 0, \pm 1 \text{ and } \Delta l = 0 \text{ (but } l = 0 \not\leftrightarrow 0) \tag{13.38}$$

The m-part of $C^{1ll'}_{Amm'}$ can similarly be evaluated. Since the magnetic quantum number of x^A $(A = 1, 2, 3)$ corresponds to $0, \pm 1$, the selection rules for $\Delta m = m - m'$ become

$$\Delta m = 0, \pm 1\,. \tag{13.39}$$

These selection rules (13.38) and (13.39) are the usual ones for radiation transitions in the hydrogen-like atoms. In the electric dipole approximation, correlation between polarizations and intensions of spectral lines are given by the Clebsch-Gordan coefficient $C^{1ll'}_{Amm'}$ in (13.37).

The next order correction can be considered by introducing the quadrupole coupling $\partial_A E_B \left(x^A x^B - \frac{1}{3}\delta^{AB}x^2\right)$. In this case the Clebsch-Gordan coefficient of interest becomes $C^{2ll'}_{(AB)mm'}$ and the corresponding selection rules are given by

$$\begin{aligned} &\Delta l = 0, \pm 1, \pm 2 \text{ (but } l = 0 \not\leftrightarrow 0, 1 \text{ and } l = \tfrac{1}{2} \not\leftrightarrow \tfrac{1}{2}), \\ &\Delta m = 0, \pm 1, \pm 2. \end{aligned} \tag{13.40}$$

The above examples illustrate the usefulness of the Wigner-Eckart theorem in the study of spectroscopy. In many cases the theorem provides a formal answer to a problem of spectroscopy.

13.3.2 Hadron spectroscopy

We have already discussed the spectroscopy of hadrons in Chapter 5. In what follows, we shall make a brief review of it in connection with the Wigner-Eckart theorem.

Hadrons are strongly interacting particles and they are categorized into mesons and baryons. Underlying dynamics of hadrons is governed by quantum chromodynamics (QCD) and is calculable only in perturbative region of the theory. Low energy mesons and baryons consist of three light quarks (u, d, s). Our strategy is to use group theory for the analysis of meson/baryon multiplets, masses and interactions.

In making such an analysis, we should notice that there is a hierarchy of strength among interactions:

$$\text{I. strong interaction} \gg \underbrace{\text{II. mass effect} \gg \text{III. interaction with } W_\pm, Z, A}_{\text{weak interaction}} \tag{13.41}$$

where $W_\pm, Z$ denote weak bosons and A denotes gluons. Symmetries of QCD are good at strong level I, somewhat broken at level II, and further broken at level III. At level I, strong forces treat all three quarks (u, d, s) identically. This means that there are unitary transformations among (u, d, s) states, $(u', d', s')^T = g(u, d, s)^T$ where g is a 3×3 matrix element of the $U(3)$ group. Any $U(3)$ matrix element can be written as $g = e^{i\theta}\tilde{g}$ where $\det\tilde{g} = 1$ and the phase factor $e^{i\theta}$ corresponds to the $U(1)$ factor of $U(3) = U(1) \times SU(3)$. Thus $|q\rangle = (u', d', s')^T$ transforms as the **3** representation of the $SU(3)$ group (See Section 5.2 on irreducible representations of $SU(3)$.) In other words, *mesons and baryons fall into degenerate multiplets which are irreducible representations of $SU(3)$*. Notice that we do not know the details of QCD dynamics but the symmetry of it, and we shall utilize this fact to consider spectroscopy of hadrons at level II of interactions.

Mesons are made of a pair of quark and antiquark. Thus, as discussed in (5.35), there exist octet and singlet mesons because of the following relation

$$\mathbf{3} \otimes \mathbf{3}^* \;=\; \mathbf{1} \oplus \mathbf{8}\,. \tag{13.42}$$

Typical examples for these multiplets are listed in Table 5.2. The basic energy scale for strong interaction is about 1 GeV. This scale is also called the QCD scale. The masses of octet mesons can be regarded as identical in the QCD scale. One can in fact expect many other octet mesons up to about 2 GeV. Above that energy scale, the mesons decay and the world is described by the so-called quark-gluon plasma.

Similarly, multiplets of baryons which are bound states of three quarks can be found from

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} \;=\; \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}\,. \tag{13.43}$$

Examples of octet and decuplet baryons are given in (5.3) and (5.4).

Now we consider the mass effects. As discussed in Section 5.3, the mass matrix can be written as $\langle q|m|q\rangle = \mathrm{diag}(m_u, m_d, m_s) \approx \mathrm{diag}(m_u, m_u, m_u + \Delta)$ at this level where $|q\rangle = (u, d, s)^T$ and $m_u \approx m_d \ll m_s$ with $\Delta \equiv m_s - m_u$. The general mass operator M for mesons and baryons can then be written as $M = M_0\mathbf{1} + M^a$ where M^a transforms as generators of $SU(3)$. Using the Gell-Mann matrices λ^a ($a = 1, 2, \cdots, 8$) in (1.49), we find $c\mathbf{1} - \sqrt{3}c\,\lambda^8 = \mathrm{diag}(0, 0, 3c)$ where c is a constant and $\mathbf{1}$ is the 3×3 identity matrix. Thus we can specify the index a of M^a to $a = 8$.

To illustrate a Wigner-Eckart type application, we here consider the mass matrix for the octet baryons. The mass matrix is expressed as

$$\begin{aligned} \mathcal{M}_{bc} \;&\equiv\; \langle \mathbf{8}, b|M|\mathbf{8}, c\rangle \;=\; M_0\,\delta_{bc} \,+ \langle \mathbf{8}, b|M^a|\mathbf{8}, c\rangle\,, \\ \langle \mathbf{8}, b|M^a|\mathbf{8}, c\rangle \;&=\; \langle \mathbf{8}, b|M^8|\mathbf{8}, c\rangle \;=\; C^{\mathbf{888}}_{8bc} \underbrace{\langle \mathbf{8}||M^{\mathbf{8}}||\mathbf{8}\rangle}_{\text{reduced matrix}} \end{aligned} \tag{13.44}$$

where indices b, c transform as the octet/adjoint representation $\mathbf{8}$ and $C^{\mathbf{888}}_{8bc}$ denotes the Clebsch-Gordan coefficient. By use of the tensor analysis, we find

$$\mathbf{8} \otimes \mathbf{8} \;=\; \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \oplus \mathbf{10}^* \oplus \mathbf{27}\,. \tag{13.45}$$

Using this result, the Clebsch-Gordan coefficients can be defined by

$$\begin{aligned} |\mathbf{8}, b\rangle \otimes |\mathbf{8}, c\rangle \;&=\; C^{\mathbf{188}}_{0bc}|\mathbf{1}, 0\rangle + C^{\mathbf{888}}_{abc}|\mathbf{8}, a\rangle + \widetilde{C}^{\mathbf{888}}_{abc}|\mathbf{8}, a\rangle + C^{\mathbf{10\,8\,8}}_{\alpha bc}|\mathbf{10}, \alpha\rangle \\ &\quad + \widetilde{C}^{\mathbf{10\,8\,8}}_{\alpha bc}|\mathbf{10}, \alpha\rangle \;+\; C^{\mathbf{27\,8\,8}}_{Abc}|\mathbf{27}, A\rangle\,. \end{aligned} \tag{13.46}$$

In this way, one can determine $C^{\mathbf{888}}_{8bc}$ and the mass formulae among octet baryons (13.45) can be obtained. This method is powerful and general, however, there exists an easier way to compute the coefficients $C^{\mathbf{888}}_{abc}$. The coefficients are constant, which literally means that they are invariant under the group transformations. Thus these can be parametrized by the Casimir invariants of the

$SU(3)$ algebra. For the rank-3 tensors (in terms of the group indices), these invariants are given by the trace

$$\begin{aligned} \mathrm{Tr}(\lambda^a\lambda^b\lambda^c) &= \mathrm{Tr}\lambda^a\frac{\left[\lambda^b,\lambda^c\right]}{2} + \mathrm{Tr}\lambda^a\frac{\left\{\lambda^b,\lambda^c\right\}}{2} \\ &= i2f^{abc} + 2d^{abc} \end{aligned} \tag{13.47}$$

where f^{abc} and d^{abc} are the structure constants and the totally symmetric symbols of $SU(3)$, respectively. Notice that the trace is invariant under $\lambda^a \to \tilde{g}\lambda^a\tilde{g}^{-1}$ where $\tilde{g}$ is the 3×3 matrix element of $SU(3)$. In terms of the adjoint representation $\mathcal{D}^{ak}(\tilde{g})\lambda^k = \lambda^a \to \tilde{g}\lambda^a\tilde{g}^{-1}$, the trace is also written as $\mathcal{D}^{ak}\mathcal{D}^{bl}\mathcal{D}^{cm}\mathrm{Tr}(\lambda^k\lambda^l\lambda^m) = \mathrm{Tr}(\lambda^a\lambda^b\lambda^c)$. The mass formula (13.44) for the octet baryons is then expressed as

$$\mathcal{M}_{bc} = A\delta_{bc} + Bf^8_{bc} + Cd^8_{bc} \tag{13.48}$$

where A, B and C are constant. From this formula, we can recover the relations in (5.46) and (5.47).

For decuplet baryons we can similarly construct mass formulae by considering the mass matrix $\langle \mathbf{10},\alpha|M^8|\mathbf{10},\beta\rangle$. These analyses can also be applicable to interactions between mesons and baryons. Consider 2-1 scattering processes $B + M \to B$ where B and M denote octet baryons and octet mesons, respectively. From the above argument, we find that the scattering amplitudes $\langle B|M|B\rangle$ can be given by $\langle \mathbf{8},b|M^a|\mathbf{8},c\rangle \sim C^{\mathbf{888}}_{abc}$ which is parametrized by f_{abc} and d_{abc}. Thus an interaction Lagrangian can be written down as

$$\mathcal{L}_{\rm int} = \lambda_1\bar{B}^bB^cM^af_{abc} + \lambda_2\bar{B}^bB^cM^ad_{abc}\,. \tag{13.49}$$

Notice that there are a lot of BM interactions, $8^3 = 512$ in principle, but the above terms are encoded only by two parameters (λ_1,λ_2). For example, the first term describes interactions among (p,n,π) for one choice of (λ_1,λ_2) and (Σ,Λ,K) for another. In this sense, the Wigner-Eckart type theorem (13.44) is a powerful and generic tool in the study of hadron dynamics.

13.3.3 Case-Gasiorowicz-Weinberg-Witten theorem

The Case-Gasiorowicz-Weinberg-Witten (CGWW) theorem consists of two statements:

1. There are no massless electrically charged particles of spin $> \frac{1}{2}$.
2. There are no massless particles of spin > 1 which have a conserved energy-momentum tensor.

A proof is as follows. Let p^μ be a four-momentum of a massless particle, $p^\mu = (\omega,\vec{p})$ with $\omega = p^0 = |\vec{p}|$. Let J^μ be a current operator. The charge

operator of the particle is given by

$$\widehat{Q} \;=\; \int d^3x\, J^0\,. \tag{13.50}$$

We now consider a matrix element of J^0 between one-particle states $|p\rangle$ and $|p'\rangle$, $\langle p'|J^0|p\rangle$. At the limit of $|p'\rangle \to |p\rangle$, it becomes

$$\lim_{p'\to p} \langle p'|J^0|p\rangle \;=\; \frac{Q}{V} \tag{13.51}$$

where V is the spacial volume and Q is the charge of the particle labeled by $|p\rangle$, *i.e.*, Q is the eigenvalue of $\widehat{Q}$ acting on the one-particle state $|p\rangle$. We have the expression (13.51) such that it is consistent with the definition (13.50).

The matrix element of the four-vector current J^μ is then expressed as

$$\lim_{p'\to p} \langle p'|J^\mu|p\rangle \;=\; Q\frac{p^\mu}{\omega V}\,. \tag{13.52}$$

Since p^μ and p'^μ are light-like vectors, $p\cdot p = p'\cdot p' = 0$, we have

$$(p+p')^2 \;=\; 2p\cdot p' \;=\; 2\omega\omega'(1-\cos\theta) > 0 \tag{13.53}$$

where θ is the angle between $\vec{p}$ and $\vec{p}'$. This means that $(p+p')^\mu$ is a time-like vector, *i.e.*, the time component dominates over the space components. One can then find a frame where $\vec{p}+\vec{p}' = 0$, corresponding to the choice of $\cos\theta = -1$. Namely, we can take the center of momentum frame:

$$p^\mu = (\omega, \vec{p})\,, \quad p^{\mu\prime} = (\omega, -\vec{p})\,. \tag{13.54}$$

In this frame, the matrix element of J^μ is given by $\langle -\vec{p}|J^\mu|+\vec{p}\rangle$.

Let $L_{\hat{p}}$ be an angular momentum operator along the direction of $\vec{p}$. Then, by definition, we have

$$e^{iL_{\hat{p}}\phi}|+\vec{p}\rangle \;=\; e^{is\phi}|+\vec{p}\rangle\,, \quad e^{iL_{\hat{p}}\phi}|-\vec{p}\rangle \;=\; e^{-is\phi}|-\vec{p}\rangle \tag{13.55}$$

where s is the spin of the massless particle and ϕ denotes the angle of counter-clockwise rotation around the axis of $\vec{p}$. The eigenvalue of $L_{\hat{p}}$ acting on $|\pm\vec{p}\rangle$ is given by $\pm s$, which is known as the helicity of the massless particle. From (13.55), we find

$$\begin{aligned}
\langle -\vec{p}|\,\vec{J}\,|+\vec{p}\rangle &= \langle -\vec{p}|\; e^{-iL_{\hat{p}}\phi}\, e^{iL_{\hat{p}}\phi}\,\vec{J}\, e^{-iL_{\hat{p}}\phi} e^{iL_{\hat{p}}\phi}\;|+\vec{p}\rangle \\
&= e^{i2s\phi}\langle -\vec{p}|\; e^{iL_{\hat{p}}\phi}\,\vec{J}\, e^{-iL_{\hat{p}}\phi}\;|+\vec{p}\rangle \\
&= e^{i2s\phi}\mathcal{D}^{ab}(\phi)\langle -\vec{p}|\,J^b\,|+\vec{p}\rangle
\end{aligned} \tag{13.56}$$

where we use $\langle -\vec{p}|e^{-iL_{\hat{p}}\phi} = \langle -\vec{p}|e^{is\phi}$ and the expression

$$e^{iL_{\hat{p}}\phi}\, J^a\, e^{-iL_{\hat{p}}\phi} \;=\; \mathcal{D}^{ab}(\phi)\, J^b\,. \tag{13.57}$$

The indices a, b denote the spacial components of the current operator. The matrix element $\langle -\vec{p}|\vec{J}|+\vec{p}\rangle$ vanishes unless $e^{i2s\phi}\mathcal{D}^{ab}(\phi) = \delta^{ab}$ at least for some a, b. Since $\langle -\vec{p}|\vec{J}|+\vec{p}\rangle \sim Q$ is not zero for charged particles, equation (13.56) leads to the condition

$$e^{2is\phi}\mathcal{D}^{ab}(\phi) = \delta^{ab} \tag{13.58}$$

for charged massless particles. Notice that the $\mathcal{D}$-function is an adjoint representation of the $SO(3)$ group (sub)elements that are *projected* onto the plane perpendicular to the $\hat{p}$ direction. Hence, it is an Abelian **c**-number, having the value of $e^{\pm i\phi}$ or 1. To satisfy the condition (13.58), we therefore need to impose $2s \leq 1$. This concludes the proof of the first statement.

The second part can analogously be proven by considering matrix elements of the rank-2 energy-momentum tensor, $\langle p'|T^{\mu\nu}|p\rangle$. One can then show that the spin of massless particles should satisfy $2s \leq 2$ in order for them to have non-vanishing matrix elements. Since the value of matrix elements are in general evaluated by the Clebsch-Goradan coefficients, these proofs can be seen as an elementary application of the Wigner-Eckart theorem.

The CGWW theorem seems to violate the existence of massless spin-1 gluons or massless spin-2 gravitons. But the theorem has rather stringent conditions on this matter and the two particles are allowed to exist. For gravitons, they do not have conserved energy-momentum tensor. Since gravitons transform with a local frame, we can not define local energy-momentum density, *i.e.*, $\partial_\mu T^{\mu\nu} \neq 0$. It is known that the conservation holds only for a pseudo-tensor $t^{\mu\nu}$, $\partial_\mu t^{\mu\nu} = 0$. Corollary of this fact is the existence of massless spin-$\frac{3}{2}$ gravitino in a supergravity theory due to local supersymmetric invariance.

For gluons, they can have charges but there is no gauge-invariant notion of charge. Namely, even if we can define a conserved current $\partial_\mu J^\mu = 0$, it is not possible to define a *covariantly conserved* current $D_\mu J^\mu = 0$ where D_μ is a covariant derivative. Thus for spin-1 gauge bosons we can not apply the CGWW theorem. Notice that the same is true for gravitons since we can interpret them as spin-2 gauge bosons.

Chapter 14

Spontaneous symmetry breaking

14.1 Goldstone's theorem

Let G be a symmetry of a Hamiltonian $\mathcal{H}$, *i.e.*, $[Q^a, \mathcal{H}] = 0$ where Q^a is a generator of the symmetry group G. The symmetry is called spontaneously broken if the ground state $|\Omega\rangle$ does not preserve the symmetry, *i.e.*, $Q^a|\Omega\rangle \neq |\Omega\rangle$.

Heisenberg ferromagnet

A typical example of spontaneous symmetry breaking is given by the Heisenberg ferromagnet. Its Hamiltonian is defined by

$$\mathcal{H} = -\sum_{ij} J_{ij} \vec{S}_i \cdot \vec{S}_j \tag{14.1}$$

where i, j denote lattice points, J_{ij}'s are interaction coefficients and $\vec{S}_i = S_i^a$ $(a = 1, 2, 3)$ is a spin vector at the point i. Since the lattice indices are all contracted, the Hamiltonian has full rotational symmetry. A rotational operator can be expressed as

$$L^a = \sum_i S_i^a \tag{14.2}$$

which satisfies

$$\left[L^a, S_i^b\right] = i\epsilon^{abc} S_i^c \,. \tag{14.3}$$

Thus we can confirm the rotational symmetry

$$[L^a, \mathcal{H}] = 0 \,. \tag{14.4}$$

On the other hand, the ground state $|\Omega\rangle$ has net magnetization so that it is not invariant under rotation, $L^a|\Omega\rangle \neq |\Omega\rangle$. In the following, we shall consider how to implement such a ground state.

Let A_M be some tensor operator for a symmetry group G which we take to be a Lie group. In the present case, we are interested in the $SO(3)$ group and the group element is given by $g = \exp(iT^a\theta^a)$ where T^a denotes an angular momentum operator ($a = 1, 2, 3$). Suppose that the ground state is invariant under G with no spontaneous symmetry breaking, $e^{iT^a\theta^a}|\Omega\rangle = |\Omega\rangle$. Then the expectation value of A_M in the ground state becomes

$$\begin{aligned} \langle\Omega|A_M|\Omega\rangle &= \langle\Omega|e^{-iT\cdot\theta}A_M e^{iT\cdot\theta}|\Omega\rangle \\ &= \mathcal{D}_{MN}(\theta)\,\langle\Omega|A_N|\Omega\rangle \end{aligned} \tag{14.5}$$

where θ is arbitrary and, as discussed in the previous chapter, $\mathcal{D}_{MN}(\theta)$ denotes the adjoint representation of G:

$$e^{-iT\cdot\theta}A_M e^{iT\cdot\theta} = \mathcal{D}_{MN}(\theta)\,A_N\,. \tag{14.6}$$

Since $\mathcal{D}_{MN}(\theta) \neq \delta_{MN}$ for arbitrary θ, (14.5) leads to the relation $\langle\Omega|A_M|\Omega\rangle = 0$. Namely, if the ground state (or the vacuum) is invariant under the symmetry group G, then the vacuum expectation value of a tensor operator A_M for G vanishes. In other words, we have the condition

$$\langle\Omega|A_M|\Omega\rangle \neq 0 \quad\longrightarrow\quad e^{iT\cdot\theta}|\Omega\rangle \neq |\Omega\rangle\,. \tag{14.7}$$

In order to tell whether the symmetry is spontaneously broken or not, the vacuum expectation value $\langle\Omega|A_M|\Omega\rangle$ should be calculated. To carry out the computation, we first need to find the ground state by minimizing the Hamiltonian.

Spontaneous breaking of $U(1)$ symmetry

As an example of how to calculate $\langle\Omega|A_M|\Omega\rangle$, we now consider a complex scalar field $\phi(x)$ with a Hamiltonian

$$\mathcal{H} = \int d^3x \left[\dot{\phi}^*\dot{\phi} + (\nabla\phi)^*(\nabla\phi) + \lambda(\phi^*\phi)^2 + \sigma\phi^*\phi\right] \tag{14.8}$$

where $\dot{\phi} = \frac{\partial}{\partial t}\phi$ and $\lambda > 0$. If $\lambda < 0$, then $\mathcal{H} \to -\infty$ at large ϕ and no grand states exist in the theory. In other words, $\lambda > 0$ is necessary for $\mathcal{H}$ to be bounded below. If σ is positive, every term in (14.8) is positive. In quantum mechanics, $\mathcal{H}$ should be bounded below, otherwise we can not define a ground state and the corresponding theory becomes disastrous. Notice that, as discussed in Chapter 2, the Hamiltonian of Hydrogen atoms $\mathcal{H} = \frac{\vec{p}^2}{2m} - \frac{e^2}{r}$ is not bounded below. For short r, however, the value of $|\vec{p}|$ goes up. Thus, quantum mechanically, we can obtain a ground state in this case.

Classically, a ground state of (14.8) is given by

$$\dot{\phi} = 0\,, \quad \nabla\phi = 0\,, \quad \phi = 0\,. \tag{14.9}$$

In quantum theory, however, finding a ground state is a complicated issue. Thus our strategy is to find a quantum ground state such that expectation values correspond to the classical ground-state values. In the present case, we require

$$\langle\Omega|\widehat{\phi}|\Omega\rangle \;=\; \langle\Omega|\dot{\widehat{\phi}}|\Omega\rangle \;=\; \langle\Omega|\nabla\widehat{\phi}|\Omega\rangle \;=\; 0 \tag{14.10}$$

where we denote the operators by putting a hat. Semi-classically, this requirement is fine but it is not exact. For instance, consider the expectation value

$$\begin{aligned}\langle\Omega|\widehat{\phi}^*\widehat{\phi}|\Omega\rangle \;&=\; \langle\Omega|\widehat{\phi}^*|\Omega\rangle\langle\Omega|\widehat{\phi}^*|\Omega\rangle \\ &\quad + \sum_{K\neq\Omega} \underbrace{\langle\Omega|\widehat{\phi}^*|K\rangle\langle K|\widehat{\phi}|\Omega\rangle}_{=\mathcal{O}(\hbar)}\end{aligned} \tag{14.11}$$

where the last term corresponds to contributions from off-diagonal elements; semi-classically, these off-diagonal elements are negligible.

For $\sigma < 0$ (or for $T < T_c$ if we use a parametrization $\sigma = c(T - T_c)$, $c > 0$), the Hamiltonian can be written as

$$\mathcal{H} \;=\; \int d^3x \left[|\dot{\phi}|^2 + |\nabla\phi|^2 + \lambda\left(|\phi|^2 - \frac{|\sigma|}{2\lambda}\right)^2 - \frac{|\sigma|}{4\lambda}\right]. \tag{14.12}$$

The minimum is then given by $\dot{\phi} = 0$, $\nabla\phi = 0$ and $|\phi|^2 = \phi^*\phi = \frac{|\sigma|}{2\lambda}$. Namely, for the minimum we have

$$\phi = \sqrt{\frac{|\sigma|}{2\lambda}}e^{i\alpha} \tag{14.13}$$

where α is a constant phase factor. Quantum theoretically, this result is expressed as

$$\langle\Omega|\widehat{\phi}|\Omega\rangle \;=\; \sqrt{\frac{|\sigma|}{2\lambda}}e^{i\alpha} \;+\; \mathcal{O}(\hbar)\,. \tag{14.14}$$

The Hamiltonian is symmetric under $\phi \to \phi e^{i\theta}$ where θ is a constant. Variation of ϕ under this $U(1)$ transformation is given by

$$\delta\phi \;=\; i\theta\phi\,. \tag{14.15}$$

The generator of this transformation for the Hamiltonian (14.12) can be defined by

$$Q \;=\; i\int d^3x\,(\phi\pi - \phi^*\pi^*) \tag{14.16}$$

where π denotes the canonical momentum which is conjugate to ϕ. This can be calculated as $\pi = \dot{\phi}^*$ and satisfies the commutation relation

$$[\pi(x), \phi(y)] \;=\; -i\delta^{(3)}(x-y)\,. \tag{14.17}$$

In terms of Q we can then calculate $\delta\phi$ as $\delta\phi = i\theta[Q,\phi]$. Alternatively, the $U(1)$ transformation of ϕ can be calculated as

$$\begin{aligned} e^{i\theta Q}\phi e^{-i\theta Q} &= \phi + i\theta[Q,\phi] + \cdots \\ &= \phi + i\theta\phi + \cdots \\ &= e^{i\theta}\phi\,. \end{aligned} \tag{14.18}$$

Notice that this is analogous to the expression (14.6). From (14.14) we can then generally state that

$$\langle\Omega|\widehat{\phi}|\Omega\rangle \neq 0 \quad\longrightarrow\quad e^{i\theta Q}|\Omega\rangle \neq |\Omega\rangle\,. \tag{14.19}$$

Thus the original $U(1)$ symmetry of $\mathcal{H}$ (for $\sigma < 0$) is spontaneously broken.

Goldstone's theorem

Goldstone's theorem states that

> If a continuous symmetry is spontaneously broken, there exist states for which the energy $\omega_p \to 0$ as the momentum $|\vec{p}| \to 0$.

The states correspond to massless particles, the so-called Goldstone bosons, in relativistic theories. The theorem also applies to non-relativistic phenomena in which case the states correspond to field excitations, known as the Goldstone modes.

The dynamics of Goldstone modes is derived purely by group theory. In the following, we shall see this in the present example of complex scalar field. Excitations of ϕ can be implemented as $\phi = \sqrt{\frac{|\sigma|}{2\lambda}}e^{i\alpha} + \eta$ where η is a complex scalar field, satisfying $\langle\Omega|\eta|\Omega\rangle = 0$. In a slightly different way, we can also parametrize ϕ as

$$\phi(x) \;=\; e^{i\chi(x)}\left(\sqrt{\frac{|\sigma|}{2\lambda}}e^{i\alpha} + \rho(x)\right) \tag{14.20}$$

where χ, ρ are real scalar fields. With this parametrization ϕ is expressed as

$$\phi \;\approx\; \phi_0 + i\chi\phi_0 + \rho + \cdots \tag{14.21}$$

where $\phi_0 = \sqrt{\frac{|\sigma|}{2\lambda}}e^{i\alpha}$. Thus, up to linear order, we can also use the parametrization (14.20) with $\langle\Omega|\chi|\Omega\rangle = \langle\Omega|\rho|\Omega\rangle = 0$. $\dot{\phi}$ and $\nabla\phi$ are calculated as

$$\dot{\phi} \;=\; e^{i\chi}\left[\dot{\rho} + i\dot{\chi}(\phi_0+\rho)\right]\,, \tag{14.22}$$

$$\nabla\phi \;=\; e^{i\chi}\left[\nabla\rho + i\nabla\chi(\phi_0+\rho)\right]\,. \tag{14.23}$$

Using these, we can rewrite the Hamiltonian (14.12) as

$$\mathcal{H} \;=\; \int d^3x\Big[\dot{\rho}^2 + \dot{\chi}^2(\phi_0+\rho)^2 + (\nabla\rho)^2 + (\nabla\chi)^2(\phi_0+\rho)^2$$

$$
\begin{aligned}
&+\lambda\left((\phi_0+\rho)(\phi_0^*+\rho)-\frac{|\sigma|^2}{2\lambda}\right)-\frac{|\sigma|^2}{4\lambda}\Bigg] \\
= &\int d^3x\left[\dot{\rho}^2+(\nabla\rho)^2+m^2\rho^2+\phi_0^*\phi_0(\dot{\chi}^2+(\nabla\chi)^2)\right] \\
&+(\text{interaction terms})
\end{aligned}
\tag{14.24}
$$

where $m^2=\lambda(\phi_0+\phi_0^*)^2$. We also neglect the constant term. From this expansion, we find that ρ is a massive field of mass m and χ is a massless field. Notice that the potential term is independent of χ and from its form we see that ρ is not an excited field, while χ can be regarded as an excited mode, the Goldstone mode, since there is a freedom to scale $\sqrt{\frac{|\sigma|}{2\lambda}}$. The emergence of χ is therefore a natural consequence of the spontaneous symmetry breaking of $U(1)$ symmetry $\phi\to e^{i\theta}\phi$ as it has a non-zero vacuum expectation value.

Spontaneous breaking of $O(N)$ symmetry

We now consider another example of spontaneous symmetry breaking. The theory is defined by an N-dimensional real vector field $\phi=(\phi_1,\phi_2,\cdots,\phi_N)^T$ with a Hamiltonian

$$
\mathcal{H}=\int d^3x\left[\frac{1}{2}\dot{\phi}_i\dot{\phi}_i+\frac{1}{2}(\nabla\phi)(\nabla\phi)+\lambda(\phi_i\phi_i)^2+\frac{\sigma}{2}(\phi_i\phi_i)\right] \tag{14.25}
$$

where $i=1,2,\cdots,N$. Obviously, the Hamiltonian has a symmetry

$$
\phi_i\ \longrightarrow\ R_{ij}\phi_j \tag{14.26}
$$

where

$$
R_{ij}R_{ik}\ =\ \delta_{jk}\,. \tag{14.27}
$$

Namely, $(R^TR)_{jk}=\delta_{jk}=\mathbf{1}$ where $\mathbf{1}$ is the $N\times N$ identity matrix. This means that R denotes an operator of $O(N)$ symmetry.

As in the previous example, we need to require $\lambda>0$ so that $\mathcal{H}$ is bounded below. Thus there are two cases depending on the sign of σ:

1. For $\sigma>0$, the ground state is given by $\phi_i=0$. The vacuum expectation value vanishes $\langle\Omega|\phi_i|\Omega\rangle=0$ and no spontaneous symmetry breaking occurs.
2. For $\sigma<0$, the potential term is given by $\lambda\left(\phi_i\phi_i-\frac{|\sigma|}{4\lambda}\right)-\frac{|\sigma|^2}{16\lambda}$. Classically, the vacuum is given by $\phi_i\phi_i=\phi_1^2+\phi_2^2+\cdots+\phi_N^2=\frac{|\sigma|}{4\lambda}$.

In the latter case, one possibility to the classical solution is given by

$$
\phi^{(0)}\ \equiv\ \langle\Omega|\phi|\Omega\rangle\ =\ \left(0,\cdots,0,\sqrt{\frac{|\sigma|}{4\lambda}}\right)^T\,. \tag{14.28}
$$

In terms of the elements, this is written as $\phi_1^{(0)} = \phi_2^{(0)} = \cdots = \phi_{N-1}^{(0)} = 0$ and $\phi_N^{(0)} = \sqrt{\frac{|\sigma|}{4\lambda}}$. Notice that $\phi^{(0)}$ is invariant under transformations of an $O(N-1)$ subgroup of the full symmetry group $O(N)$. Let H be the subgroup of G, $H \subset G$, where $H = O(N-1)$ and $G = O(N)$. H is called an isotropic group or a little group of $\phi^{(0)}$ The subgroup H is not broken. This can be seen by considering the expectation value

$$\begin{aligned} \langle\Omega|h^{-1}\phi_i h|\Omega\rangle &= R_{ij}(h)\langle\Omega|\phi_j|\Omega\rangle \\ &= R_{ij}(h)\phi_j^{(0)} = \phi_i^{(0)} = \langle\Omega|\phi_i|\Omega\rangle \end{aligned} \tag{14.29}$$

where $h \in H$. This relation leads to $h|\Omega\rangle = |\Omega\rangle$. Thus the H subgroup is *not* spontaneously broken.

Let T^A be the elements of $\underline{G} = \underline{O(N)}$ with $A = 1, 2, \cdots, N(N-1)/2$. The rotational transformations are carried out by the action of operator $R = \exp(iT^A\theta^A)$. The generators T^A are split into $t^a \in \underline{H}$ where $a = 1, 2, \cdots, (N-1)(N-2)/2$ and $S^\alpha \in \underline{G} - \underline{H}$ where $\alpha = 1, 2, \cdots, N-1$. In general, the number of independent S^α's is given by $\dim G - \dim H$. In terms of t^a an infinitesimal form of $R_{ij}(h)\phi_j^{(0)} = \phi_i^{(0)}$ is written as $(\mathbf{1} + it^a\theta^a)_{ij}\phi_j^{(0)} = \phi_i^{(0)}$. This means that actions of t^a and S^α on $\phi^{(0)}$ are given by

$$t^a\,\phi^{(0)} = 0\,, \quad S^\alpha\,\phi^{(0)} \neq 0\,. \tag{14.30}$$

In the following, we consider Goldstone's theorem in terms of these generators.

Minimization of the Hamiltonian (14.25), which is necessary to obtain the ground state, is given by $\frac{\partial}{\partial\phi_i}V(\phi_i\phi_i) = 0$ where $V(\phi_i\phi_i)$ denotes the potential terms

$$V(\phi_i\phi_i) = \lambda(\phi_i\phi_i)^2 + \frac{\sigma}{2}(\phi_i\phi_i)\,. \tag{14.31}$$

The $O(N)$ symmetry transformations are written as

$$\phi_i \longrightarrow \phi_i' = R_{ij}\phi_i \approx (\delta_{ij} + i\theta^A T_{ij}^A)\phi_j\,. \tag{14.32}$$

Namely, we have $\delta\phi_i = i\theta^A T_{ij}^A\phi_j$ infinitesimally. Invariance of V under the symmetry can be expressed as $\delta V = \frac{\partial V}{\partial\phi_i}\delta\phi_i = 0$, *i.e.*,

$$i\theta^A\frac{\partial V}{\partial\phi_i}T_{ij}^A\phi_j = 0\,. \tag{14.33}$$

Taking a derivative of it with respect to ϕ_k, we then have

$$\frac{\partial^2 V}{\partial\phi_i\partial\phi_k}T_{ij}^A\phi_j + \frac{\partial V}{\partial\phi_i}T_{ik}^A = 0\,. \tag{14.34}$$

Now, setting all ϕ's to be the solution of $\frac{\partial V}{\partial\phi_i} = 0$, this leads to

$$\left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_k}\right)_{\phi^{(0)}} T_{ij}^A\phi_j^{(0)} = 0\,. \tag{14.35}$$

Fluctuations around $\phi^{(0)}$ can be analyzed by substituting $\phi_i = \phi_i^{(0)} + \eta_i$ into the Hamiltonian:

$$\begin{aligned}\mathcal{H} &= \int d^3x \left[\dot{\phi}^2 + (\nabla\phi)^2 + V(\phi_i\phi_i)\right] \\ &= \dot{\eta}^2 + (\nabla\eta)^2 + V(\phi_i^{(0)}\phi_i^{(0)}) \\ &\quad + \underbrace{\left(\frac{\partial V}{\partial\phi_i}\right)_{\phi^{(0)}}}_{=0}\eta_i + \frac{1}{2}\underbrace{\left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_k}\right)_{\phi^{(0)}}}_{\equiv\mathcal{M}_{ik}}\eta_i\eta_k + \mathcal{O}(\eta^3)\,. \end{aligned} \tag{14.36}$$

The eigenvalues of $\mathcal{M}_{ik}$ give the masses or gaps of the η-modes. In terms of $\mathcal{M}_{ik}$, the relation (14.35) is written as

$$\mathcal{M}_{ki}\,\xi_i^A = 0 \tag{14.37}$$

where $\xi_i^A = T_{ij}^A\phi_j^{(0)}$. Thus, from (14.30) we find

$$T^A = \begin{cases} t^a: & t^a\phi^{(0)} = 0 \;\rightarrow\; \xi_i^a = 0 \\ S^\alpha: & S^\alpha\phi^{(0)} \neq 0 \;\rightarrow\; \xi_i^a \neq 0 \end{cases} \tag{14.38}$$

The second case corresponds to a zero mode of $\mathcal{M}$. To summarize, the result can be stated as follows:

> For every generator S^α which corresponds to the broken symmetry ($S^\alpha\phi^{(0)} \neq 0$), there is a zero mode (zero mass and zero gap) for η's.

This is a general statement of Goldstone's theorem since this holds for a general symmetry group G and an isotropy group H. For the completion of this statement, we need to show linear independence of ξ^α's. If ξ^α's are linearly dependent, then there exists a non-null vector V^α that satisfies $\xi^\alpha V^\alpha = 0$. This means that the matrix $N^{\alpha\beta} = \xi_i^\alpha\xi_i^\beta$ obeys $N^{\alpha\beta}V^\beta = 0$, *i.e.*, $\det N = 0$. Since $N^{\alpha\beta}$ is symmetric, it can be diagonalized by an orthogonal matrix. If any of the diagonal elements vanishes, then we can have a larger isotropy subgroup for the vacuum expectation value $\phi^{(0)}$. Thus the condition $\det N = 0$ is not allowed and we can conclude the linear independence of ξ^α's.

14.2 Dynamics of Goldstone bosons

To continue the above analysis, we first consider the $O(N)$ theory

$$\mathcal{H} = \int d^3x \left[\dot{\phi}^2 + (\nabla\phi)^2 + V(\phi_i\phi_i)\right] \tag{14.39}$$

where the potential term is given by (14.31). Fluctuations around $\phi^{(0)}$ in (14.14) are parametrized as

$$\phi_i = R_{ij}(x)\chi_j\,, \tag{14.40}$$

$$\chi \;=\; (0,\cdots,0,v+\rho)^T \tag{14.41}$$

where $v = \phi_N^{(0)} = \sqrt{\frac{|\sigma|}{4\lambda}}$ and $R_{ij}(x)$ denotes a spacetime-dependent rotation matrix. Since the isotropy group of $\phi^{(0)}$ is $O(N-1)$, the fluctuation (14.40) is invariant under $O(N-1)$-part rotation of R_{ij}. In other words, the number of degrees of freedom in R_{ij} is given by $\dim\left[\frac{O(N)}{O(N-1)}\right] = N-1$. Note that the remaining freedom is encoded by ρ in the parametrization (14.41).

Form (14.40) we obtain

$$\begin{aligned}
\phi^2 &= \phi_i\phi_i \,=\, (v+\rho)^2 && (14.42)\\
\dot\phi &= \dot R\chi + R\dot\chi \,=\, R(R^{-1}\dot R\chi + \dot\chi) && (14.43)\\
\dot\phi^2 &= \dot\phi_i\dot\phi_i \,=\, (R^{-1}\dot R\chi+\dot\chi)_i(R^{-1}\dot R\chi+\dot\chi)_i\\
&= (R^{-1}\dot R)_{ij}\chi_j(R^{-1}\dot R)_{ik}\chi_k + \dot\chi_i(R^{-1}\dot R)_{ij}\chi_j + (R^{-1}\dot R)_{ij}\chi_j\dot\chi_i + \dot\chi_i\dot\chi_i\\
&= \chi_j^T(R^{-1}\dot R)^T_{ji}(R^{-1}\dot R)_{ik}\chi_k + 2\dot\rho(R^{-1}\dot R)_{NN}(v+\rho) + \dot\rho^2\\
&= -\chi^T(R^{-1}\dot R R^{-1}\dot R)\chi && (14.44)
\end{aligned}$$

where we use the fact that $R^{-1}\dot R$ is antisymmetric, *i.e.*,

$$(R^{-1}\dot R)^T \,=\, \dot R^T R^{-1\,T} \,=\, \dot R^{-1}R \,=\, -R^{-1}\dot R\,. \tag{14.45}$$

Notice that (14.45) automatically leads to $(R^{-1}\dot R)_{NN} = 0$. Substituting (14.40) into (14.39), we can express the Hamiltonian as

$$\begin{aligned}
\mathcal{H} \;=\; \int d^3x\Bigg[\frac{1}{2}\left[\dot\rho^2 - \chi^T(R^{-1}\dot R R^{-1}\dot R)\chi + (\nabla\rho)^2 - \chi^T(R^{-1}\nabla R\, R^{-1}\nabla R)\chi\right]\\
+ V((v+\rho)^2)\Bigg]\,. \qquad (14.46)
\end{aligned}$$

The potential term depends on ρ and generally gives mass term or a gap for this field. On the other hand, R has no gap and represents $(N-1)$ Goldstone bosons. Dynamics of these bosons is given by

$$\mathcal{H}_R \,=\, -\frac{1}{2}\int d^3x\,\chi^T\left[R^{-1}\dot R R^{-1}\dot R + R^{-1}\nabla R\, R^{-1}\nabla R\right]\chi\,. \tag{14.47}$$

Since χ is parametrized by (14.41), it produces interaction terms between R and ρ. At low energy $E \ll m_\rho$ (mass or gap of ρ), one can neglect ρ and interaction of ρ with Goldstone bosons. Effective low-energy dynamics of the Goldstone bosons is then described by the effective Hamiltonian

$$\begin{aligned}
\mathcal{H}_{\rm eff} &= -\frac{1}{2}v^2\int d^3x\left[(R^{-1}\dot R R^{-1}\dot R)_{NN} + (R^{-1}\nabla R\, R^{-1}\nabla R)_{NN}\right]\\
&= -\frac{v^2}{2}\int d^3x\,{\rm Tr}\left[\Big((R^{-1}\dot R)^2 + (R^{-1}\nabla R)^2\Big)W\right] && (14.48)
\end{aligned}$$

where $R \in O(N)$ and

$$W = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{14.49}$$

Neglecting ρ-terms is classically true but it is not generally correct in quantum mechanics. A nontrivial effect of ρ comes in at quantum level. For example, we need to take account of ρ-field vacuum fluctuation in the case of $[\rho, \dot{\rho}] \neq 0$. Nevertheless, there is a theorem regarding the dynamics of Goldstone bosons:

> The $\mathcal{H}_{\text{eff}}$ in (14.48) is the low-energy effective Hamiltonian including all quantum interactions.

Notice that the above mentioned quantum corrections can be taken care of by redefining (or renormalizing) the vacuum value v.

The effective Hamiltonian $\mathcal{H}_{\text{eff}}$ in (14.48) can be simplified as

$$\begin{aligned} \mathcal{H}_{\text{eff}} &= -\frac{1}{2}v^2 \int d^3x \left[(R^{-1}\dot{R}R^{-1}\dot{R})_{NN} + (R^{-1}\nabla R\, R^{-1}\nabla R)_{NN} \right] \\ &= -\frac{1}{2}v^2 \int d^3x \left[(R^{-1}\dot{R}R^{-1}\dot{R})_{NN} - \nabla(R^{-1}\nabla R)_{NN} + (R^{-1}\nabla^2 R)_{NN} \right] \\ &\approx -\frac{v^2}{2} \int d^3x \, \mathrm{Tr}(R^{-1}\dot{R}R^{-1}\dot{R})W \end{aligned} \tag{14.50}$$

where we neglect the term $R^{-1}\nabla^2 R \sim \vec{k}^2$ for the low energy theory. Let R_{N-1} be an $(N-1)\times(N-1)$ matrix element of the isotropy group $H = O(N-1)$. In terms of R_{N-1} we can define an $N \times N$ matrix element $h \in H$ as

$$h = \begin{pmatrix} & & 0 \\ & R_{N-1} & \vdots \\ & & 0 \\ 0 & \cdots 0 & 1 \end{pmatrix}. \tag{14.51}$$

Under the h-transformations, $R \to R' = Rh$, we have

$$\dot{R}' = \dot{R}h + R\dot{h} = R(R^{-1}\dot{R}h + \dot{h}), \tag{14.52}$$

$$R'^{-1}\dot{R}' = h^{-1}R^{-1}(\dot{R}h + R\dot{h}) = h^{-1}R^{-1}\dot{R}h + h^{-1}\dot{h}. \tag{14.53}$$

Thus we can explicitly show the invariance of the Hamiltonian (14.50) under the $O(N-1)$ rotation:

$$\begin{aligned} \mathrm{Tr}R'^{-1}\dot{R}'R'^{-1}\dot{R}'W &= \mathrm{Tr}h^{-1}R^{-1}\dot{R}hh^{-1}R^{-1}\dot{R}hW + \mathrm{Tr}h^{-1}\dot{h}h^{-1}R^{-1}\dot{R}hW \\ &\quad + \mathrm{Tr}h^{-1}R^{-1}\dot{R}hh^{-1}\dot{h}W + \mathrm{Tr}h^{-1}\dot{h}h^{-1}\dot{h}W \end{aligned}$$

$$
\begin{aligned}
&= \mathrm{Tr} R^{-1}\dot{R}R^{-1}\dot{R}hWh^{-1} + \mathrm{Tr} h^{-1}R^{-1}\dot{R}hWh^{-1}\dot{h} \\
&= \mathrm{Tr} R^{-1}\dot{R}R^{-1}\dot{R}W \qquad (14.54)
\end{aligned}
$$

where we use $\dot{h}W = 0$ and $hWh^{-1} = W$. The Hamiltonian (14.50) is then written as $\mathcal{H}_{\rm eff} = \int \epsilon(R) d^3x$ where $\epsilon(R) = -\frac{v^2}{2}\mathrm{Tr} R^{-1}\dot{R}R^{-1}\dot{R}W$, satisfying $\epsilon(R) = \epsilon(Rh)$. Suitable geometry relevant to the low energy dynamics of Goldstone bosons is therefore given by a coset space G/H. We shall consider such a geometric notion of spontaneous symmetry breaking in the following.

Coset manifolds

Consider a compact Lie group G and a subgroup $H \subset G$. As discussed in Chapter 12, G can be interpreted as a nice differentiable Riemannian manifold. Every element of the group, $g \in G$, corresponds to a point of the manifold. Now, we identify a point g with gh where $h \in H$. In other words, we impose equivalence $g \sim gh$ on the manifold G. This leads to a smaller manifold G/H, called a coset manifold. The dimension of G/H is given by

$$
\dim G/H \,=\, \dim G - \dim H \,. \qquad (14.55)
$$

Using the previous notation (14.38) for generators, we find that the group elements are expressed as $g \approx 1 + iT^A\theta^A$, $h^{-1} \approx 1 - it^\alpha\theta^\alpha$ and $gh^{-1} \approx 1 + iT^A\theta^A - it^\alpha\theta^\alpha$. Thus, infinitesimally (near the origin at $\theta^A = 0$), the dimension of the coset space is given by $\dim(T^A - t^\alpha)$ and, hence, by (14.55).

Functions on the G/H space must obey $f(gh) = f(g)$. In other words, functions on G/H can be obtained as H-invariant functions on G. As discussed in Section 13.1, thanks to the Peter-Weyl theorem, these functions form a complete set. This is a well-known result in representation theory of mathematics, based on the works of Harish-Chandra, Gelfand, Naimark and others.

The Cartank-Killing metric on G is defined in Section 12.2. As shown in (12.23) and (12.24), the metric ds^2 is defined as

$$
\begin{aligned}
g^{-1}dg &= iT^A E_i^A(\theta) d\theta^i \,, \qquad (14.56) \\
ds^2 &= -2\mathrm{Tr}(g^{-1}dg\, g^{-1}dg) \,=\, E_i^A E_j^A d\theta^i d\theta^j \qquad (14.57)
\end{aligned}
$$

where we use the normalization $\mathrm{Tr}(t^A t^B) = \frac{1}{2}\delta^{AB}$. The metric tensor on G is defined by $g_{ij} = E_i^A E_j^A$. In terms of the generators $t^a \in \underline{H}$ and $S^\alpha \in \underline{G} - \underline{H}$ in (14.38), the frame field one-form $g^{-1}dg$ is written as

$$
g^{-1}dg \,=\, it^a E_i^a d\theta^i + iS^\alpha E_i^\alpha d\theta^i \,. \qquad (14.58)
$$

Using $\mathrm{Tr}(S^\alpha S^\beta) = \frac{1}{2}\delta^{\alpha\beta}$, we can express the the frame field relevant to the G/H space as

$$
E_i^\alpha(g) d\theta^i \,= -i2\mathrm{Tr}(S^\alpha g^{-1}dg) \,. \qquad (14.59)
$$

Now, under the transformation $g \to gh$, we have

$$
g^{-1}dg \;\longrightarrow\; h^{-1}g^{-1}d(gh) \;=\; h^{-1}g^{-1}dg\, h \,+\, h^{-1}dh \,, \qquad (14.60)
$$

$$E_i^\alpha(g)d\theta^i \longrightarrow E_i^\alpha(gh)d\theta^i = -i2\mathrm{Tr}(S^\alpha h^{-1}g^{-1}dg\,h) - i2\underbrace{\mathrm{Tr}(S^\alpha h^{-1}dh)}_{\mathrm{Tr}(S^\alpha t^a)=0}$$
$$= -i2\mathcal{D}^{\alpha\beta}(h)\mathrm{Tr}(S^\beta g^{-1}dg)$$
$$= \mathcal{D}^{\alpha\beta}(h)\,E_i^\beta(g)d\theta^i \tag{14.61}$$

where $\mathcal{D}^{\alpha\beta}(h)$ denotes the adjoint representation of S^α,

$$hS^\alpha h^{-1} = \mathcal{D}^{\alpha\beta}(h)\,S^\beta\,. \tag{14.62}$$

Notice that $\mathcal{D}^{\alpha\beta}(h)$ has orthogonality $\mathcal{D}^{\alpha\beta}(h)\mathcal{D}^{\alpha\gamma}(h) = \delta^{\beta\gamma}$. Therefore the metric on G/H is defined as

$$ds^2 = E_i^\alpha(g)E_j^\alpha(g)\,d\theta^i d\theta^j$$
$$= E_i^\alpha(gh)E_j^\alpha(gh)\,d\theta^i d\theta^j\,. \tag{14.63}$$

Note that for unitary groups the orthogonality relation should be written as $\mathcal{D}^{*\alpha\beta}(h)\mathcal{D}^{\alpha\gamma}(h) = \delta^{\beta\gamma}$ with $(h^*)^T = h^\dagger = h^{-1}$. Accordingly, the product of $E_i^\alpha E_j^\alpha$ should be interpreted as $E_i^{*\alpha}E_j^\alpha$ in such cases.

For $G = SU(2)$ and $H = U(1)$, the coset manifold becomes two-sphere S^2. As explicitly shown in (12.43), the metric on $SU(2)/U(1)$ can naturally be extracted from the Cartan-Killing metric on $SU(2)$. To recapitulate, the metric on $G = SU(2)$ is parametrized as

$$g = \frac{1}{\sqrt{1+z\bar{z}}}\begin{pmatrix} 1 & z \\ -\bar{z} & 1 \end{pmatrix}\begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}, \tag{14.64}$$

$$ds^2_{SU(2)} = -2\mathrm{Tr}(g^{-1}dg\,g^{-1}dg)$$
$$= 4\frac{dzd\bar{z}}{(1+z\bar{z})^2} - \left(\frac{zd\bar{z}-\bar{z}dz}{1+z\bar{z}} + id\theta\right)^2 \tag{14.65}$$

where the first term in (14.65) corresponds to the metric on $S^2 = SU(2)/U(1)$:

$$ds^2_{\frac{SU(2)}{U(1)}} = 4\frac{dzd\bar{z}}{(1+z\bar{z})^2}\,. \tag{14.66}$$

Going back to the effective Hamiltonian (14.50), $\mathcal{H}_{\mathrm{eff}} \sim \mathrm{Tr}(R^{-1}\dot{R}R^{-1}\dot{R}W)$, we find that it is proportional to the metric on G/H. Notice that in this case we have $G = O(N)$ and $H = O(N-1)$. Generally, the metric on G/H is defined as

$$ds^2 = G_{AB}(\theta)\,d\theta^A d\theta^B \tag{14.67}$$

where $G_{AB}(\theta)$ is the metric tensor on G/H as a function of a real parameter θ^A $(A = 1, 2, \cdots, \mathrm{dim}G - \mathrm{dim}H)$. There exists a theorem regarding low energy dynamics of Goldstone bosons. In terms of the above notations, it can be stated as follows:

If a continuous symmetry G is spontaneously broken down to a subgroup $H \subset G$ (a little group of $\langle\Omega|\phi|\Omega\rangle$ is H), then

1. there exist $\dim G - \dim H$ Goldstone bosons (with no gap);
2. low-energy dynamics of Goldstone modes is given by an effective Lagrangian

$$\begin{aligned}\mathcal{L}_{\rm eff} &= \frac{v^2}{2}\left(G_{AB}(\theta)\dot{\theta}^A\dot{\theta}^B - G_{AB}(\theta)\nabla\theta^A\nabla\theta^B\right) \\ &= \frac{v^2}{2}\,G_{AB}(\theta)\,\partial_\mu\theta^A\partial_\mu\theta^A \qquad (14.68)\end{aligned}$$

where $G_{AB}(\theta)d\theta^A d\theta^B$ is the metric on the coset space G/H.

Example 1: Heisenberg ferromanget

As discussed in the beginning of this chapter, the simplest example of spontaneous symmetry breaking is given by the Heisenberg ferromanget. In this model, there exists a transition temperature T_c below which spontaneous symmetry breaking of $O(3) \to O(2)$ occurs so that nonzero net magnetization arises below T_c. The relevant coset space is $O(3)/O(2) = S^2$ and there emerge two Goldstone modes, also known as spin waves in this particular case.

As shown in (14.66), the S^2 metric is given by $4\frac{dzd\bar{z}}{(1+z\bar{z})^2}$. Spin wave dynamics is then described by an effective Hamiltonian

$$\mathcal{H}_{\rm eff} \,=\, \alpha\int d^3x\,\left(\frac{\dot{Z}\dot{\bar{Z}}}{(1+Z\bar{Z})^2} + \frac{\nabla Z\nabla\bar{Z}}{(1+Z\bar{Z})^2}\right) \qquad (14.69)$$

where α is some constant and the complex scalar field Z denotes spin waves. The corresponding Lagrangian is given by

$$\mathcal{L}_{\rm eff} \,=\, \alpha\,\frac{\partial_\mu z\,\partial_\mu\bar{z}}{(1+z\bar{z})^2}\,. \qquad (14.70)$$

Using $\mathcal{H}_{\rm eff}$ or $\mathcal{L}_{\rm eff}$, we can carry out relativistic calculations of low energy spin-wave dynamics.

Example 2: Chiral symmetry breaking of strong interactions

The $SU(3)_L \times SU(3)_R$ chiral symmetry of quantum chromodynamics (QCD) is spontaneously broken to $SU(3)_V$ by the effects of strong or gluonic interactions.

$$G = SU_L(3) \times SU_R(3) \;\longrightarrow\; SU_V(3) \qquad (14.71)$$

The coset space of interest is $G/H = SU(3)$ and there are eight Goldstone bosons. These correspond to the pseudoscalar octet mesons $\pi^\pm$, π^0, K^0, $\overline{K^0}$,

K^+, K^- and η. As discussed earlier, the initial symmetry G is imperfect due to mass differences among light quarks. Thus these mesons are not true Goldstone bosons. They are in fact massive and are called pseudo-Goldstone bosons. The actual masses are listed in Table 5.2. Notice that these masses are well below the strong interaction scale, approximately 1 GeV. In terms of the hierarchy of interactions which we have discussed in (13.41), we can therefore interpret these mesons massless in the region where electroweak interactions are negligible. In other words, in the limit of no electroweak effect where the electroweak couplings and the quark masses are taken to zero, the pseudo-Goldstone bosons should become the true Goldstone bosons.

The (pseudo) Goldstone particles can be represented by U, an element of $SU(3)$. Let (g_L, g_R) be an element of the chiral symmetry group $SU_L(3) \times SU_R(3)$. Then the chiral transformations of U are given by

$$U \to U' \,=\, g_L U g_R^\dagger \,. \tag{14.72}$$

Choosing the vacuum as $\langle\Omega|U|\Omega\rangle \,=\, 1$ shows that $g_L \,=\, g_R$ is the isotropy group of the vacuum. This gives $H \,=\, SU_V(3)$, the vector subgroup of $G \,=\, SU_L(3)\times SU_R(3)$. Thus, from (14.68), we can construct the effective Lagrangian for the pseudoscalar octet mesons as

$$\mathcal{L}_{\rm eff} \,=\, f_\pi^2 \,{\rm Tr}(\partial_\mu U^\dagger \partial_\mu U) \tag{14.73}$$

where f_π is some constant.

14.3 Sigma models and patterns of $G \to H$

Let $x^\mu(t)$ be a trajectory of a point particle on a Riemannian manifold $\mathcal{M}$. Namely, it provides a mapping $x^\mu(t) : \mathbf{R} \to \mathcal{M}$. The geodesic equation is given by the extremization of the action

$$\mathcal{S} \,=\, \frac{1}{2}\int dt\, g_{\mu\nu}\frac{\partial x^\mu}{\partial t}\frac{\partial x^\nu}{\partial t} \tag{14.74}$$

where $g_{\mu\nu}$ denotes the metric tensor of $\mathcal{M}$. Now consider a mapping from a spacetime to a coset space,

$$\phi^A(x) \,:\, \mathbf{R}^4 \,\longrightarrow\, G/H \,. \tag{14.75}$$

The corresponding action that is analogous to the geodesic version (14.74) can be written as

$$\mathcal{S} \,=\, \frac{1}{2}\int d^4x\, G_{AB}\partial_\mu\phi^A\partial_\mu\phi^B \tag{14.76}$$

where $G_{AB}\partial_\mu\phi^A\partial_\mu\phi^B$ denotes the metric of G/H. This action is proportional to the effective action for Goldstone bosons (14.68).

Sigma models

To generalize, the action (14.76) can be expressed as

$$\mathcal{S} = \frac{1}{2}\int d^4x\, G_{AB}\partial_a\phi^A\partial_b\phi^B\,\sqrt{-\det g}\, g^{ab} \tag{14.77}$$

where $\phi^A(x)$ gives a mapping from a Riemannian manifold $\mathcal{N}$ with metric g^{ab} to another Riemannian manifold $\mathcal{M}$ with G^{AB}

$$\phi^A(x) : \mathcal{N} \longrightarrow \mathcal{M}\,. \tag{14.78}$$

The action (14.77) defines a sigma model on $\mathcal{M}$. Classical solutions to the sigma model, which are obtained by extremization of (14.77), are called Harmonic maps. $\mathcal{N}$ and $\mathcal{M}$ are called a base space and a target space of the sigma model, respectively. As mentioned above, the action (14.76) is proportional to the effective action for Goldstone bosons. Thus we can conclude that *low energy dynamics of Goldstone bosons is given by a sigma model with G/H as the target space.*

Patterns of $G \to H$

We now consider how to enforce a required pattern of $G \to H$. Once such a pattern is determined, we can construct Hamiltonian $\mathcal{H}$ (or Lagrangian $\mathcal{L}$) of the spontaneous symmetry breaking, satisfying $[\mathcal{H}, Q^a] = 0$ (or $[\mathcal{L}, Q^a] = 0$) where Q^a is a generator of the symmetry group G. As we have seen so far, our strategy is to start with a vector field ϕ which is in the irreducible representation of a Lie group G. We then consider a nonzero vacuum expectation value $\langle\Omega|\phi|\Omega\rangle = \phi_0 \neq 0$ such that an isotropy group of ϕ_0 is the Lie subgroup $H \subset G$. Namely, ϕ_0 is invariant under h-transformations ($h \in H$). This means that ϕ_0 is a singlet under h. Thus, group theoretically, the required pattern of $G \to H$ is given by reducing representations of G into those of H.

For example, consider the simple case $G/H = SU(2)/U(1)$. In this case, the h-transformations corresponds to the σ_3 action on ϕ where σ_i denotes the 2×2 Pauli matrices ($i = 1, 2, 3$). In the real **2** representation (or the ordinary spin-$\frac{1}{2}$ representation), there are no singlets under the action of

$$h = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \tag{14.79}$$

unless $\theta = 0$. The first nontrivial case is given by the **3** representation or the spin-1 representation. Under the σ_3 action, this decomposes as $\mathbf{3} \to \mathbf{2}\oplus\mathbf{1}$ so that the vacuum expectation value can be chosen as $\phi_0^1 = \phi_0^2 = 0$ and $\phi_0^3 = v \neq 0$. The potential term which brings about the spontaneous symmetry breaking is then given by $V = \lambda(\phi^a\phi^a - v^2)^2$ with $\lambda > 0$ and $a = 1, 2, 3$. The potential V corresponds to the simplest case of the $O(N)$ Hamiltonian considered in (14.25). Requirement of V being invariant under $G = SU(2) \simeq SO(3)$ leads to $\phi^a\phi^a = v^2$ where $\phi^a = R_{ab}\phi_0^b$, with R_{ab} denoting an $SO(3)$ rotation operator.

From the condition $\phi^2 = v^2$, we find that the coset $SU(2)/U(1)$ is geometrically equivalent to the two-sphere S^2.

Secondly, consider the case of $G/H = SU(3)/SU(2)$. Goldstone bosons are denoted by complex fields ϕ^a which are in an irreducible representation of $SU(3)$. The **3** representation of $SU(3)$ splits into representations of $SU(2)$ as

$$\begin{array}{ccc} \mathbf{3} & = & \mathbf{2} \oplus \mathbf{1}\,. \\ SU(3) & & SU(2) \end{array} \tag{14.80}$$

Since it includes the singlet **1** of $SU(2)$, we can take $V = \lambda(\phi^{*a}\phi^a - v^2)^2$ $(a = 1, 2, 3)$ as a potential for the spontaneous symmetry breaking of $SU(3) \to SU(2)$. As before, the value of v is determined by the classical vacuum ϕ_0^a, *e.g.*, $\phi_0^1 = \phi_0^2 = 0$ and $\phi_0^3 = v$. This leads to the condition $\phi^{*a}\phi^a = v^2$, *i.e.*,

$$(\mathrm{Re}\phi)^a(\mathrm{Re}\phi)^a + (\mathrm{Im}\phi)^a(\mathrm{Im}\phi)^a \;=\; v^2 \tag{14.81}$$

for $a = 1, 2, 3$. Thus we find that the coset $SU(3)/U(2)$ is geometrically equivalent to the five-sphere S^5.

Lastly, we consider spontaneous symmetry breaking of

$$SU(3) \;\longrightarrow\; SU(2) \times U(1) \;\simeq\; U(2)\,. \tag{14.82}$$

In terms of the Gell-Mann matrices λ^a $(a = 1, 2, \cdots, 8)$ in (1.49), the $SU(1)$ and $U(1)$ part of the $SU(3)$ algebra are assigned by $(\lambda^1, \lambda^2, \lambda^3)$ and λ^8, respectively. As in the case of (14.80), the **3** representation of $SU(3)$ decomposes into representations of $SU(2) \times U(1)$ as

$$\begin{array}{ccc} \mathbf{3} & = & (\mathbf{2}, \frac{1}{2\sqrt{3}}) \;\oplus\; (\mathbf{1}, -\frac{1}{\sqrt{3}}) \\ SU(3) & & SU(2) \times U(1) \end{array} \tag{14.83}$$

where the $U(1)$ representation is labeled by the corresponding value of $\frac{\lambda^8}{2}$ which is known as the hypercharge in particle physics. Neither of the right-hand side terms in (14.83) is a singlet of $H = SU(2) \times U(1)$. In order to obtain a required pattern of spontaneous symmetry breaking (14.82), we need to decompose an $SU(3)$ irreducible representation such that it includes a singlet of $SU(2) \times U(1)$.

For this purpose, consider the relation

$$\begin{array}{ccc} \mathbf{3} \otimes \mathbf{3}^* & = & \mathbf{1} \oplus \mathbf{8} \\ SU(3) & & SU(3) \end{array} \tag{14.84}$$

where **1** denotes the singlet of $SU(3)$. We then consider the decomposition of the adjoint representation **8**. Using the tensor analysis we can split this representation as

$$T_b^a \;\longrightarrow\; T_\beta^\alpha + T_3^\alpha + T_\alpha^3 + T_3^3 \tag{14.85}$$

where $a, b = 1, 2, 3$ and $\alpha, \beta = 1, 2$. In terms of the $SU(2)$ representation this is written as

$$\begin{array}{ccc} \mathbf{8} & = & \mathbf{3} \oplus \mathbf{2}^* \oplus \mathbf{2} \oplus \mathbf{1}\,. \\ SU(3) & & SU(2) \end{array} \tag{14.86}$$

Including the hypercharge contributions, this leads to

$$\begin{array}{ccc} \mathbf{8} & = & (\mathbf{3},0)\oplus(\mathbf{2}^*,-\frac{3}{2\sqrt{3}})\oplus(\mathbf{2},\frac{3}{2\sqrt{3}})\oplus(\mathbf{1},0) \\ SU(3) & & SU(2)\times U(1) \end{array} \tag{14.87}$$

which contains the singlet of $H = SU(2) \times U(1)$. Thus, we can construct a set of Goldstone bosons ϕ^A $(A = 1, 2, \cdots, 8)$ with a potential $V = \lambda(\phi^A\phi^A - v^2)^2$ and $\phi_0^8 = v$, $\phi_0^a = 0$ $(a = 1, 2, \cdots, 7)$. The invariance of the potential under G leads to $\phi^A\phi^A = v^2$. Thus, as in the previous case, we can naively think of ϕ^A as coordinates of S^7. But this is not correct since the coset space of interest is four-dimensional. The singlet $(\mathbf{1}, 0)$ is not sensitive to the $SU(2)$ transformations. Upon these constraints we can properly consider the dimension of ϕ^A and, indeed, we find that ϕ^A correspond to coordinates of the coset space

$$\frac{SU(3)}{SU(2)\times U(1)} \simeq \frac{SU(3)}{SU(2)}\Big/U(1) \simeq S^5/S^1 = \mathbf{CP}^2\,. \tag{14.88}$$

Complex projective spaces

To visualize a projective space, let us identify a line in $\mathbf{R}^3$ as the same. This means that a two-dimensional projective space can be considered as a set of all straight lines in $\mathbf{R}^3$. Namely, S^2 modulo antipolar point gives a real projective space $\mathbf{RP}^2$. Another way of saying this is that $\mathbf{RP}^2$ is defined by

$$x^a \sim \alpha x^a \quad (\alpha \in \mathbf{R},\, \alpha \neq 0) \tag{14.89}$$

where $a = 1, 2, 3$.

Complex version of the relation (14.89) is written as

$$z^a \sim \lambda z^a \quad (\lambda \in \mathbf{C},\, \lambda \neq 0) \tag{14.90}$$

where λ is a nonzero complex number. Real part of λ can be fixed by a choice of normalization

$$z^{*a}z^a = 1\,. \tag{14.91}$$

This means that the complex coordinate z^a and its conjugate define a five-sphere S^5. But z^a still has a phase degree of freedom, *i.e.*, $z^a \to e^{i\theta}z^a$. Thus, this leads to the coset space S^5/S^1. By definition, this is a complex projective space $\mathbf{CP}^2$,

$$S^5/S^1 = \mathbf{CP}^2\,. \tag{14.92}$$

In this particular case, it is also called the complex projective plane.

Generally, the complex projective space $\mathbf{CP}^N$ is defined by $N + 1$ non-zero complex variables z^a $(a = 1, 2, \cdots, N+1)$ which satisfy the projection condition (14.91). In terms of coset spaces, $\mathbf{CP}^N$ is expressed as

$$\mathbf{CP}^N = S^{2N+1}/S^1 \simeq \frac{SU(N+1)}{U(N)}\,. \tag{14.93}$$

Chapter 15

Solitons

15.1 The sine-Gordon solitons

Solitons are classical solutions of nonlinear equations. In this section, we consider $(1+1)$-dimensional solitons in the so-called sine-Gordon model. The Lagrangian of this model is given by

$$\mathcal{L} = \frac{1}{2}\left(\dot{\phi}^2 - \phi'^2\right) - \lambda(1-\cos\phi) \tag{15.1}$$

where $\dot{\phi} = \frac{\partial}{\partial t}\phi(t,x)$ and $\phi' = \frac{\partial}{\partial x}\phi(t,x)$. λ is a positive constant. The equation of motion then becomes

$$\Box\phi + \lambda\sin\phi = 0 \tag{15.2}$$

where $\Box = \partial_t^2 - \partial_x^2$. This is called the sine-Gordon equation. For $\lambda = 0$, this reduces to the massless Klein-Gordon equation. (The term "sine-Gordon" is named after this fact.) The corresponding Hamiltonian is given by

$$\mathcal{H} = \int dx \left(\frac{\dot{\phi}^2 + \phi'^2}{2} + \lambda(1-\cos\phi)\right). \tag{15.3}$$

Consider one class of solutions to the sine-Gordon equation (15.2). For small ϕ, we have $(\Box\phi + \lambda\phi) + \mathcal{O}(\phi^3) = 0$. Since $\mathcal{H}$ is positive, we have finite energy solutions by imposing the boundary conditions

$$\begin{aligned} \phi' &\to 0 \quad \text{as } x \to \pm\infty\,, \\ (1-\cos\phi) &\to 0 \quad \text{as } x \to \pm\infty\,. \end{aligned} \tag{15.4}$$

For such solutions, $\phi(\pm\infty) = 2\pi n$ $(n = 0, \pm1, \pm2, \cdots)$ is allowed. For example, we can fix the boundary by

$$\phi(-\infty) = 0\,, \quad \phi(\infty) = 2\pi\,. \tag{15.5}$$

Notice that classical evolution of ϕ is a smooth deformation of ϕ. Thus, if such a solution exists, classically this is absolutely stable.

The solution (15.5) in particular is called a kink. Under smooth deformations of ϕ, the kink can move but it never goes away. Namely, the kink configuration is conserved. The number of kinks is then defined by

$$Q = \frac{\phi(+\infty) - \phi(-\infty)}{2\pi} . \tag{15.6}$$

This is called the soliton number. The case of (15.5) corresponds to $Q = 1$. Since the soliton number is conserved, we can interpret it as a charge. Thus Q can also be expressed as

$$Q = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\partial\phi}{\partial x} dx \equiv \int J_0 \, dx \tag{15.7}$$

where $J_0 = \frac{1}{2\pi}\partial_x\phi$ denotes a charge density. The corresponding current vector is defined by

$$J_\mu = \frac{1}{2\pi}\epsilon_{\mu\nu}\partial_\nu\phi . \tag{15.8}$$

Conservation of the current can easily be checked as

$$\partial_\mu J_\mu = \frac{1}{2\pi}\epsilon_{\mu\nu}\partial_\mu\partial_\nu\phi = 0 . \tag{15.9}$$

These results are obtained without using the equations of motion. This is a peculiarity of solition solutions. Namely, we do not use symmetries but only topological constraints to show the conservation of Q. This is therefore a mathematical identity. For example, a Q-soliton solution ϕ in general always has a freedom to add an extra field χ, satisfying $\chi(-\infty) = \chi(+\infty) = 0$ because for $\widetilde{\phi} = \phi + \chi$ we can calculate the charge $\widetilde{Q}$ as

$$\begin{aligned} \widetilde{Q} &= \frac{1}{2\pi}\int dx(\partial_x\phi + \partial_x\chi) \\ &= Q + \frac{1}{2\pi}\int dx\partial_x\chi = Q . \end{aligned} \tag{15.10}$$

The essence of these soliton solutions is the fact that the $Q = 1$ solution (15.5) is allowed.

We now introduce the field

$$u(x) = \exp(i\phi) \tag{15.11}$$

with $u(-\infty) = u(\infty) = 1$. Using this parametrization, we have $\dot{\phi} = -iu^\dagger\dot{u}$, $\phi' = -iu^\dagger\partial_x u$. The Hamiltonian (15.3) is then written as

$$\mathcal{H} = \int dx \left[\frac{1}{2}\dot{u}^\dagger\dot{u} + \frac{1}{2}\partial_x u^\dagger\partial_x u + \lambda\left(1 - \frac{u + u^\dagger}{2}\right)\right] . \tag{15.12}$$

In terms of $u(x)$, the configuration of ϕ is described only by a circle. The energy density then becomes a function on an S^1-valued function,

$$u(x) \,:\, \mathbf{R} \longrightarrow S^1 \,. \tag{15.13}$$

To visualize Q-soliton solutions, we can naturally interpret $Q = \frac{1}{2\pi}\int dx \partial_x \phi$ as a winding number which counts how many times $\phi(x)$ winds around the S^1 as we go from $x = -\infty$ to $x = +\infty$. By definition, this Q is an integer and $\dot{Q} = 0$ identically. As long as $u(x)$ keeps the S^1 structure, coming back to $u(+\infty) = 1$, this result still holds, even with perturbation. Quantum mechanically, one must calculate an action for connecting different Q's. There are, however, no finite action configurations of this sort (in the Euclidean metric). Thus Q is preserved under quantum mechanical time evolution as well.

Properties of the current

$$J_\mu \,=\, \frac{1}{2\pi}\epsilon_{\mu\nu}\,\partial_\nu \phi \,=\, -\frac{i}{2\pi}\epsilon_{\mu\nu}\,u^\dagger \partial_\nu u \tag{15.14}$$

are summarized as follows.

1. $\partial_\mu J_\mu = 0$. This is an identity. We do not need to use the equations of motion.
2. $Q = \int J_0 dx$ is a winding number for the mapping $u(x) : \mathbf{R} \to S^1$.
3. Variation of J_μ is a total derivative. Q is therefore a topological invariant.

Existence of solution

Consider a set of all functions $u(x) : \mathbf{R} \to S^1$. This forms an ∞-dimensional space which consists of an infinite number of disconnected sectors, see Fig. 15.1.

In the following, we consider static solutions. The static Hamiltonian is given by

$$\mathcal{H} \,=\, \int dx \left(\frac{1}{2}(\partial_x \phi)^2 + \lambda(1-\cos\phi)\right) . \tag{15.15}$$

Let C_1 denote the configuration of ϕ in the sector of $Q = 1$. Then the Hamiltonian gives a mapping $\mathcal{H} : C_1 \to \mathbf{R}$. By picking a configuration which minimizes $\mathcal{H}$, we can find a solution. This, of course, leads to the static sine-Gordon equation

$$-\partial_x^2 \phi + \lambda \sin\phi \,=\, 0 \,. \tag{15.16}$$

General solutions to this can be given by

$$\phi \,=\, b\,\tanh ax \,+\, c \tag{15.17}$$

where a, b, c are, in general, λ-dependent constants. Imposing the boundary conditions (15.5), the solutions become

$$\phi \,=\, \pi \,+\, \pi \tanh ax \,. \tag{15.18}$$

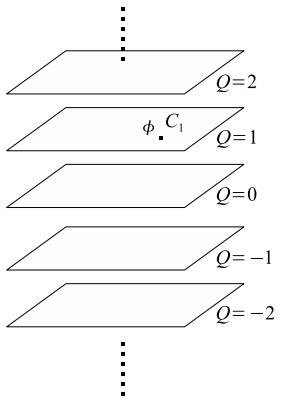

Figure 15.1: ∞-dimensional space that consists of disconnected sectors

Under Lorentz transformations, we can also obtain a moving solution

$$\phi \;=\; \pi \;+\; \pi \tanh a \frac{(x - vt)}{\sqrt{1 - v^2}} \,. \tag{15.19}$$

This describes a solitary wave and this is why the solution is called a soliton. Owing to solitary properties, the shape of solitions is superposed upon collision and the total charge $Q_{\rm tot}$ should be conserved. For a soliton-antisoliton collision, we have $Q_{\rm tot} = 0$, that is, the two solitary waves annihilate each other. In this sense, solitons are quite a bit like particles.

15.2 Solitons in (3+1) dimensions

In this section, we consider solitons in spatial three dimensions. Let x_i $(i = 1, 2, 3)$ be coordinates on $\mathbf{R}^3$. In terms of x_i, a three-sphere S^3 can be parametrized as

$$y_i \;=\; \frac{2x_i}{1 + x^2} \,, \quad y_4 \;=\; \frac{1 - x^2}{1 + x^2} \,. \tag{15.20}$$

Using the stereographic projection of S^3 (see Fig. 15.2), we can then identify configurations on $\mathbf{R}^3$ which have the same boundary value (at spatial infinity) as configurations on a three-sphere S^3. Notice that this is a one-dimensional higher version of the stereographic projection for S^2 considered in (12.44).

Since $S^3 \simeq SU(2)$, we can use the element $g \in SU(2)$ to represent the coordinates on S^3. The element satisfies $g^\dagger g = 1$ and $\det g = 1$. In terms of the 2×2 Pauli matrices σ_i, it is parametrized as

$$g(x) \;=\; a(x) + i b_i(x) \sigma_i \tag{15.21}$$

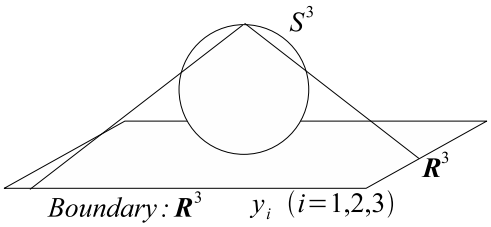

Figure 15.2: Stereographic projection for the three-sphere $y_i^2 + y_4^2 = 1$

where a, b_i $(i = 1, 2, 3)$ are functions of $\vec{x}$, satisfying

$$a^2 + b_i^2 \;=\; 1\,. \tag{15.22}$$

This means that $g(x)$ gives a mapping

$$g(x) \;:\; \mathbf{R}^3 \;\longrightarrow\; S^3\,. \tag{15.23}$$

We can think of it as an analog of $u(x)$ in (15.13).

In the case of $u(x)$, the current is defined by (15.8). Similarly, we can guess the corresponding current in (3+1)-dimensions as

$$J_\mu \;=\; C\,\epsilon_{\mu\nu\alpha\beta}\mathrm{Tr}\left[g^{-1}\partial_\nu g\, g^{-1}\partial_\alpha g\, g^{-1}\partial_\alpha g\right] \tag{15.24}$$

where C is some constant. What we would like to check in the following is whether J_μ is conserved or not without using any equations of motion. Using the relation

$$\partial_\mu(g^{-1}\partial_\nu g) \;=\; -g^{-1}\partial_\mu g\, g^{-1}\partial_\nu g \;+\; g^{-1}\partial_\mu\partial_\nu g \tag{15.25}$$

and antisymmetric properties in the indices $(\mu\nu\alpha\beta)$, we find

$$\partial_\mu J_\mu \;=\; -\,C\,\epsilon_{\mu\nu\alpha\beta}\mathrm{Tr}\,(I_\mu I_\nu I_\alpha I_\beta + I_\nu I_\mu I_\alpha I_\beta + I_\nu I_\alpha I_\mu I_\beta) \tag{15.26}$$

where $I_\mu \equiv g^{-1}\partial_\mu g$. The first term is calculated as

$$\begin{aligned}\epsilon_{\mu\nu\alpha\beta}\mathrm{Tr}(I_\mu I_\nu I_\alpha I_\beta) \;&=\; \epsilon_{\mu\nu\alpha\beta}\mathrm{Tr}(I_\beta I_\mu I_\nu I_\alpha) \;=\; -\epsilon_{\beta\mu\nu\alpha}\mathrm{Tr}(I_\beta I_\mu I_\nu I_\alpha)\\ &=\; -\epsilon_{\mu\nu\alpha\beta}\mathrm{Tr}(I_\mu I_\nu I_\alpha I_\beta) \;=\; 0\,.\end{aligned} \tag{15.27}$$

Similarly, we find that the rest of the terms also vanish. Thus, the current is indeed conserved, $\partial_\mu J_\mu = 0$.

As in the previous section, the associated charge is defined as

$$Q[g] \;=\; \int d^3x\, J_0 \;=\; C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(I_i I_j I_k) \tag{15.28}$$

where $I_i = g^{-1}\partial_i g$. Now consider the transformation $g \to gh$ where h is another element of $SU(2)$. Since I_i transforms as

$$I_i = g^{-1}\partial_i g \longrightarrow h^{-1}(g^{-1}\partial_i g)h + h^{-1}\partial_i h = h^{-1}(I_i + \partial_i h\, h^{-1})h\,, \qquad (15.29)$$

the soliton number becomes

$$\begin{aligned} Q[gh] &= C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}\left[(I_i + \partial_i h\, h^{-1})(I_j + \partial_j h\, h^{-1})(I_k + \partial_k h\, h^{-1})\right] \\ &= C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(I_i I_j I_k) \\ &\quad + C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1} I_j I_k + I_i \partial_j h\, h^{-1} I_k + I_i I_j \partial_k h\, h^{-1}) \\ &\quad + C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1}\partial_j h\, h^{-1} I_k + \partial_i h\, h^{-1} I_j d_k h\, h^{-1} + I_i \partial_j h\, h^{-1}\partial_k h\, h^{-1}) \\ &\quad + C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1}\partial_j h\, h^{-1}\partial_k h\, h^{-1}) \\ &= Q[g] + Q[h] + 3C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1} I_j I_k + \partial_i h\, h^{-1}\partial_j h\, h^{-1} I_k)\,. \qquad (15.30) \end{aligned}$$

The integrand of the last term is a total derivative. This can be seen as follows. Notice that

$$\begin{aligned} \partial_j(\partial_i h\, h^{-1}) &= \partial_j\partial_i h\, h^{-1} + \partial_i h \partial_j h^{-1} \\ &= \partial_j\partial_i h\, h^{-1} - \partial_i h\, h^{-1}\partial_j h\, h^{-1}\,, \qquad (15.31) \\ \partial_j I_k &= -g^{-1}\partial_j g\, g^{-1}\partial_k g + g^{-1}\partial_j\partial_k g\,. \qquad (15.32) \end{aligned}$$

Then, we find

$$\begin{aligned} & 3C\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1} I_j I_k + \partial_i h\, h^{-1}\partial_j h\, h^{-1} I_k) \\ = & \int d^3x\;\partial_j\Big[-3C\,\epsilon_{ijk}\mathrm{Tr}(\partial_i h\, h^{-1}\, g^{-1}\partial_k g)\Big] = 0\,. \qquad (15.33) \end{aligned}$$

Thus $Q[gh]$ is calculated as

$$Q[gh] = Q[g] + Q[h]\,. \qquad (15.34)$$

If h is a small perturbation (close to the identity for all $\vec{x}$), then $Q[h] = 0$. Namely, $Q[gh] = Q[g]$ in this case. This means that $Q[g]$ is invariant under small deformations of g. In other words, $Q[g]$ is a topological invariant for the map $g(x) : \mathbf{R}^3 \to SU(2)$.

In terms of the parametrization (15.21), a particular configuration is given by

$$g_1(x) = a_1(x) + ib_{1i}(x)\sigma_i \qquad (15.35)$$

where

$$a_1(x) = \frac{1-x^2}{1+x^2}\,, \quad b_{1i}(x) = \frac{2x_i}{1+x^2}\,. \tag{15.36}$$

Namely, we choose these parameters as coordinates on S^3 defined in (15.20). Thus, by construction, the winding number should be one, *i.e.*, $Q[g_1] = 1$. Since the volume of S^3 is 2π, using the above parametrization, we can fix the normalization factor of $Q[g]$ as

$$Q[g] \,=\, \frac{1}{24\pi^2}\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(g^{-1}\partial_i g\, g^{-1}\partial_k g\, g^{-1}\partial_k g\,)\,. \tag{15.37}$$

In terms of this $Q[g]$, the above arguments are summarized as

$$Q[\mathbf{1}] = 0\,, \quad Q[g_1] = 1\,. \tag{15.38}$$

From (15.34), the winding number for $g_2 = g_1 g_1$ becomes

$$Q[g_2] \,=\, Q[g_1] + Q[g_1] \,=\, 2\,. \tag{15.39}$$

Similarly, we can obtain configurations for $Q[g] = 3, 4, \cdots$. Using the identity $\mathbf{1} = g_1^{-1} g_1$, on the other hand, we find $Q[\mathbf{1}] = Q[g_1^{-1}] + Q[g_1] = 0$, *i.e.*,

$$Q[g_1^{-1}] = -1\,. \tag{15.40}$$

To summarize, the winding number takes a value of integer $Q[g] \in \mathbf{Z}$. Accordingly, there exist infinite configurations for the map $g(x) : \mathbf{R}^3 \to SU(2) = S^3$, with $g(x) \to 1$ (or any fixed value) as $|x| \to \infty$. As discussed in the previous section, a set of all such maps form an infinite dimensional space which consists of disconnected sectors, see Fig. 15.1. Notice that if the boundary of the base space is homogeneous the above map can be expressed as $g(x) : S^3_x \to S^3_g$. This allows us to associate the winding number to a homotopy class in mathematics.

Invariance of Q under coordinate transformations can explicitly be shown as follows. We first define the coordinate transformation $x^i \to x'^i$ by $M_i^m = \frac{\partial x^m}{\partial x'^i}$. Then we have

$$\left\{\begin{array}{rcl} I'_i &=& g^{-1}\frac{\partial g}{\partial x'^i} \,=\, g^{-1}\frac{\partial g}{\partial x^m}\left(\frac{\partial x^m}{\partial x^i}\right) \,=\, M_i^m I_m\,, \\ dx'^i &=& (M^{-1})^i_{\ m} dx^m\,, \\ d^3x' &=& (\det M^{-1}) d^3x\,. \end{array}\right. \tag{15.41}$$

Substituting these into (15.37), we find

$$Q[g(x')] \,=\, \frac{1}{24\pi^2}\int (\det M^{-1}) d^3x\,\epsilon^{ijk} M_i^l M_j^m M_k^n\,\mathrm{Tr}(I_l I_m I_n) \,=\, Q[g(x)] \tag{15.42}$$

where we use the identity $\epsilon^{ijk} M_i^l M_j^m M_k^n = \det M\,\epsilon^{lmn}$. Q is also independent of the metric of space. Thus, as mentioned before, it is a topological invariant.

Static solitons

To discuss static solutions associated with the Q-soliton, one need to define a Lagrangian or a Hamiltonian as in the case of the sine-Gordon model. A naive guess leads to a static Hamiltonian

$$\mathcal{H}[g] \,=\, \frac{1}{2}\int \mathrm{Tr}(g^{-1}\partial_i g\, g^{-1}\partial_i g)\, d^3x \tag{15.43}$$

which is quadratic in spatial derivatives. For a specific function $g = \widetilde{g}(x)$, the Hamiltonian takes a nonzero value $\mathcal{H}[\widetilde{g}] \neq 0$ classically. We now consider a scale transformation

$$\widetilde{g}(x) \,\longrightarrow\, \widetilde{g}'(x) \,=\, \widetilde{g}\left(\frac{x}{R}\right) \,\equiv\, \widetilde{g}(y) \tag{15.44}$$

where R is a positive real number and $y \equiv \frac{x}{R}$. Under this transformation, the Hamiltonian becomes

$$\mathcal{H}[\widetilde{g}'] \,=\, \frac{1}{2}\int \mathrm{Tr}\left(\widetilde{g}^{-1}(y)\frac{\partial}{\partial y^i}\frac{\widetilde{g}(y)}{R}\,\widetilde{g}^{-1}(y)\frac{\partial}{\partial y^i}\frac{\widetilde{g}(y)}{R}\right) d^3y\, R^3 \,=\, R\,\mathcal{H}[\widetilde{g}]\,. \tag{15.45}$$

The Hamiltonian is linear in R under the scale transformation (15.44). This means that there exist no minima and we can not have finite energy solutions.

Notice that the number of dimensions is crucial in deriving this result. The winding number (15.37), which contains three derivatives in the integrand, is thus invariant under the scale transformations:

$$Q[\widetilde{g}'] \,=\, Q[\widetilde{g}]\,. \tag{15.46}$$

In order to obtain finite energy solutions, we need to modify the Hamiltonian. One of such modifications is given by

$$\mathcal{H}[g] \,=\, \frac{1}{2}\int \mathrm{Tr}(g^{-1}\partial_i g\, g^{-1}\partial_i g)\, d^3x \,+\, \int \mathrm{Tr}(g^{-1}\partial g)^4\, d^3x\,. \tag{15.47}$$

Then, under the scale transformations, the above Hamiltonian can classically be evaluated as $\alpha R + \beta/R$ where $\alpha, \beta > 0$. Thus we have a minimum and we can use (15.47) to obtain static solitons. The quartic term in (15.47) therefore induces soliton solutions.

An example of (3+1)-dimensional solitions is given by the so-called skyrmions, named after pioneering work by Skyrme. The skyrmions are considered in the context of spontaneous chiral symmetry breaking of QCD, $SU(3) \times SU(3) \to SU(3)$. As discussed in (14.71), this leads to an effective theory for octet mesons described by the Hamiltonian (15.43) with $g \in SU(3)$. Thus, using the modified Hamiltonian like (15.47), one can obtain static soliton solutions, the skyrmions, which turn out to have many of the properties of baryons. The associated winding number Q then corresponds to the baryon number of the skyrmion.

15.3 Solitons in (2+1) dimensions

In the (2+1)-dimensional case, the map of interest is given by

$$\phi^a(x) \,:\, S^2_x \,\longrightarrow\, S^2_\phi \tag{15.48}$$

where both x_a and ϕ^a $(a = 1, 2, 3)$ denote coordinates of S^2, satisfying $x_a x_a = 1$, $\phi^a \phi^a = 1$. Notice that ϕ^a are not elements of a Lie group. The target space is now a coset space $S^2 = SU(2)/U(1)$. As before, the unit map $\phi^a = x_a$ is given by the stereographic projection

$$x_i = \frac{2y_i}{1+y^2}\,, \quad x_3 = \frac{1-y^2}{1+y^2} \tag{15.49}$$

where $i = 1, 2$, $y_a \in \mathbf{R}^3$ and $y^2 = y_a y_a$. The volume of S^2_x is given by $\frac{1}{2}\epsilon_{abc} x_a dx^b dx^c$. The corresponding winding number is then defined by

$$\begin{aligned} Q[\phi] &= \frac{1}{4\pi} \times \big[\text{volume traced out by } \phi^a(x)\big] \\ &= \frac{1}{8\pi} \int d^2x\, \epsilon_{abc}\, \partial_\mu \phi^a \partial_\nu \phi^b\, \phi^c\, \epsilon^{\mu\nu}\,. \end{aligned} \tag{15.50}$$

Since $Q = \int d^2x J^0$, the topological current J^α is expressed as

$$J^\alpha = \frac{1}{8\pi}\, \epsilon^{\mu\nu\alpha} \epsilon_{abc}\, (\partial_\mu \phi^a \partial_\nu \phi^b) \phi^c\,. \tag{15.51}$$

The conservation of current can be seen from

$$\partial_\alpha J^\alpha = \frac{1}{8\pi}\, \epsilon^{\mu\nu\alpha} \epsilon_{abc}\, \partial_\mu \phi^a \partial_\nu \phi^b \partial_\alpha \phi^c = \frac{3}{4\pi} \det(\partial\phi) = 0 \tag{15.52}$$

where we use $\det(\partial\phi) = \frac{1}{3!} \epsilon^{\mu\nu\alpha} \epsilon_{abc} \partial_\mu \phi^a \partial_\nu \phi^b \partial_\alpha \phi^c$. Notice that $\det(\partial\phi) = 0$ follows from the constraint $\phi^a \phi^a = 1$.

We now check that Q is invariant under small deformations of ϕ^a.

$$\delta Q \equiv Q[\phi + \delta\phi] - Q[\phi] = \frac{3}{8\pi} \int d^2x\, \epsilon_{abc}\, \delta\phi^a \partial_\mu \phi^b \partial_\nu \phi^c\, \epsilon^{\mu\nu}\,. \tag{15.53}$$

Notice that we still have the constraint $(\phi + \delta\phi)^a (\phi + \delta\phi)^a = 1$. Thus the condition $\phi^a\, \delta\phi^a = 0$ should be imposed. We can then parametrize $\delta\phi^a$ as $\delta\phi^a = \epsilon^{amn} \phi^m \theta^n$. Using this, (15.53) becomes

$$\delta Q = \frac{3}{8\pi} \int d^2x\, (\phi_b \theta_c - \phi_c \theta_b) \partial_\mu \phi^b \partial_\nu \phi^c = 0\,. \tag{15.54}$$

Thus Q is indeed invariant under small deformations of ϕ^a. From the same argument in the previous section, we find $Q \in \mathbf{Z}$ as well.

Physical models

As in the (3+1)-dimensional case, we can postulate the Hamiltonian for static solitons in the form of

$$\mathcal{H}[\phi] = \frac{1}{2} \int d^2x\, \partial_i \phi^a \partial_i \phi^a\,. \tag{15.55}$$

In the present (2+1)-dimensional case, this Hamiltonian is scale invariant. The Hamiltonian therefore gives finite-energy solutions for the configuration of ϕ^a with $Q \neq 0$.

We now look at the following inequality

$$\int d^2x \left(\partial_i \phi^a - \frac{1}{8\pi} \epsilon_{abc} \partial_j \phi^b \phi^c \epsilon_{ij} \right)^2 \geq 0 \,. \tag{15.56}$$

The right-hand side is expanded as

$$\begin{aligned} & \int d^2x \, \partial_i \phi^a \partial_i \phi^a - 2Q + \int d^2x \left(\frac{1}{8\pi} \right)^2 \partial_i \phi^b \, \phi^c (\partial_i \phi^b \, \phi^c - \partial_i \phi^c \, \phi^b) \\ = \quad & 2\mathcal{H}[\phi] - 2Q + 2 \left(\frac{1}{8\pi} \right)^2 \mathcal{H}[\phi] \end{aligned} \tag{15.57}$$

where we use $\phi^c \phi^c = 1$ and $\partial_i \phi^c \phi^c = 0$. Thus we find that the Hamiltonian (15.55) has a lower bound

$$\mathcal{H}[\phi] \geq \frac{(8\pi)^2}{1 + (8\pi)^2} Q[\phi] \,. \tag{15.58}$$

This is called the Bogomol'nyi bound. At the minimum, the equality in (15.56) holds, *i.e.*,

$$\partial_i \phi^a - \frac{1}{8\pi} \epsilon_{abc} \partial_j \phi^b \phi^c \epsilon_{ij} = 0 \,. \tag{15.59}$$

This is known as the Bogomol'nyi equations.

The existence of the Bogomol'nyi bound (15.58) explicitly shows that we can use the Hamiltonian (15.55) to construct physical models in (2+1) dimensions. For example, it describes two-dimensional skyrmions. As discussed in earlier chapters, it can also be used for the study of the ferromagnet on S^2 and the quantum Hall effect.

Index