

# **LIE GROUPS IN PHYSICS<sup>1</sup>**

version 25/06/07

**Institute for Theoretical Physics  
Utrecht University  
Beta Faculty**

**2007**

English version by G. 't Hooft  
Original text by M.J.G. Veltman  
B.Q.P.J. de Wit  
and G. 't Hooft

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Quantum mechanics and rotation invariance</b>	<b>7</b>
<b>3</b>	<b>The group of rotations in three dimensions</b>	<b>14</b>
<b>4</b>	<b>More about representations</b>	<b>22</b>
<b>5</b>	<b>Ladder operators</b>	<b>26</b>
<b>6</b>	<b>The group <math>SU(2)</math></b>	<b>31</b>
<b>7</b>	<b>Spin and angular distributions</b>	<b>39</b>
<b>8</b>	<b>Isospin</b>	<b>45</b>
<b>9</b>	<b>The Hydrogen Atom</b>	<b>48</b>
<b>10</b>	<b>The group <math>SU(3)</math></b>	<b>55</b>
<b>11</b>	<b>Representations of <math>SU(N)</math>; Young tableaux</b>	<b>60</b>
<b>12</b>	<b>Beyond these notes</b>	<b>61</b>
	Appendix A. Summary of the properties of matrices	62
	Appendix B. Differentiation of matrices	64
	Appendix C. Functions of matrices	65
	Appendix D. The Campbell-Baker-Hausdorff formula	66
	Appendix E. Complex inner product, unitary and hermitian matrices	70

# 1. Introduction

Many systems studied in physics show some form of symmetry. In physics, this means the following: we can consider some *transformation rule*, like a rotation, a displacement, or the reflection by a mirror, and we compare the original system with the transformed system. If they show some resemblance, we have a symmetry. A snow flake looks like itself when we rotate it by  $60^\circ$  or when we perform a mirror reflection. We say that the snow flake has a symmetry. If we replace a proton by a neutron, and *vice versa*, the replaced particles behave very much like the originals; this is also a symmetry. Many laws of Nature have symmetries in this sense. Sometimes the symmetry is perfect, but often it is not exact; the transformed system is then slightly different from the original; the symmetry is broken.

If system  $A$  resembles system  $B$ , and system  $B$  resembles  $C$ , then  $A$  resembles  $C$ . Therefore, the product of two symmetry transformations is again a symmetry transformation. Thus, the set of all symmetry transformations that characterize the symmetry of a system, are elements of a *group*. For example, the reflections with respect to a plane form a group that contains just two elements: the reflection operation and the identity — the identity being the one operation that leaves everything the same. The rotations in three-dimensional space, the set of all Lorentz transformations, and the set of all parallel displacements also form groups, which have an unlimited number of elements. For obvious reasons, groups with a finite (or denumerable) number of elements are called *discrete* groups; groups of transformations that continuously depend on a number of parameters, such as the rotations, which can be defined in terms of a few angular parameters, are called *continuous* groups.

The symmetry of a system implies certain relations among observable quantities, which may be obeyed with great precision, independently of the nature of the forces acting in the system. In the hydrogen atom, for example, one finds that the energies of different states of the atom, are exactly equal, as a consequence of the rotational invariance of the system. However, one also often finds that the symmetry of a physical system is only approximately realized. An infinite crystal, for example, is invariant under those translations for which the displacement is an integral multiple of the distance between two adjacent atoms. In reality, however, the crystal has a definite size, and its surface perturbs the translational symmetry. Nevertheless, if the crystal contains a sufficiently large number of atoms, the disturbance due to the surface has little effects on the properties at the interior.

An other example of a symmetry that is only approximately realized, is encountered in elementary particle physics. The so-called  $\Delta^+$  particle, which is one of the excited states of the nucleons, decays into a nucleon and an other particle, the  $\pi$ -meson, also called *pion*. There exist two kinds of nucleons, neutrons and protons, and there are three types of pions, the electrically charged pions  $\pi^+$  and  $\pi^-$ , and the neutral one,  $\pi^0$ . Since the total electric charge of the  $\Delta^+$  must be preserved during its decay, one distinguishes

---

<sup>1</sup>This lecture course was originally set up by M. Veltman, and subsequently modified and extended by B. de Wit and G. 't Hooft.

nucleons	pions	$\Delta$ particles
$m_{\text{proton}} \approx 938 \text{ MeV}/c^2$	$m_{\pi^+} \approx 140 \text{ MeV}/c^2$	$m_{\Delta^{++}} \approx 1231 \text{ MeV}/c^2$
$m_{\text{neutron}} \approx 939 \text{ MeV}/c^2$	$m_{\pi^0} \approx 135 \text{ MeV}/c^2$	$m_{\Delta^+} \approx 1232 \text{ MeV}/c^2$
	$m_{\pi^-} \approx 140 \text{ MeV}/c^2$	$m_{\Delta^0} \approx 1233 \text{ MeV}/c^2$
		$m_{\Delta^-} \approx 1235 \text{ MeV}/c^2$

Table 1: Masses of nucleons, pions and  $\Delta$  particles, expressed in  $\text{MeV}/c^2$ .

two possible decay modes:

$$\Delta^+ \rightarrow n \pi^+ \quad \text{and} \quad \Delta^+ \rightarrow p \pi^0. \quad (1.1)$$

Remarkably, the second decay occurs twice as often as the the first one, a fact that seems to be difficult to explain as being due to the differences in the charges of the decay products. A natural explanation of this factor 2 could follow from symmetry considerations. This is not as strange as it might seem, because protons and neutrons have nearly identical masses, just as the three species of pions and the four  $\Delta$  particles that are found in Nature (see table).

It will be demonstrated that the near equality of the masses, and also the factor 2 in the two decay modes (1.1), can be explained by assuming nature to be invariant under so-called *isospin* transformations. The notion of ‘isobaric spin’, or ‘isospin’ for short, was introduced by Heisenberg in 1932. He was puzzled by the fact that protons and neutrons have nearly equal masses, while, apart from the obvious differences in electrical charge, also other properties are much alike. Thus, the nucleons form a *doublet*, just like electrons that show a doublet structure as a consequence of the fact that there are two possible spin orientations for the electron states — hence the term *isobaric spin*. Later, it turned out that elementary particles with nearly equal masses can always be arranged in so-called isospin multiplets. The nucleons form an isospin doublet, the pions an isospin triplet, and the  $\Delta$  particles an isospin quadruplet. Particles inside a single multiplet all have approximately identical masses, but different electric charges. The charge arrangement is as indicated in the table: no two particles in one multiplet have the same charge, and the particles can always be arranged in such a way that the charge difference between two successive particles is exactly one elementary charge unit.

However, it will be clear that isospin invariance can only be an approximation, since the masses of the nucleons, pions and  $\Delta$  particles turn out to depend somewhat on their electric charges. The mass differences within a multiplet are only of the order of a few percent, and this is the degree of accuracy that one can expect for theoretical predictions based upon isospin invariance.

The above example is an application of group theory in the physics of elementary particles, but invariance principles play an important role in nearly all branches of physics. In atomic physics we frequently notice the consequences of rotation invariance, in nuclear physics we have rotation and isospin invariance, in solid state physics also invariance under discrete translations and rotations. Also in (quantum) field theory, symmetry transformations are important. A very special kind of transformations are encountered

for example in electrodynamics. Here, electric and magnetic fields can be expressed in terms of the so-called *vector potential*  $A_\mu(x)$ , for which we use a relativistic four-vector notation ( $\mu = 0, 1, 2, 3$ ):

$$A_\mu(x) = (-c^{-1}\phi(x), \mathbf{A}(x)), \quad x^\mu = (ct, \mathbf{x}), \quad (1.2)$$

where  $\phi$  denotes the potential, and  $\mathbf{A}$  the three-dimensional vector potential field;  $c$  is the velocity of light. The electric and magnetic fields are defined by

$$\mathbf{E} = -\nabla\phi - c^{-1}\frac{\partial\mathbf{A}}{\partial t}, \quad (1.3)$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (1.4)$$

An electrically charged particle is described by a complex wave function  $\psi(\vec{x}, t)$ . The Schrödinger equation obeyed by this wave function remains valid when one performs a rotation in the complex plane:

$$\psi(\vec{x}, t) \rightarrow e^{i\Lambda} \psi(\vec{x}, t). \quad (1.5)$$

Is the phase factor  $\Lambda$  allowed to vary in space and time?

The answer to this is yes, however only if the Schrödinger equation depends on the vector potential in a very special way. Wherever a derivative  $\partial_\mu$  occurs, it must be in the combination

$$D_\mu = \partial_\mu - ieA_\mu, \quad (1.6)$$

where  $e$  is the electric charge of the particle in question. If  $\Lambda(\vec{x}, t)$  depends on  $\vec{x}$  and  $t$ , then (1.5) must be associated with the following transformation rules for the potential fields:

$$\mathbf{A}(x) \rightarrow \mathbf{A}(x) + e^{-1}\nabla\Lambda(x), \quad (1.7)$$

$$\phi(x) \rightarrow \phi(x) - (ce)^{-1}\frac{\partial}{\partial t}\Lambda(x), \quad (1.8)$$

or, in four-vector notation,

$$A_\mu(x) \rightarrow A_\mu(x) + e^{-1}\partial_\mu\Lambda(x). \quad (1.9)$$

It can now easily be established that  $\mathbf{E}$  en  $\mathbf{B}$  will not be affected by this so-called *gauge transformation*. Furthermore, we derive:

$$D_\mu\psi(x) \rightarrow e^{i\Lambda(x)}D_\mu\psi(x). \quad (1.10)$$

Notice that the substitution (1.6) in the Schrödinger equation is all that is needed to include the interaction of a charged particle with the fields  $\mathbf{E}$  en  $\mathbf{B}$ .

These phase factors define a group, called the group of  $1 \times 1$  unitary matrices,  $U(1)$ . In this case, the group is quite a simple one, but it so happens that similar theories exist

that are based on other (continuous) groups that are quite a bit more complicated such as the group  $SU(2)$  that will be considered in these lectures. Theories of this type are known as *gauge theories*, or Yang-Mills theories, and the field  $A_\mu$  is called a *gauge field*. The fact that  $\mathbf{E}$  en  $\mathbf{B}$  are invariant under gauge transformations implies that electromagnetic phenomena are *gauge-invariant*. For more general groups it turns out that several of these gauge fields are needed: they form multiplets.

Surprisingly, the theory of gravitation, Einstein's general relativity theory, turns out to be a gauge theory as well, be it of a somewhat different type. This theory can be considered to be the gauge theory of the general coordinate transformations, the most general reparametrizations of points in space and time,

$$x^\mu \rightarrow x^\mu + \xi^\mu(x) . \quad (1.11)$$

The gauge field here is the gravitational field, taking the form of a metric, which is to be used in the definitions of distances and angles in four-dimensional space and time. All of this is the subject of an entire lecture course, *Introduction to General Relativity*.

The fact that gauge transformations are associated to an abstract group, and can depend on space and time as well, can give rise to interesting phenomena of a topological nature. Examples of this are flux quantization in super conductors, the Aharonov-Bohm effect in quantum mechanics, and magnetic monopoles. To illustrate the relevance of topology, we consider again the group of the  $U(1)$  gauge transformations, but now in two-dimensional space (or equivalently, in a situation where the fields only depend on two of the three space coordinates). Let  $\psi(x, y)$  be a complex function, such as a wave function in quantum mechanics, transforming under these gauge transformations, *i.e.*

$$\psi(x, y) \rightarrow e^{i\Lambda(x,y)}\psi(x, y) . \quad (1.12)$$

From the fact that the phase of  $\psi$  can be modified everywhere by applying different gauge transformations, one might conclude that the phase of  $\psi$  is actually irrelevant for the description of the system. This however is not quite the case. Consider for instance a function that vanishes at the origin. Now take a closed curve in the  $x$ - $y$  plane, and check how the phase of  $\psi(x, y)$  changes along the curve. After a complete run along the curve the phase might not necessarily take the same value as at the beginning, but if we assume that  $\psi(x, y)$  is single-valued on the plane, then the phase difference will be equal to  $2\pi n$ , where  $n$  is an arbitrary *integral* number. This number is called the *winding number*. An example of a situation with winding number  $n = 1$  is pictured in Fig. 1; the phase angle makes a full turn over  $2\pi$  when we follow the function  $\psi(x, y)$  along a curve winding once around the origin. One can easily imagine situations with other winding numbers. The case  $n = 0$  for instance occurs when the phase of  $\psi(x, y)$  is constant.

If we change the function  $\psi(x, y)$  continuously, the winding number will not change. This is why the winding number is called a *topological invariant*. This also implies that the winding number will not change under the gauge transformations (1.12), provided that we limit ourselves to gauge transformations that are well-defined in the entire plane. Note also that the winding number does not depend on the choice of the closed curve around the origin, as long as it is not pulled across the origin or any other zero of the function

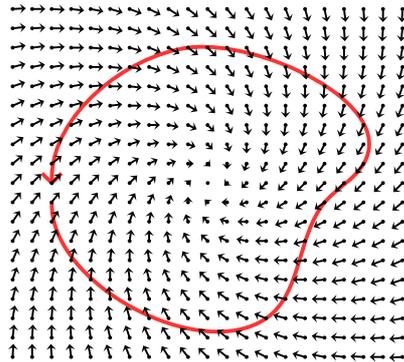


Figure 1: De phase angle of  $\psi(x, y)$  indicated by an arrow (whose length is immaterial, but could be given for instance by  $|\psi(x, y)|$ ) at various spots in the  $x - y$  plane. This function has a zero at the origin.

$\psi(x, y)$ . All this implies that although locally, that is, at one point and its immediate neighborhood, the phase of  $\psi$  can be made to vanish, this can be realized globally, that is, on the entire plane, only if the winding number for any closed curve equals zero.

A similar situation can be imagined for the vector potential. Once more consider the two-dimensional plane, and assume that we are dealing with a magnetic field that is everywhere equal to zero, except for a small region surrounding the origin. In this region,  $\mathbf{A}$  cannot be equal to zero, because of the relation (1.4). However, in the surrounding region, where  $\mathbf{B}$  vanishes, there may seem to be no reason why not also  $\mathbf{A}$  should vanish. Indeed, one can show that, at every given point and its neighborhood, a suitably chosen gauge transformation can ensure  $\mathbf{A}(x)$  to vanish there. This result, however, can only hold locally, as we can verify by considering the following loop integral:

$$\Phi[C] = \oint_C A_i dx^i, \quad (1.13)$$

where  $C$  is a given closed curve. It is easy to check that  $\Phi[C]$  does not change under a gauge transformation (1.5). Indeed, we know from the theory of magnetism that  $\Phi[C]$  must be proportional to the total magnetic flux through the surface enclosed by the curve  $C$ .

Applying this to the given situation, we take the curve  $C$  to surround the origin and the region where  $\mathbf{B} \neq \mathbf{0}$ , so that the  $\mathbf{B}$  field vanishes on the curve itself. The quantity  $\Phi[C]$  equals the total flux through  $C$ , which may well be different from zero. If this is the case, we cannot transform  $\mathbf{A}$  away in the entire outside region, even if it can be transformed away locally<sup>2</sup>. Note that the magnetic flux here plays the same role as the winding number of the previous example. Indeed, in superconducting material, the gauge phases can be chosen such that  $\mathbf{A}$  vanishes, and consequently, magnetic flux going through a superconducting coil is limited to integral values: the flux is *quantized*.

---

<sup>2</sup>This causes an interesting quantum mechanical effect in electrons outside a magnetic field, to wit, the *Aharonov-Bohm effect*.

Under some circumstances, magnetic field lines can penetrate superconducting materials in the form of *vortices*. These vortices again are quantized. In the case of more complicated groups, such as  $SU(2)$ , other situations of a similar nature can occur: magnetic monopoles are topologically stable objects in three dimensions; even in four dimensions one can have such phenomena, referred to as “instantons”.

Clearly, group theory plays an essential role in physics. In these lectures we will primarily limit ourselves to the group of three-dimensional rotations, mostly in the context of quantum mechanics. Many of the essentials can be clarified this way, and the treatment can be made reasonably transparent, physically and mathematically. The course does not intend to give a complete mathematical analysis; rather, we wish to illustrate as clearly as possible the relevance of group theory for physics. Therefore, some physical applications will be displayed extensively. The rotation group is an example of a so-called *compact Lie group*. In most applications, we consider the *representations* of this group. Representation theory for such groups is completely known in mathematics. Some advance knowledge of linear algebra (matrices, inner products, traces, functions and derivatives of matrices, etc.) will be necessary. For completeness, some of the most important properties of matrices are summarized in a couple of appendices.

## 2. Quantum mechanics and rotation invariance

Quantum mechanics tells us that any physical system can be described by a (usually complex) wave function. This wave function is a solution of a differential equation (for instance the Schrödinger equation, if a non-relativistic limit is applicable) with boundary conditions determined by the physical situation. We will not indulge in the problems of determining this wave function in all sorts of cases, but we are interested in the properties of wave functions that follow from the fact that Nature shows certain symmetries. By making use of these symmetries we can save ourselves a lot of hard work doing calculations.

One of the most obvious symmetries that we observe in nature around us, is invariance of the laws of nature under rotations in three-dimensional space. An observer expects that the results of measurements should be independent of the orientation of his or her apparatus in space, assuming that the experimental setup is not interacting with its environment, or with the Earth's gravitational field. For instance, one does not expect that the time shown by a watch will depend on its orientation in space, or that the way a calculator works changes if we rotate it. Rotational symmetry can be found in many fundamental equations of physics: Newton's laws, Maxwell's laws, and Schrödinger's equation for example do not depend on orientation in space. To state things more precisely: Nature's laws are invariant under rotations in three-dimensional space.

We now intend to find out what the consequences are of this invariance under rotation for wave functions. From classical mechanics it is known that rotational invariance of a system with no interaction with its environment, gives rise to conservation of angular momentum: in such a system, the total angular momentum is a constant of the motion. This conservation law turns out to be independent of the details of the dynamical laws; it simply follows from more general considerations. It can be deduced in quantum mechanics as well. There turns out to be a connection between the behavior of a wave function under rotations and the conservation of angular momentum.

The equations may be hard to solve explicitly. But consider a wave function  $\psi$  depending on all sorts of variables, being the solution of some linear differential equation:

$$\mathcal{D}\psi = 0 . \quad (2.1)$$

The essential thing is that the exact form of  $\mathcal{D}$  does not matter; the only thing that matters is that  $\mathcal{D}$  be invariant under rotations. An example is Schrödinger's equation for a particle moving in a spherically symmetric potential  $V(r)$ ,

$$\left[ \frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) - V(r) + i\hbar \frac{\partial}{\partial t} \right] \psi(\vec{x}, t) = 0 , \quad r \stackrel{\text{def}}{=} \sqrt{\vec{x}^2} . \quad (2.2)$$

Consider now the behavior of this differential equation under rotations. When we rotate, the position vector  $\vec{x}$  turns into an other vector with coordinates  $x'_i$ :

$$x'_i = \sum_j R_{ij} x_j . \quad (2.3)$$

Here, we should characterize the rotation using a  $3 \times 3$  matrix  $R$ , that is orthogonal and has determinant equal to 1 (orthogonal matrices with determinant  $-1$  correspond to mirror reflections). The orthogonality condition for  $R$  implies that

$$\tilde{R}R = R\tilde{R} = \mathbf{1}, \quad \text{or} \quad \sum_i R_{ij} R_{ik} = \delta_{jk}; \quad \sum_j R_{ij} R_{kj} = \delta_{ik}, \quad (2.4)$$

where  $\tilde{R}$  is the transpose of  $R$  (defined by  $\tilde{R}_{ij} = R_{ji}$ ).

It is not difficult now to check that equation (2.2) is rotationally invariant. To see this, consider<sup>3</sup> the function  $\psi'(\vec{x}, t) \stackrel{\text{def}}{=} \psi(\vec{x}', t)$ ,

$$\frac{\partial}{\partial x_i} \psi'(\vec{x}, t) = \frac{\partial}{\partial x_i} \psi(\vec{x}', t) = \sum_j \frac{\partial x'_j}{\partial x_i} \frac{\partial}{\partial x'_j} \psi(\vec{x}', t) = \sum_j R_{ji} \frac{\partial}{\partial x'_j} \psi(\vec{x}', t), \quad (2.5)$$

where use was made of Eq. (2.3). Subsequently, we observe that

$$\begin{aligned} \sum_i \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_i} \psi(\vec{x}', t) &= \sum_{i,j,k} R_{ji} R_{ki} \frac{\partial}{\partial x'_j} \frac{\partial}{\partial x'_k} \psi(\vec{x}', t) \\ &= \sum_i \frac{\partial}{\partial x'_i} \frac{\partial}{\partial x'_i} \psi(\vec{x}', t), \end{aligned} \quad (2.6)$$

where we made use of Eq. (2.4). Since  $\vec{x}'^2 = \vec{x}^2$ , the potential  $V(r)$  also remains the same after a rotation. From the above, it follows that Equation (2.2) is invariant under rotations: if  $\psi(\vec{x}, t)$  is a solution of Eq. (2.2), then also  $\psi'(\vec{x}, t)$  must be a solution of the same equation.

In the above, use was made of the fact that rotations can be represented by real  $3 \times 3$  matrices  $R$ . Their determinant must be  $+1$ , and they must obey the orthogonality condition  $R\tilde{R} = \mathbf{1}$ . Every rotation in 3 dimensions can be represented by three angles (this will be made more precise in Chapter 3) Let  $R_1$  and  $R_2$  both be matrices belonging to some rotations; then their product  $R_3 = R_1 R_2$  will also be a rotation. This statement is proven as follows: assume that  $R_1$  and  $R_2$  are orthogonal matrices with determinant 1. From the fact that

$$\tilde{R}_1 = R_1^{-1}, \quad \tilde{R}_2 = R_2^{-1}, \quad (2.7)$$

it follows that also  $R_3 = R_1 R_2$  is orthogonal:

$$\tilde{R}_3 = \widetilde{R_1 R_2} = \tilde{R}_2 \tilde{R}_1 = R_2^{-1} R_1^{-1} = (R_1 R_2)^{-1} = R_3^{-1}. \quad (2.8)$$

Furthermore, we derive that

$$\det R_3 = \det(R_1 R_2) = \det R_1 \det R_2 = 1, \quad (2.9)$$

---

<sup>3</sup>By rotating  $\vec{x}$  first, and taking the old function at the new point  $\vec{x}'$  afterwards, we actually rotate the wave function into the opposite direction. This is a question of notation that, rather from being objectionable, avoids unnecessary complications in the calculations.

so that  $R_3$  is a rotation as well. Note that also the product  $R_4 = R_2R_1$  is a rotation, but that  $R_3$  en  $R_4$  need not be the same. In other words, rotations are not *commutative*; when applied in a different order, the result will be different, in general.

We observe that the rotations form what is known as a *group*. A set of elements (here the set of real  $3 \times 3$  matrices  $R$  with determinant 1 and  $\tilde{R}R = \mathbf{1}$ ) is called a group if an operation exists that we call ‘multiplication’ (here the ordinary matrix multiplication), in such a way that the following demands are obeyed:

1. If  $R_1$  en  $R_2$  are elements of the group, then also the product  $R_1R_2$  is an element of the group.
2. The multiplication is associative:  $R_1(R_2R_3) = (R_1R_2)R_3$ . So, one may either first multiply  $R_2$  with  $R_3$ , and then multiply the result with  $R_1$ , or perform these two operations in the opposite order. Note that the order in which the matrices appear in this expression *does* have to stay the same.
3. There exists a unity element  $\mathbf{1}$ , such that  $\mathbf{1}R = R$  for all elements  $R$  of the group. This unity element is also an element of the group.
4. For all elements  $R$  of the group, there exists in the group an *inverse* element  $R^{-1}$  such that  $R^{-1}R = \mathbf{1}$ .

The set of rotation matrices possesses all these properties. This set forms a group with infinitely many elements.

Every group is fully characterized by its multiplication structure, *i.e.* the relation between the elements *via* the multiplication rules. Later, we will attempt to define this notion of “structure” more precisely in terms of formulae. Note that a group does not possess notions such as “add” or “subtract”, only “multiply”. There is no “zero-element” in a group.

Much use is made of the fact that the set of all transformations that leave a system invariant, together form a group. If we have two invariance transformations, we can immediately find a third, by subjecting the quantities in terms of which the theory is defined, to the two transformations in succession. Obviously, the resulting transformation must leave the system invariant as well, and so this “product transformation” belongs to our set. Thus, the first condition defining a group is fulfilled; the others usually are quite obvious as well.

For what follows, the time dependence of the wave function is immaterial, and therefore we henceforth write a rotation  $R$  of a wave function as:

$$\psi'(\vec{x}) = \psi(\vec{x}') = \psi(R\vec{x}) . \quad (2.10)$$

Applying a second rotation  $S$ , gives us

$$\psi'' = \psi'(S\vec{x}) = \psi(RS\vec{x}) . \quad (2.11)$$

In what follows now, we will make use of the fact that the equation  $\mathcal{D}\psi = 0$  is a *linear* equation. This is in contrast to the invariance transformation  $R$ , which may or may not be linear: the sum of two matrices  $R$  and  $S$  usually is not a legitimate rotation. It *is* true that if we have two solutions  $\psi_1$  and  $\psi_2$  of the equation (2.1), then every linear combination of these is a solution as well:

$$\mathcal{D}(\lambda\psi_1 + \mu\psi_2) = \lambda\mathcal{D}\psi_1 + \mu\mathcal{D}\psi_2 = 0. \quad (2.12)$$

In general: if  $\psi_1, \dots, \psi_n$  are solutions of the equation in (2.1) then also every linear combination

$$\lambda_1\psi_1 + \lambda_2\psi_2 + \dots + \lambda_n\psi_n \quad (2.13)$$

is a solution of (2.1).

Regarding the behavior under rotations, we now distinguish two possible situations. Either the wave function  $\psi$  is rotationally invariant, that is, upon a rotation,  $\psi$  turns into itself,

$$\psi'(\vec{x}) = \psi(\vec{x}) \iff \psi(\vec{x}') = \psi(\vec{x}), \quad (2.14)$$

or we have sets of linearly independent solutions  $\psi_1, \dots, \psi_n$ , that, upon a rotation, each transform into some linear combination of the others. To illustrate the second possibility, we can take for example the set of solutions of particles moving in all possible directions. In this case, the set  $\psi_1, \dots, \psi_n$  contains an infinite number of solutions. In order to avoid complications due to the infinite number of elements in this set, we can limit ourselves either to particles at rest, or omit the momentum dependence of the wave functions. Upon a rotation, a particle at rest turns into itself, but the internal structure might change. In this case, the set of wave functions that rotate into one another usually only contains a finite number of linearly independent solutions. If the particle is in its ground state, the associated wave function is often rotationally invariant; in that case, the set only contains one wave function. If the particle is in an excited state, different excited states can emerge after a rotation.

Now let there be given such a set  $\Psi = (\psi_1, \dots, \psi_n)$  of wave functions transforming into one another upon a rotation. This means that after a rotation,  $\psi_1$  turns into some linear combination of  $\psi_1, \dots, \psi_n$ ,

$$\psi'_1(\vec{x}) \equiv \psi_1(R\vec{x}) = d_{11}\psi_1(\vec{x}) + d_{12}\psi_2(\vec{x}) + \dots + d_{1n}\psi_n(\vec{x}), \quad (2.15)$$

and a similar expression holds for  $\psi_2, \dots, \psi_n$ . In general, we can write

$$\psi'_A = \sum_B d_{AB}\psi_B. \quad (A, B = 1, \dots, n). \quad (2.16)$$

The coefficients  $d_{AB}$  depend on  $R$  and form a matrix  $D(R)$ , such that

$$\Psi'(\vec{x}) = \Psi(R\vec{x}) = D(R)\Psi(\vec{x}), \quad (2.17)$$

where we indicated the wave functions  $\psi_1, \dots, \psi_n$  as a column vector  $\Psi$ . In the cases to be discussed next, there is only a limited number of linearly independent solutions of

the equation  $\mathcal{D}\psi = 0$ , and therefore the space of all solutions (2.15) that we obtain by rotating one of them, must be finite-dimensional.

The matrices  $D(R)$  in (2.15)-(2.16) are related to the rotation matrices  $R$  in the sense that for every rotation  $R$  in 3-dimensional space a matrix  $D(R)$  exists that turns the solutions  $\psi_A$  into linear combinations of the same solutions. One can, however, say more. A given rotation can either be applied at once, or be the result of several rotations performed in succession. Whatever is the case, the final result should be the same. This implies that the matrices  $D(R)$  must possess certain multiplication properties. To derive these, consider two successive rotations,  $R$  and  $S$  (see Eq. (2.11)). Let  $R$  be associated with a matrix  $D(R)$ , and  $S$  with a matrix  $D(S)$ . In formulae:

$$\begin{aligned}\Psi(R\vec{x}) &= D(R)\Psi(\vec{x}), \\ \Psi(S\vec{x}) &= D(S)\Psi(\vec{x}).\end{aligned}\tag{2.18}$$

Obviously, the combined rotation  $RS$  must be associated with a matrix  $D(RS)$ , so that we have

$$\Psi(RS\vec{x}) = D(RS)\Psi(\vec{x}).\tag{2.19}$$

But we can also determine  $\Psi(RS)$  using Eq. (2.18),

$$\Psi(RS\vec{x}) = D(R)\Psi(S\vec{x}) = D(R)D(S)\Psi(\vec{x}).\tag{2.20}$$

therefore, one must have<sup>4</sup>

$$D(R)D(S) = D(RS).\tag{2.21}$$

Thus, the matrices  $D(R)$  must have the same multiplication rules, the same multiplication structure, as the matrices  $R$ . A mapping of the group elements  $R$  on matrices  $D(R)$  with this property is said to be a ‘representation’ of the group. We shall study various kinds of representations of the group of rotations in three dimensions.

Summarizing: a set of matrices forms a *representation* of a group, if one has

1. Every element  $a$  of the group is mapped onto a matrix  $A$ ,
2. The product of two elements is mapped onto the product of the corresponding matrices, *i.e.* if  $a$ ,  $b$  and  $c$  are associated to the matrices  $A$ ,  $B$ , and  $C$ , and  $c = ab$ , then one must have  $C = AB$ .

We found the following result: Upon rotations in three-dimensional space, the wave functions of a physical system must transform as linear mappings that form a representation of the group of rotations in three dimensions.

---

<sup>4</sup>In this derivation, note the order of  $R$  en  $S$ . The correct mathematical notation is:  $D(R)\Psi = \Psi \cdot R$ , so  $D(R) \cdot (D(S) \cdot \Psi) = D(R) \cdot \Psi \cdot S = (\Psi \cdot R) \cdot S = D(RS) \cdot \Psi$ . It is *not* correct to say that this should equal  $D(R) \cdot (\Psi \cdot S) \stackrel{?}{=} (\Psi \cdot S) \cdot R$  because the definitions (2.18) *only* hold for the given wave function  $\Psi$ , *not* for  $\Psi \cdot S$ .

As a simple example take the three functions

$$\psi_1(\vec{x}) = x_1 f(r) , \quad \psi_2(\vec{x}) = x_2 f(r) , \quad \psi_3(\vec{x}) = x_3 f(r) , \quad (2.22)$$

where  $f(r)$  only depends on the radius  $r = \sqrt{\vec{x}^2}$ , which is rotationally invariant. These may be, for instance, three different solutions of the Schrödinger equation (2.2). Upon a rotation, these three functions transform with a matrix  $D(R)$  that happens to coincide with  $R$  itself. The condition (2.21) is trivially obeyed.

However, the above conclusion may not always hold. According to quantum mechanics, two wave functions that only differ by a factor with absolute value equal to 1, must describe the same physical situation. The wave functions  $\psi$  and  $e^{i\alpha}\psi$  describe the same physical situation, assuming  $\alpha$  to be real. This leaves us the possibility of a certain multivaluedness in the definition of the matrices  $D(R)$ . In principle, therefore, the condition (2.21) can be replaced by a weaker condition

$$D(R_1) D(R_2) = \exp [i\alpha(R_1, R_2)] D(R_1 R_2) , \quad (2.23)$$

where  $\alpha$  is a real phase angle depending on  $R_1$  and  $R_2$ . Matrices  $D(R)$  obeying (2.23) with a non-trivial phase factor form what we call a *projective representation*. Projective representations indeed occur in physics. We shall discover circumstances where every matrix  $R$  of the rotation group is associated to *two* matrices  $D(R)$  en  $D'(R)$ , differing from one another by a phase factor, to wit, a factor  $-1$ . One has  $D'(R) = -D(R)$ . This is admitted because the wave functions  $\psi$  and  $-\psi$  describe the same physical situation. This multivaluedness implies that the relation (2.21) is obeyed only up to a sign, so that the phase angle  $\alpha$  in (2.23) can be equal to 0 or  $\pi$ . Particles described by wave functions transforming according to a projective representation, have no analogue in classical mechanics. Examples of such particles are the electron, the proton and the neutron. Their wave functions will transform in a more complicated way than what is described in Eq. (2.10). We shall return to this topic (Chapter 6).

The physical interpretation of the quantum wave function has another implication, in the form of an important constraint that the matrices  $D(R)$  must obey. A significant role is attributed to the *inner product*, a mapping that associates a complex number to a pair of wave functions,  $\psi_1$  and  $\psi_2$ , to be written as  $\langle \psi_1 | \psi_2 \rangle$ , and obeying the following relations (see Appendix E):

$$\begin{aligned} \langle \psi | \psi \rangle &\geq 0 , \\ \langle \psi | \psi \rangle &= 0 , \quad \text{then and only then if } |\psi\rangle = 0 , \end{aligned} \quad (2.24)$$

$$\langle \psi_1 | \lambda \psi_2 + \mu \psi_3 \rangle = \lambda \langle \psi_1 | \psi_2 \rangle + \mu \langle \psi_1 | \psi_3 \rangle , \quad (2.25)$$

for every pair of complex numbers  $\lambda$  and  $\mu$ ,

$$\langle \psi_1 | \psi_2 \rangle^* = \langle \psi_2 | \psi_1 \rangle . \quad (2.26)$$

For wave functions depending on just one coordinate, such an inner product is defined by

$$\langle \psi_1 | \psi_2 \rangle = \int_{-\infty}^{\infty} dx \psi_1^*(x) \psi_2(x) , \quad (2.27)$$

but for our purposes the exact definition of the inner product is immaterial.

According to quantum mechanics, the absolute value of the inner product is to be interpreted as a probability. More explicitly, consider the state described by  $|\psi\rangle$ . The probability that a measurement will establish the system to be in the state  $|\varphi\rangle$  is given by  $|\langle\varphi|\psi\rangle|^2$ . Now subject the system, including the measurement device, to a rotation. According to (2.17), the states will change into

$$|\psi\rangle \rightarrow D|\psi\rangle, \quad |\varphi\rangle \rightarrow D|\varphi\rangle. \quad (2.28)$$

The corresponding change of the inner product is then

$$\langle\varphi|\psi\rangle \longrightarrow \langle\varphi|D^\dagger D|\psi\rangle. \quad (2.29)$$

However, if nature is invariant under rotations, the probability described by the inner product, should *not* change under rotations. The two inner products in (2.29) must be equal. Since this equality must hold for all possible pairs of states  $|\psi\rangle$  and  $|\varphi\rangle$ , we can conclude that the matrices themselves must obey the following condition:

$$D^\dagger D = \mathbf{1}, \quad (2.30)$$

in other words,  $D$  must be a *unitary* matrix.<sup>5</sup> Since this has to hold for every matrix  $D(R)$  associated to a rotation, this demand should hold for the entire representation. Thus, in this context, we shall be exclusively interested in *unitary* representations.

---

<sup>5</sup>The condition is that the absolute value of the inner product should not change, so one might suspect that it suffices to constrain  $D^\dagger D$  to be equal to unity apart from a phase factor. However,  $D^\dagger D$  is a hermitian, positive definite matrix, so we must conclude that this phase factor can only be equal to 1.

### 3. The group of rotations in three dimensions

A rotation in three-dimensional space can be represented by a  $3 \times 3$  matrix of real numbers. Since upon a rotation of a set of vectors, the angles between them remain the same, the matrix in question will be *orthogonal*. These orthogonal matrices form a group, called  $O(3)$ . From the demand  $R\tilde{R} = \mathbf{1}$ , one derives that  $\det(R) = \pm 1$ . If we restrict ourselves to the orthogonal matrices with  $\det(R) = +1$ , then we call the group  $SO(3)$ , the *special orthogonal group in 3 dimensions*.

A rotation in three-dimensional space is completely determined by the rotation axis and the angle over which we rotate. The rotation axis can for instance be specified by a three-dimensional vector  $\vec{\alpha}$ ; the length of this vector can then be chosen to be equal to the angle over which we rotate (in radians). Since rotations over angles that differ by a multiple of  $2\pi$ , are identical, we can limit ourselves to rotation axis vectors  $\vec{\alpha}$  inside (or on the surface of) a three-dimensional sphere with radius  $\pi$ . This gives us a natural parametrization for all rotations. Every point in this sphere of parameters corresponds to a possible rotation: the rotation axis is given by the line through this point and the center of the sphere, and the angle over which we rotate (according to a left-handed screw for instance) varies from 0 to  $\pi$  (rotations over angles between  $-\pi$  and 0 are then associated with the vector in the opposite direction). Two opposite points on the surface of the sphere, that is,  $\vec{\alpha}$  and  $-\vec{\alpha}$  with  $|\vec{\alpha}| = \pi$ , describe the same rotation, one over an angle  $\pi$  and one over an angle  $-\pi$ , around the same axis of rotation. However, apart from this identification of diametrically opposed points on the surface of the sphere, two different points inside this parameter sphere always describe two different rotations.

From the above, it is clear that rotations can be parameterized in terms of three independent parameters, being the three components of the vectors  $\vec{\alpha}$ , and furthermore that the rotations depend on these parameters in a continuous fashion. To study this dependence further, consider infinitesimal rotations, or, rotations corresponding to vectors  $|\vec{\alpha}| \approx 0$ . First, let us limit ourselves to rotations around the  $z$  axis, so that  $\vec{\alpha} = (0, 0, \alpha)$ . The associated rotation follows from

$$\begin{aligned}x &\rightarrow \cos \alpha x + \sin \alpha y, \\y &\rightarrow \cos \alpha y - \sin \alpha x, \\z &\rightarrow z.\end{aligned}\tag{3.1}$$

This leads to a matrix  $R(\alpha)$ , equal to

$$R(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.\tag{3.2}$$

The rotation by an angle  $\alpha$  can also be regarded as being the result of  $n$  successive rotations over an angle  $\alpha/n$ . For very large values of  $n$ , the rotation by a small angle  $\alpha/n$  will differ from the identity only infinitesimally; ignoring terms of order  $(\alpha/n)^2$ , we

find for the associated  $3 \times 3$  matrix,

$$\begin{aligned} R(\alpha/n) &= \begin{pmatrix} 1 & \alpha/n & 0 \\ -\alpha/n & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + O\left(\frac{\alpha^2}{n^2}\right) \\ &= \mathbf{1} + \frac{\alpha}{n} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + O\left(\frac{\alpha^2}{n^2}\right). \end{aligned} \quad (3.3)$$

It is now possible to reconstruct the finite rotation over an angle  $\alpha$  by taking the  $n^{\text{th}}$  power of (3.3),

$$R(\alpha) = [R(\alpha/n)]^n = \left[ \mathbf{1} + \frac{\alpha}{n} T + O\left(\frac{\alpha^2}{n^2}\right) \right]^n, \quad (3.4)$$

where the matrix  $T$  is given by

$$T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.5)$$

In the limit  $n \rightarrow \infty$ , we expect to be able to ignore terms of order  $1/n^2$ ; furthermore, we make use of the formula

$$e^A = \lim_{n \rightarrow \infty} \left( \mathbf{1} + \frac{1}{n} A \right)^n. \quad (3.6)$$

This results in

$$R(\alpha) = \exp(\alpha T). \quad (3.7)$$

The exponent of this matrix can be elaborated by using the series expansion

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \quad (3.8)$$

Next, we remark that

$$T^{2n} = (-)^n \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (n \geq 1) \quad (3.9)$$

from which it follows immediately that  $T^{2n+1} = (-)^n T$  for  $n \geq 0$ . Using this, we can perform the exponentiation by separately selecting the even and odd powers. This leads to

$$\begin{aligned} \exp(\alpha T) &= \mathbf{1} + \sum_{n=1}^{\infty} \frac{(-)^n \alpha^{2n}}{(2n)!} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sum_{n=0}^{\infty} \frac{(-)^n \alpha^{2n+1}}{(2n+1)!} T \\ &= \mathbf{1} + (\cos \alpha - 1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sin \alpha T, \end{aligned} \quad (3.10)$$

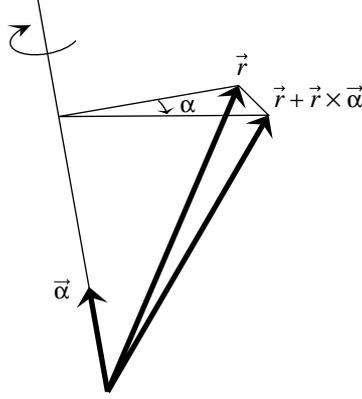


Figure 2: Infinitesimal rotation of a vector  $\vec{r}$ , around a rotation axis  $\vec{\alpha}$

which indeed coincides with the original matrix (3.2).

Let us now consider the relation between finite and infinitesimal transformations as given by Eq. (3.7), for more general rotations. For rotations over a small angle, every  $\vec{r}$  gets a small vector added to it that is orthogonal both to the vector  $\vec{r}$  and the rotation axis (see Figure 2). This tiny vector is exactly equal to the outer product of  $\vec{r}$  and the rotation axis vector  $\vec{\alpha}$  (where it was assumed that  $|\vec{\alpha}| \approx 0$ ), so that

$$\vec{r} \rightarrow \vec{r} + \vec{r} \times \vec{\alpha} + O(|\vec{\alpha}|^2) . \quad (3.11)$$

therefore, in case of a general rotation axis vector  $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$  one can write

$$\begin{aligned} x &\rightarrow x + \alpha_3 y - \alpha_2 z + O(|\vec{\alpha}|^2) , \\ y &\rightarrow y + \alpha_1 z - \alpha_3 x + O(|\vec{\alpha}|^2) , \\ z &\rightarrow z + \alpha_2 x - \alpha_1 y + O(|\vec{\alpha}|^2) . \end{aligned} \quad (3.12)$$

Infinitesimal rotations can therefore be written as follows:

$$R(\vec{\alpha}) = \mathbf{1} + i(\alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3) + O(|\vec{\alpha}|^2) , \quad (3.13)$$

where we added a factor  $i$  in order to conform to the usual notations, and the hermitian matrices  $L_1$ ,  $L_2$  en  $L_3$  are defined by

$$\begin{aligned} L_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} , \\ L_2 &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} , \\ L_3 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \end{aligned} \quad (3.14)$$

Above result can be compressed in one expression by using the completely skew-symmetric epsilon tensor,

$$(L_i)_{jk} = -i\epsilon_{ijk} . \quad (3.15)$$

Indeed, we can easily check that

$$\begin{aligned} (L_1)_{23} &= -(L_1)_{32} = -i\epsilon_{123} = -i , \\ (L_2)_{31} &= -(L_2)_{13} = -i\epsilon_{231} = -i , \\ (L_3)_{12} &= -(L_3)_{21} = -i\epsilon_{312} = -i . \end{aligned} \quad (3.16)$$

Again, we can consider  $R(\vec{\alpha})$  as being formed out of  $n$  successive rotations with rotation axis  $\vec{\alpha}/n$ ,

$$\begin{aligned} R(\vec{\alpha}) &= [R(\vec{\alpha}/n)]^n \\ &= \left[ \mathbf{1} + \frac{1}{n} (i\alpha_1 L_1 + i\alpha_2 L_2 + i\alpha_3 L_3) + \mathcal{O}\left(\frac{|\vec{\alpha}|^2}{n^2}\right) \right]^n . \end{aligned} \quad (3.17)$$

Employing (3.4), we find then the following expression in the limit  $n \rightarrow \infty$ ,

$$R(\vec{\alpha}) = \exp\left(i \sum_k \alpha_k L_k\right) . \quad (3.18)$$

The correctness of Eq. (3.18) can be checked in a different way. First, we note that the following multiplication rule holds for rotations around one common axis of rotation, but with different rotation angles:

$$R(s\vec{\alpha}) R(t\vec{\alpha}) = R((s+t)\vec{\alpha}) , \quad (3.19)$$

where  $s$  and  $t$  are real numbers. The rotations  $R(s\vec{\alpha})$  with one common axis of rotation define a commuting subgroup of the complete rotation group. This is not difficult to see: The matrices  $R(s\vec{\alpha})$  (with a fixed vector  $\vec{\alpha}$  and a variable  $s$ ) define a group, where the result of a multiplication does not depend on the order in the product,

$$R(s\vec{\alpha}) R(t\vec{\alpha}) = R(t\vec{\alpha}) R(s\vec{\alpha}) . \quad (3.20)$$

This subgroup is the group  $SO(2)$ , the group of the two-dimensional rotations (the axis of rotation stays the same under these rotations, only the components of a vector that are orthogonal to the axis of rotation are rotated). Using Eq. (3.19), we can simply deduce the following differential equation for  $R(s\vec{\alpha})$ ,

$$\begin{aligned} \frac{d}{ds} R(s\vec{\alpha}) &= \lim_{\Delta \rightarrow 0} \frac{R((s+\Delta)\vec{\alpha}) - R(s\vec{\alpha})}{\Delta} \\ &= \lim_{\Delta \rightarrow 0} \frac{R(\Delta\vec{\alpha}) - \mathbf{1}}{\Delta} R(s\vec{\alpha}) \\ &= \left( i \sum_k \alpha_k L_k \right) R(s\vec{\alpha}) , \end{aligned} \quad (3.21)$$

where first Eq. (3.19) was used, and subsequently (3.13). Now it is easy to verify that the solution of this differential equation is exactly given by Eq. (3.18).

Yet another way to ascertain that the matrices (3.18) represent rotations, is to prove that these matrices are orthogonal and have determinant equal to 1, which means that the following relations are fulfilled

$$\widetilde{R(\vec{\alpha})} = [R(\vec{\alpha})]^{-1} = R(-\vec{\alpha}) , \quad \det R(\vec{\alpha}) = 1 , \quad (3.22)$$

The proof follows from the following properties for a general matrix  $A$  (see also Appendix C),

$$\widetilde{(e^A)} = e^{\widetilde{A}} , \quad \det (e^A) = e^{\text{Tr } A} . \quad (3.23)$$

From this, it follows that the matrices (3.18) obey Eqs. (3.22) provided that the matrix  $i \sum_k \alpha_k L_k$  be real and skew-symmetric. This indeed turns out to be the case; from the definitions (3.15) it follows that  $i \sum_k \alpha_k L_k$  in fact represents the most general real, and skew-symmetric  $3 \times 3$  matrix.

The above question may actually be turned around: can *all* rotations be written in the form of Eq. (3.18)? The answer to this question is not quite so easy to give. In principle, the exponentiation in (3.18) can be performed explicitly *via* the power series expansion (3.8), and the result can be compared with the most general rotation matrix. It will turn out that the answer is affirmative: all rotations can indeed be written in the form of Eq. (3.18). This, however, is not the case for all groups. The so-called *non-compact* groups contain elements that cannot be written as a product of a finite number of such exponentials. These groups are called non-compact, because the volume of parameter space is non-compact. The rotation group, where all possible group elements are defined in terms of the parameters  $\alpha_k$  that are restricted to the insides of a sphere with radius  $\pi$ , is a compact group. Within the frame of these lectures, non-compact groups will play no role, but such groups are not unimportant in physics. The Lorentz group, for example, which is the group consisting of all Lorentz transformations, is an example of a non-compact group.

From the preceding discussion it will be clear that the matrices  $L_k$ , associated with the infinitesimal transformations, will be important, and at least for the compact groups, they will completely determine the group elements, by means of the exponentiation (3.18). This is why these matrices are called the *generators* of the group. Although our discussion was confined to the rotation group, the above can be applied to all *Lie groups*<sup>6</sup>: a group whose elements depend analytically on a finite number of parameters, in our case  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ . In the case that the group elements take the form of matrices, this means that the matrix elements must be differentiable functions of the parameters.<sup>7</sup> The number of linearly independent parameters defines the *dimension* of the Lie group, not to be confused

---

<sup>6</sup>Named after the Norwegian mathematician Sophus Lie, 1842-1899

<sup>7</sup>This is clearly the case for the rotation group. In the general case, the above requirement can be somewhat weakened; for a general Lie group it suffices to require the elements as functions of the parameters to be twice differentiable.

with the dimension of the matrices considered.<sup>8</sup> The number of linearly independent generators must obviously be equal to the dimension of the group.

One of the most essential ingredients of a group, is its multiplication structure, according to which the product of two rotations  $R(\vec{\alpha})$  and  $R(\vec{\beta})$ , again should be a rotation,

$$R(\vec{\alpha}) R(\vec{\beta}) = R(\vec{\gamma}) , \quad (3.24)$$

where  $\vec{\gamma}$  depends on  $\vec{\alpha}$  and  $\vec{\beta}$ . The exact dependence fixes the multiplication structure of the group. The fact that such a vector function  $\vec{\gamma}(\vec{\alpha}, \vec{\beta})$  must exist, has implications for the product of generators. To derive these, we expand (3.24) in powers<sup>9</sup> of  $\alpha$  en  $\beta$ ,

$$\begin{aligned} e^{i\vec{\alpha}\cdot\vec{L}} e^{i\vec{\beta}\cdot\vec{L}} &= \left( \mathbf{1} + i\alpha_k L_k + O(\alpha^2) \right) \left( \mathbf{1} + i\beta_l L_l + O(\beta^2) \right) \\ &= \mathbf{1} + i(\alpha + \beta)_k L_k - \alpha_k \beta_l L_k L_l + O(\alpha^2) + O(\beta^2) \\ &= \mathbf{1} + i(\alpha + \beta)_k L_k - \frac{1}{2}(\alpha + \beta)_k (\alpha + \beta)_l L_k L_l \\ &\quad - \frac{1}{2}\alpha_k \beta_l [L_k, L_l] + O(\alpha^2) + O(\beta^2) . \end{aligned} \quad (3.25)$$

The first three terms are recognized as the beginning of the power series of  $\exp(i(\vec{\alpha} + \vec{\beta})\vec{L})$ . If the fourth term would vanish, that is, if the matrices  $L_k$  and  $L_l$  commute, then indeed  $\gamma_k = \alpha_k + \beta_k$ . However, it will turn out that the generators of the rotation group do *not* commute. Since it must be possible in any case to write the r.h.s. of the equation again in the form of the power series for  $\exp(i\vec{\gamma}\cdot\vec{L})$ , it must be possible to rewrite the commutators of the generators in terms of some linear combination of the generators. In other words, we must have

$$[L_i, L_j] = c_{ij}^k L_k , \quad (3.26)$$

where the constants  $c_{ij}^k$  are called the *structure constants* of the group, because they (nearly) completely determine the multiplication structure of the group. Note that, since the generators  $L_k$  are hermitian, the structure constants must be purely imaginary.

Before continuing, we first verify whether the generators (3.15) obey to the demand (3.26). After explicit matrix multiplications, we find this indeed to be the case:

$$[L_1, L_2] = iL_3 , \quad [L_2, L_3] = iL_1 , \quad [L_3, L_1] = iL_2 , \quad (3.27)$$

or,

$$[L_i, L_j] = i\epsilon_{ijk} L_k . \quad (3.28)$$

Making use of Eq. (3.26), we can now deduce the following result for  $\vec{\gamma}(\vec{\alpha}, \vec{\beta})$ :

$$\gamma_k = \alpha_k + \beta_k + \frac{i}{2}c_{mn}^k \alpha_m \beta_n + O(\alpha^2) + O(\beta^2) . \quad (3.29)$$

---

<sup>8</sup>For the rotation group in three dimensions the dimension of the group and that of the matrices are both 3, but this is a coincidence: the dimension of the rotation group in  $d$  dimensions is  $\frac{1}{2}d(d-1)$ .

<sup>9</sup>The notation  $\vec{\alpha}\cdot\vec{L}$  is here intended to mean  $\alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3$ . In Eq. (3.25) we also used summation convention: if in one term of an expression an index occurs twice, this means that it is summed over, even if the summation sign is not explicitly shown. So,  $\alpha_k L_k \equiv \sum_k \alpha_k L_k$ . From now on, this convention will be frequently used.

In principle, the higher order contributions can be determined by means of iteration; for example, we find

$$\gamma_k = \alpha_k + \beta_k + \frac{i}{2} c_{mn}^k \alpha_m \beta_n - \frac{1}{12} (\alpha_m \alpha_n \beta_p + \beta_m \beta_n \alpha_p) c_{mq}^k c_{np}^q + \dots \quad (3.30)$$

The fact that all terms in this iteration can be expressed in terms of the structure constants follows from the Campbell-Baker-Hausdorff formula, which expresses the logarithm of  $(\exp A \exp B)$  in terms of a power series consisting exclusively of repeated commutators of the matrices  $A$  and  $B$ . Thus, the multiplication structure of the group is determined by the structure constants (at least for all those group elements that reside in some finite domain in the neighborhood of the identity). The CBH formula is explained in Appendix D.

Imagine that we can find matrices  $A_k$ , different from the matrices  $L_k$ , obeying the same commutation relations (3.26) as the  $L_k$ . In that case, by means of exponentiation, we can determine the corresponding group elements, which will have the same multiplication rules as the elements of the original group. In other words, we find a representation of the group this way. On the other hand, for every representation of the group, we can construct the corresponding generators, using the infinitesimal transformations, and they will obey the same commutation rules (3.26), with the same structure constants. Thus, we have found a direct relation between group representations and the matrix relations (3.26) (In more mathematical terms: the generators  $L_k$ , together with the commutation relations (3.26), define an algebra, called the *Lie algebra*. Matrices  $A_k$  with the same commutation relations then define a *representation* of the Lie algebra.)

One can easily check that the structure constants also must obey certain relations. This follows from the so-called *Jacobi identity*, which holds for any triple of matrices  $A$ ,  $B$  and  $C$ ,

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \quad (3.31)$$

This identity can be proven by explicitly writing the commutators and using the associativity of the multiplication (See chapter 2); one then obtains 12 terms that cancel out pairwise. Using the Jacobi identity with  $A = L_i$ ,  $B = L_j$  en  $C = L_k$ , we deduce the following equation for the structure constants,

$$c_{ij}^m c_{mk}^n + c_{jk}^m c_{mi}^n + c_{ki}^m c_{mj}^n = 0 \quad (3.32)$$

where use was made of (3.26). The equation (3.32) is also called the Jacobi identity. For the rotation group, this implies the following equation for the  $\epsilon$ -tensors:

$$\epsilon_{ijm} \epsilon_{mkn} + \epsilon_{jkm} \epsilon_{min} + \epsilon_{kim} \epsilon_{mjn} = 0 \quad (3.33)$$

which will be frequently used later. The validity of Eq. (3.33) can be derived directly from the identity

$$\epsilon_{ijm} \epsilon_{mkl} = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk} \quad (3.34)$$

which is easy to prove (for instance by choosing a couple of values for two of the indices).

Equation (3.32) has another consequence. Let us define  $n \times n$  matrices  $C_i$  according to

$$(C_i)_j^k \equiv -c_{ij}^k, \quad (3.35)$$

where  $n$  is the dimension of the Lie group. We can then write (3.32) as

$$(c_{ij}^m C_m)_k^n + (C_j C_i)_k^n - (C_i C_j)_k^n = 0, \quad \text{or} \quad C_i C_j - C_j C_i = c_{ij}^k C_k. \quad (3.36)$$

These are exactly the same commutation relations as the ones we used to define the structure constants, in Eq. (3.26). The matrices  $C_i$  thus define a representation of the Lie algebra based on (3.26). Through exponentiation of the matrices  $C_i$ , we can then define a group with the same multiplication properties (at least in some finite region surrounding the identity) as the original Lie group, consisting of  $n \times n$  matrices, where  $n$  is the dimension of the Lie group. This representation is called the *adjoint representation*.

Applying the above to the case of the rotation group leads to something of a disappointment. Since in this case  $c_{ij}^k = i\epsilon_{ijk}$ , the matrices  $C_i$  are simply equal to the matrices  $L_i$  (see Eq. (3.15), and so we recovered the original three-dimensional rotations. The adjoint representation thus coincides with the original group. This, however, is rather the exception than the rule, as will be seen later.

## 4. More about representations

In the previous chapter the properties of the group of three-dimensional rotations were discussed. Now, we return to the representations of this group. First, we note that, starting from a given representation, for instance by the matrices  $D$  acting on the wave functions that we combined in a column vector  $\psi$ , we can obtain an other representation, by constructing an other vector  $\hat{\psi}$ . For instance, rearrange  $\psi$  in wave functions  $\hat{\psi}$  according to

$$\hat{\psi} = U\psi . \quad (4.1)$$

Under rotations,  $\hat{\psi}$  then transforms according to

$$\hat{\psi} \rightarrow \hat{\psi}' = \hat{D}\hat{\psi} , \quad (4.2)$$

where  $\hat{D}$  is given by

$$\hat{D} = UDU^{-1} . \quad (4.3)$$

Both the original matrices  $D$  and the matrices  $\hat{D}$  define a representation of the rotation group, but such representations will not be considered as fundamentally different. This is why representations that are related according to (4.3), are called *equivalent representations*. This allows us to formulate an important result in representation theory:

*All finite dimensional representations of finite or compact groups are unitary.*

With this we mean that all representations can be chosen to be unitary *via* a redefinition (4.3), so that all matrices  $D$  belonging to the representation obey  $D^\dagger = D^{-1}$ . We will not prove this here.

Up to here, we have primarily discussed one special representation of the group of rotations, being the representation defined by rotating the three-dimensional vector  $\vec{x} = (x_1, x_2, x_3)$ . There is an easy way to construct larger representations: just consider two vectors,  $\vec{x}$  and  $\vec{y}$ , both transforming the usual way under rotations. Together, they form a six-dimensional vector  $\vec{z} = (x_1, x_2, x_3, y_1, y_2, y_3)$ , transforming under rotations as

$$\vec{z} \rightarrow \vec{z}' = D\vec{z} , \quad (4.4)$$

where the matrix  $D$  can be decomposed in  $3 \times 3$  matrices in the following way:

$$D = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} . \quad (4.5)$$

Such a representation is called *reducible*, because the six-dimensional space can be split up in two invariant three-dimensional subspaces. This reducible six-dimensional representation can therefore be regarded as the *direct sum* of two three-dimensional representations, and we write

$$\mathbf{6} = \mathbf{3} \oplus \mathbf{3} . \quad (4.6)$$

The sum representation can occur if we consider a particle that can be in a superposition of two different kinds of quantum states.

It will be clear that representations that do *not* leave any subspace invariant, and therefore cannot be described in a block diagonal form such as in Eq. (4.5), are considered to be *irreducible representations*.

Other representations can be obtained by constructing so-called *product representations*. Consider for instance a system of two (free) particles with wave functions  $\psi_1(\vec{x})$  and  $\psi_2(\vec{y})$ , where  $\vec{x}$  and  $\vec{y}$  are the coordinates of the particles. The wave functions  $\Psi(\vec{x}, \vec{y})$  of the combined system then consist of all possible products of wave functions  $\psi_1$  and  $\psi_2$ . We call this a *tensor product*, which is denoted by

$$\Psi = \psi_1 \otimes \psi_2 . \quad (4.7)$$

Under rotations of both  $\vec{x}$  and  $\vec{y}$ , this  $\Psi$  transforms accordingly, but the corresponding representation is more complicated than the ones associated to the separate wave functions  $\psi_1$  and  $\psi_2$ . Often, such a product representation is not irreducible, and can be decomposed into a number of distinct representations that are irreducible. Let us demonstrate this phenomenon first in the following example. Let three possible functions  $\psi_i^1$  be given by the coordinates  $x_i$  and three possible functions  $\psi_j^2$  by the coordinates  $y_j$ . Thus, both the  $\psi_i^1$ 's and the  $\psi_j^2$ 's transform according to the three-dimensional representation of the rotation group. The product representation works on all possible products of  $\psi_i^1$  and  $\psi_j^2$ , and therefore we can distinguish nine independent functions,

$$T_{ij}(\vec{x}, \vec{y}) = x_i y_j , \quad (4.8)$$

transforming under rotations as

$$T_{ij} \rightarrow T'_{ij} = R_{ii'} R_{jj'} T_{i'j'} . \quad (4.9)$$

This nine-dimensional representation however is not irreducible. For instance, the symmetric part and the skew-symmetric part of  $T_{ij}$ , defined by  $T_{(ij)} \equiv \frac{1}{2}(T_{ij} + T_{ji})$ , and  $T_{[ij]} \equiv \frac{1}{2}(T_{ij} - T_{ji})$ , transform separately and independently under rotations. This follows directly by restricting ourselves only to the (skew-)symmetric part of  $T'_{ij}$ , and observing that the (anti)symmetry in  $i$  and  $j$  of  $\frac{1}{2}(R_{ii'} R_{jj'} \pm R_{j'i'} R_{i'j'})$  implies the (anti)symmetry in  $i'$  en  $j'$ . This is why we write

$$T_{(ij)} \rightarrow T'_{(ij)} = R_{ii'} R_{jj'} T_{(i'j')} , \quad T_{[ij]} \rightarrow T'_{[ij]} = R_{ii'} R_{jj'} T_{[i'j']} . \quad (4.10)$$

The skew-symmetric part of  $T_{ij}$  contains three independent components, transforming as a three-dimensional representation of the rotation group. The symmetric part of  $T_{ij}$  contains the remaining six components, which however do not transform as an irreducible transformation. This follows immediately from the fact that the trace of  $T_{ij}$  is equal to

$$T_{ii} = \vec{x} \cdot \vec{y} , \quad (4.11)$$

and therefore invariant under rotations. We must conclude that  $T_{ij}$  can be decomposed

in three independent tensors<sup>10</sup>,

$$T_{ij} \rightarrow \left\{ \begin{array}{l} T = \vec{x} \cdot \vec{y} \\ T_i = \epsilon_{ijk} x_j y_k \\ S_{ij} = x_i y_j + x_j y_i - \frac{2}{3} \delta_{ij} (\vec{x} \cdot \vec{y}) \end{array} \right\} . \quad (4.12)$$

Note that we used the epsilon symbol to describe the skew-symmetric part of  $T_{ij}$  again as a three-dimensional vector  $\vec{T}$  (it is nothing but the outer product  $\vec{x} \times \vec{y}$ ). Furthermore, we made the symmetric part  $S_{ij}$  traceless by adding an extra term proportional to  $\delta_{ij}$ . The consequence of this is that  $S_{ij}$  consists of only *five* independent components. Under rotations, the terms listed above transform into expressions of the same type; the five independent components of  $S_{ij}$  transform into one another.<sup>11</sup> In short, the product of two three-dimensional representations can be written as

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5} , \quad (4.13)$$

where the representations are characterized by their dimensions (temporarily ignoring the fact that inequivalent irreducible representations might exist with equal numbers of dimensions; they don't here, as we will see later).

The procedure followed in this example, rests on two features; first, we use that the symmetry properties of tensors do not change under the transformations, and secondly we make use of the existence of two invariant tensors, to wit:

$$T_{ij} = \delta_{ij} , \quad T_{ijk} = \epsilon_{ijk} . \quad (4.14)$$

An invariant tensor is a tensor that does not change at all under the group transformations, as they act according to the index structure of the tensor, so that

$$T_{ijk\dots} \rightarrow T'_{ijk\dots} = R_{ii'} R_{jj'} R_{kk'} \cdots T_{i'j'k'\dots} = T_{ijk\dots} . \quad (4.15)$$

Indeed, both tensors  $\delta_{ij}$  and  $\epsilon_{ijk}$  obey (4.15), since the equation

$$R_{ii'} R_{jj'} \delta_{i'j'} = \delta_{ij} \quad (4.16)$$

is fulfilled because the  $R_{ij}$  are orthogonal matrices, and

$$R_{ii'} R_{jj'} R_{kk'} \epsilon_{i'j'k'} = \det R \epsilon_{ijk} = \epsilon_{ijk} \quad (4.17)$$

<sup>10</sup>In the second equation, again summation convention is used, see an earlier footnote.

<sup>11</sup>For each of these representations, we can indicate the matrices  $D(R)$  that are defined in chapter 2. For the first representation, we have that  $D(R) = 1$ . In the second representation, we have  $3 \times 3$  matrices  $D(R)$  equal to the matrix  $R$ . For the third representation, we have  $5 \times 5$  matrices  $D(R)$ . The indices of this correspond to the symmetric, traceless index pairs  $ij$ . The matrices  $D(R)$  can be written as

$$D(R)_{(ij)(kl)} = \frac{1}{2} (R_{ik} R_{jl} + R_{il} R_{jk}) - \frac{1}{3} \delta_{ij} \delta_{kl} .$$

holds because the rotation matrices  $R_{ij}$  have  $\det R = 1$ . For every given tensor  $T_{ijk\dots}$  we can contract the indices using invariant tensors. It is then evident that tensors contracted that way span invariant subspaces, in other words, under rotations they will transform into tensors that are formed the same way. For example, let  $T_{ijk\dots}$  be a tensor transforming like

$$T_{ijk\dots} \rightarrow T'_{ijk\dots} = R_{ii'} R_{jj'} R_{kk'} \cdots T_{i'j'k'\dots} . \quad (4.18)$$

Now, form the tensor

$$\hat{T}_{klm\dots} \equiv \delta_{ij} T_{ijklm\dots} \quad (4.19)$$

which has two indices less. By using Eq. (4.16), it is now easy to check that  $\hat{T}$  transforms as

$$\hat{T}_{klm\dots} \rightarrow \hat{T}'_{klm\dots} = R_{kk'} R_{ll'} R_{mm'} \cdots \hat{T}_{k'l'm'\dots} , \quad (4.20)$$

and, in a similar way, we can verify that contractions with one or more  $\delta$  and  $\epsilon$  tensors, produce tensors that span invariant subspaces. Using the example discussed earlier, we can write the expansion as

$$T_{ij} = \frac{1}{2} \epsilon_{ijk} (\epsilon_{klm} T_{lm}) + \frac{1}{2} \left( T_{ij} + T_{ji} - \frac{2}{3} \delta_{ij} T_{kk} \right) + \frac{1}{3} \delta_{ij} T_{kk} , \quad (4.21)$$

where the first term can also be written as  $\frac{1}{2}(T_{ij} - T_{ji})$ , by using the identity (3.34),

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} , \quad (4.22)$$

and the second term in (4.21) is constructed in such a way that it is traceless:

$$\delta_{ij} \left( T_{ij} + T_{ji} - \frac{2}{3} \delta_{ij} T_{kk} \right) = 0 . \quad (4.23)$$

## 5. Ladder operators

Let us consider a representation of the rotation group, generated by hermitian matrices  $I_1$ ,  $I_2$  and  $I_3$ , which obey the same commutation rules as  $L_1$ ,  $L_2$  and  $L_3$ , given in Eq. (3.15),

$$[I_1, I_2] = iI_3, \quad [I_2, I_3] = iI_1, \quad [I_3, I_1] = iI_2, \quad (5.1)$$

or in shorthand:

$$[I_i, I_j] = i\epsilon_{ijk} I_k. \quad (5.2)$$

We demand the matrices  $\exp(i\alpha_k I_k)$  to be unitary; therefore, the  $I_i$  are hermitian:  $I_i^\dagger = I_i$ . Starting from this information, we now wish to determine all sets of irreducible matrices  $I_i$  with these properties. This is the way to determine all (finite-dimensional, unitary) representations of the group of rotations in three dimensions.

To this end, first define the linear combinations

$$I_\pm = I_1 \pm iI_2, \quad (5.3)$$

so that  $(I_\pm)^\dagger = I_\mp$ , and

$$[I_3, I_\pm] = [I_3, I_1] \pm i[I_3, I_2] = iI_2 \pm I_1 = \pm I_\pm. \quad (5.4)$$

So we have for any state  $|\psi\rangle$ ,

$$I_3 \left( I_+ |\psi\rangle \right) = I_+ (I_3 + 1) |\psi\rangle. \quad (5.5)$$

A *Casimir operator* is a combination of operators for a representation constructed in such a way that it commutes with all generators. *Schur's lemma* states the following: if and only if the representation is irreducible, every Casimir operator will be a multiple of the unit matrix.

In the case of the three-dimensional rotations, we have such a Casimir operator:

$$\vec{I}^2 \equiv I_1^2 + I_2^2 + I_3^2. \quad (5.6)$$

We derive from Eq. (5.1):

$$[\vec{I}^2, I_1] = [\vec{I}^2, I_2] = [\vec{I}^2, I_3] = 0. \quad (5.7)$$

Since  $\vec{I}^2$  and  $I_3$  are two commuting matrices, we can find a basis of states such that  $\vec{I}^2$  and  $I_3$  both at the same time take a diagonal form, with real eigenvalues. Furthermore, the eigenvalues of  $\vec{I}^2$  must be positive (or zero), because we have

$$\langle \psi | \vec{I}^2 | \psi \rangle = \left| I_1 | \psi \rangle \right|^2 + \left| I_2 | \psi \rangle \right|^2 + \left| I_3 | \psi \rangle \right|^2 \geq 0. \quad (5.8)$$

It will turn out to be convenient to write the eigenvalues of  $\vec{I}^2$  as  $\ell(\ell + 1)$ , where  $\ell \geq 0$  (The reason for this strange expression will become clear shortly; for the time being, consider this merely as a notation).

Now, consider a state  $|\ell, m\rangle$  that is an eigenstate of  $\vec{I}^2$  and  $I_3$ , with eigenvalues  $\ell(\ell + 1)$  and  $m$ ,

$$\vec{I}^2 |\ell, m\rangle = \ell(\ell + 1) |\ell, m\rangle, \quad I_3 |\ell, m\rangle = m |\ell, m\rangle. \quad (5.9)$$

From Eqs. (5.5) and (5.7), one derives that

$$\begin{aligned} I_3 (I_+ |\ell, m\rangle) &= (m + 1) (I_+ |\ell, m\rangle), \\ \vec{I}^2 (I_+ |\ell, m\rangle) &= \ell(\ell + 1) (I_+ |\ell, m\rangle). \end{aligned} \quad (5.10)$$

Substituting  $I_+ |\ell, m\rangle = |\psi\rangle$ , we have

$$I_3 |\psi\rangle = (m + 1) |\psi\rangle, \quad \vec{I}^2 |\psi\rangle = \ell(\ell + 1) |\psi\rangle, \quad (5.11)$$

in other words,  $|\psi\rangle$  is a new eigenvector of  $I_3$  and  $\vec{I}^2$  with eigenvalues  $m' = m + 1$ , and  $\ell' = \ell$ , unless

$$|\psi\rangle \equiv I_+ |\ell, m\rangle \stackrel{?}{=} 0. \quad (5.12)$$

Furthermore, we find

$$\begin{aligned} \langle \psi | \psi \rangle &= \langle \ell, m | I^- I^+ | \ell, m \rangle \\ &= \langle \ell, m | I_1^2 + I_2^2 + i[I_1, I_2] | \ell, m \rangle \\ &= \langle \ell, m | I_1^2 + I_2^2 - I_3 | \ell, m \rangle \\ &= \langle \ell, m | \vec{I}^2 - I_3(I_3 + 1) | \ell, m \rangle, \end{aligned} \quad (5.13)$$

where we made use of:  $I_+^\dagger = I_-$ . And so, using Eq. (5.9), we find

$$\langle \psi | \psi \rangle = \left( \ell(\ell + 1) - m(m + 1) \right) \langle \ell, m | \ell, m \rangle. \quad (5.14)$$

If we now assume that  $|\ell, m\rangle$  is a normalized state (so,  $\langle \ell, m | \ell, m \rangle = 1$ ), then  $|\psi\rangle$  can be written as a normalized state  $|\ell, m + 1\rangle$  multiplied by a proportionality factor that is given by (5.14). This factor is fixed up to a phase factor, which we absorb in the definition of  $|\ell, m + 1\rangle$ . This way, we conclude that

$$I_+ |\ell, m\rangle = \sqrt{\ell(\ell + 1) - m(m + 1)} |\ell, m + 1\rangle. \quad (5.15)$$

Repeating this procedure, the operator  $I_+$  produces states with ever increasing eigenvalues of  $I_3$ :

$$|\ell, m\rangle \xrightarrow{I_+} |\ell, m + 1\rangle \xrightarrow{I_+} |\ell, m + 2\rangle \xrightarrow{I_+} |\ell, m + 3\rangle \xrightarrow{I_+} \text{etc.} \quad (5.16)$$

This is why  $I_+$  will be called ‘‘ladder operator’’ or ‘‘step operator’’. However, we are interested in finite matrices  $I_i$ , and this implies that the series (5.16) has to come to

an end somewhere. According to Eq. (5.15), this only happens if, in the series (5.16), a state emerges for which the eigenvalue  $m$  of  $I_3$  equals  $\ell$ . This, in turn, requires that the original eigenvalue  $m$  of the state we started off with, differs from  $\ell$  by an integer. The necessity of this in fact already follows from Eq. (5.14): since  $\langle \psi | \psi \rangle$  and  $\langle \ell, m | \ell, m \rangle$  must have non negative norms, one must have  $\ell(\ell + 1) - m(m + 1) \geq 0$ , and also  $-\ell - 1 \leq m \leq \ell$ . In order to ensure that the series (5.15) terminates, as soon as  $m$  approaches values greater than its allowed limit, we must demand that  $\ell - m$  be a positive integer. therefore, we find

$$| \ell, m \rangle \xrightarrow{I_+} | \ell, m + 1 \rangle \xrightarrow{I_+} \dots \xrightarrow{I_+} | \ell, \ell \rangle , \quad (5.17)$$

where the vector  $| \ell, \ell \rangle$  with the highest eigenvalue of  $I_3$  obeys

$$I_+ | \ell, \ell \rangle = 0 . \quad (5.18)$$

It is now easy to continue by observing that the matrix  $I_-$  is also a ladder operator, but one generating lower eigenvalues of  $I_3$ . Starting from a state  $| \ell, m \rangle$ , we can construct states with decreasing eigenvalues of  $I_3$ :

$$\text{etc. } \xleftarrow{I_-} | \ell, m - 3 \rangle \xleftarrow{I_-} | \ell, m - 2 \rangle \xleftarrow{I_-} | \ell, m - 1 \rangle \xleftarrow{I_-} | \ell, m \rangle \quad (5.19)$$

Repeating the same manipulations as the ones for  $I_+$ , shows that for  $| \psi \rangle = I_- | \ell, m \rangle$ ,

$$\langle \psi | \psi \rangle = \left[ \ell(\ell + 1) - m(m - 1) \right] \langle \ell, m | \ell, m \rangle , \quad (5.20)$$

so it follows that we must have  $\ell(\ell + 1) - m(m - 1) \geq 0$ , and subsequently  $\ell(\ell + 1) - m(m - 1) \geq 0$ , that is,  $-\ell \leq m \leq \ell + 1$ . Since we must require the series (5.19) to terminate as well, there must be a state in the series with minimal eigenvalue  $m = -\ell$ , which guarantees that

$$I_- | \ell, -\ell \rangle = 0 . \quad (5.21)$$

Again, we encounter an undetermined phase factor. It seems that we have the freedom to choose it any way we like, so again we fix the phase factor to be  $+1$ , but we return to this phase factor shortly:

$$I_- | \ell, m \rangle = \sqrt{\ell(\ell + 1) - m(m - 1)} | \ell, m - 1 \rangle . \quad (5.22)$$

Starting from a given state  $| \ell, m \rangle$ , we now have constructed  $\ell - m$  states with eigenvalues  $m + 1, m + 2, \dots, \ell$  and  $\ell + m$  states with  $I_3$  eigenvalues  $m - 1, m - 2, \dots, -\ell$ . Thus, in total we found  $1 + (\ell - m) + (\ell + m) = 2\ell + 1$  states. This is why  $2\ell + 1$  must be an integral number, so that  $\ell$ , and therefore also  $m$ , are either both integers or both an integer plus  $\frac{1}{2}$ .

Above arguments do not quite suffice to prove that we indeed found all states. In principle, it might be possible to apply arbitrary sequences of  $I_+$  and  $I_-$  operators, to find many more states. Suppose we apply  $I_+$  and subsequently  $I_-$ . We get a state with

the same values of both  $\ell$  and  $m$  as before. But is this the same state? Indeed, the answer is yes — and also the phase is  $+1$  ! Note that

$$I_- I_+ = I_1^2 + I_2^2 + i(I_1 I_2 - I_2 I_1) = (\vec{I})^2 - I_3^2 - I_3 = \left( \ell(\ell + 1) - m(m + 1) \right) \mathbf{1} . \quad (5.23)$$

This ensures that, if we apply (5.15) and (5.22) in succession, we get back exactly the same state as the one we started off with (correctly normalized, and with a phase factor  $+1$ ).

By way of exercise, we verify that the operators  $I_+$ ,  $I_-$  and  $I_3$  exclusively act on this single series of states  $|\ell, m\rangle$  as prescribed by Eqs. (5.9), (5.15), and (5.22). Checking the commutation rules,

$$[I_3, I_{\pm}] = \pm I_{\pm} , \quad [I_+, I_-] = 2I_3 , \quad (5.24)$$

we indeed find

$$\begin{aligned} (I_3 I_{\pm} - I_{\pm} I_3) |\ell, m\rangle &= (m \pm 1) \sqrt{\ell(\ell + 1) - m(m \pm 1)} |\ell, m \pm 1\rangle \\ &\quad - m \sqrt{\ell(\ell + 1) - m(m \pm 1)} |\ell, m \pm 1\rangle \\ &= \pm \sqrt{\ell(\ell + 1) - m(m \pm 1)} |\ell, m \pm 1\rangle \\ &= \pm I_{\pm} |\ell, m\rangle , \end{aligned} \quad (5.25)$$

$$\begin{aligned} (I_+ I_- - I_- I_+) |\ell, m\rangle &= \sqrt{\ell(\ell + 1) - (m - 1)m} \sqrt{\ell(\ell + 1) - (m - 1)m} |\ell, m\rangle \\ &\quad - \sqrt{\ell(\ell + 1) - (m + 1)m} \sqrt{\ell(\ell + 1) - (m + 1)m} |\ell, m\rangle \\ &= 2m |\ell, m\rangle \\ &= 2I_3 |\ell, m\rangle . \end{aligned} \quad (5.26)$$

Summarizing, we found that an irreducible representation of  $I_1$ ,  $I_2$ ,  $I_3$  can be characterized by a number  $\ell$ , and it acts on a space spanned by  $2\ell + 1$  states  $|\ell, m\rangle$  for which

$$\begin{aligned} \vec{I}^2 |\ell, m\rangle &= \ell(\ell + 1) |\ell, m\rangle , \\ I_3 |\ell, m\rangle &= m |\ell, m\rangle , \\ I_{\pm} |\ell, m\rangle &= \sqrt{\ell(\ell + 1) - m(m \pm 1)} |\ell, m \pm 1\rangle , \end{aligned} \quad (5.27)$$

with  $m = -\ell, -\ell + 1, -\ell + 2, \dots, \ell - 2, \ell - 1, \ell$ . Either both  $\ell$  and  $m$  are integers, or they are both integers plus  $\frac{1}{2}$ . Of course, we always have  $I_1 = \frac{1}{2}(I_+ + I_-)$  and  $I_2 = \frac{1}{2i}(I_+ - I_-)$ .

We now provide some examples, being the representations for  $\ell = 0, \frac{1}{2}, 1$ , and  $\frac{3}{2}$ :

- For  $\ell = 0$ , we find the trivial representation. There is only one state,  $|0, 0\rangle$ , and  $I_i |0, 0\rangle = 0$  for  $i = 1, 2, 3$ .

- For  $\ell = \frac{1}{2}$ , we find a two-dimensional representation. There are two basis elements,  $|\frac{1}{2}, \frac{1}{2}\rangle$  and  $|\frac{1}{2}, -\frac{1}{2}\rangle$ , for which, according to Eq. (5.27), we have

$$\begin{aligned} I_+ |\frac{1}{2}, -\frac{1}{2}\rangle &= |\frac{1}{2}, \frac{1}{2}\rangle, \\ I_+ |\frac{1}{2}, \frac{1}{2}\rangle &= 0, \\ I_- |\frac{1}{2}, \frac{1}{2}\rangle &= |\frac{1}{2}, -\frac{1}{2}\rangle, \\ I_- |\frac{1}{2}, -\frac{1}{2}\rangle &= 0. \end{aligned} \quad (5.28)$$

This way, we find the matrices

$$I_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad I_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (5.29)$$

The matrices  $I_1$ ,  $I_2$  en  $I_3$  following from this calculation, are the matrices  $\frac{1}{2}\tau_i$  that will be introduced in Chapter 6.

- For  $l = 1$  we find a three-dimensional representation. There are three basis elements,  $|1, 1\rangle$ ,  $|1, 0\rangle$  and  $|1, -1\rangle$ , for which, according to Eq. (5.27), we have

$$\begin{aligned} I_+ |1, -1\rangle &= \sqrt{2} |1, 0\rangle, \\ I_+ |0\rangle &= \sqrt{2} |1, 1\rangle, \\ I_+ |1, 1\rangle &= 0, \\ I_- |1, 1\rangle &= \sqrt{2} |1, 0\rangle, \\ I_- |1, 0\rangle &= \sqrt{2} |1, -1\rangle, \\ I_- |-1, -1\rangle &= 0. \end{aligned} \quad (5.30)$$

This way, we find the matrices

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}, \quad I_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}. \quad (5.31)$$

The matrices  $I_1$ ,  $I_2$  en  $I_3$  are here equal to the matrices  $L_i$ , but in a different (complex) basis, where  $L_3$  is diagonal.

- For  $l = \frac{3}{2}$ , we find a four dimensional representation. We have the basis elements  $|\frac{3}{2}, \frac{3}{2}\rangle$ ,  $|\frac{3}{2}, \frac{1}{2}\rangle$ ,  $|\frac{3}{2}, -\frac{1}{2}\rangle$  en  $|\frac{3}{2}, -\frac{3}{2}\rangle$ , for which, according to Eq. (5.27),

$$\begin{aligned} I_+ |\frac{3}{2}, -\frac{3}{2}\rangle &= \sqrt{3} |\frac{3}{2}, -\frac{1}{2}\rangle, \\ I_+ |\frac{3}{2}, -\frac{1}{2}\rangle &= 2 |\frac{3}{2}, \frac{1}{2}\rangle, \\ I_+ |\frac{3}{2}, \frac{1}{2}\rangle &= \sqrt{3} |\frac{3}{2}, \frac{3}{2}\rangle, \\ I_+ |\frac{3}{2}, \frac{3}{2}\rangle &= 0. \end{aligned} \quad (5.32)$$

This way, we find the marices

$$I_3 = \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.33)$$

The matrix  $I_-$  can be derived in a similar way from Eq. (5.27), or can be obtained directly by hermitian conjugation:  $I_- = I_+^\dagger$ .

## 6. The group $SU(2)$

In Chapter 4, we only saw irreducible representations of the three-dimensional rotation group that were all odd dimensional. Chapter 5, however, showed the complete set of all irreducible representations of this group, and as many of them are even as there are odd ones. More understanding of the even-dimensional representations is needed. To this end, we subject the simplest example of these, the one with  $\ell = \frac{1}{2}$ , to a closer inspection. Clearly, we have vectors forming a two-dimensional space, which will be called *spinors*. Every rotation in a three-dimensional space must be associated to a unitary transformation in this spinor space. If  $R = \exp(i \sum_k \alpha_k L_k)$ , then the associated transformation  $X$  is written as  $X = \exp(i \sum_k \alpha_k I_k)$ , where the generators  $I_k$  follow from Eq. (5.29):

$$I_1 = \frac{I_+ + I_-}{2} = \frac{1}{2} \tau_1, \quad I_2 = \frac{I_+ - I_-}{2i} = \frac{1}{2} \tau_2, \quad I_3 = \frac{1}{2} \tau_3. \quad (6.1)$$

Here, we have introduced the following three fundamental  $2 \times 2$  matrices:<sup>12</sup>

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (6.2)$$

These  $\tau$ -matrices obey the following product rules:

$$\tau_i \tau_j = \delta_{ij} \mathbf{1} + i \epsilon_{ijk} \tau_k, \quad (6.3)$$

as can easily be established. Since  $[\tau_i, \tau_j] = \tau_i \tau_j - \tau_j \tau_i$ , we find that the generators  $I_k$  indeed obey the correct commutation rules:

$$\left[ \frac{\tau_i}{2}, \frac{\tau_j}{2} \right] = i \epsilon_{ijk} \frac{\tau_k}{2}. \quad (6.4)$$

The three  $\tau$  matrices are *hermitian* and *traceless*:

$$\tau_i = \tau_i^\dagger; \quad \text{Tr}(\tau_i) = 0. \quad (6.5)$$

For rotations over tiny angles,  $|\vec{\alpha}| \ll 1$ , the associated matrix  $X(\vec{\alpha})$  takes the following form:

$$X(\vec{\alpha}) = \mathbf{1} + iB + O(B^2); \quad B = \alpha_i \frac{\tau_i}{2}. \quad (6.6)$$

One readily verifies that  $X(\vec{\alpha})$  is unitary and that its determinant equals 1:

$$\begin{aligned} (\mathbf{1} + iB + O(B^2))^\dagger &= (\mathbf{1} + iB + O(B^2))^{-1} = \mathbf{1} - iB + O(B^2); \\ \det(\mathbf{1} + iB + O(B^2)) &= 1 + i \text{Tr} B + O(B^2) = 1, \end{aligned} \quad (6.7)$$

since

$$B^\dagger = B, \quad \text{Tr} B = 0. \quad (6.8)$$

---

<sup>12</sup>Also called Pauli matrices, and often indicated as  $\sigma_i$ .

The finite transformation  $X(\vec{\alpha})$  is found by exponentiation of (6.6), exactly in accordance with the limiting procedure displayed in Chapter 3:

$$X(\vec{\alpha}) = \lim_{n \rightarrow \infty} \left\{ \mathbf{1} + i \frac{\alpha_i \tau_i}{n} \right\}^n = \exp \left( i \alpha_i \frac{\tau_i}{2} \right) . \quad (6.9)$$

The matrices  $\frac{1}{2}\tau_i$  are therefore the generators of the rotations for the  $\ell = \frac{1}{2}$  representation. They do require the coefficients  $\frac{1}{2}$  in order to obey exactly the same commutation rules as the generators  $L_i$  of the rotation group in three dimensions, see Eq. (6.4).

By making use of the product property of the  $\tau$ -matrices, we can calculate the exponential expression for  $X(\vec{\alpha})$ . This is done as follows:

$$\begin{aligned} X(\vec{\alpha}) &= e^{i\alpha_i \tau_i / 2} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{i\alpha_j \tau_j}{2} \right)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} \left( \frac{i\alpha_j \tau_j}{2} \right)^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \left( \frac{i\alpha_j \tau_j}{2} \right)^{2n+1} , \end{aligned} \quad (6.10)$$

where, in the last line, we do the summation over the even and the odd powers of  $(i\alpha_j \tau_j)$  separately. Now we note that

$$(i\alpha_j \tau_j)^2 = -\alpha_j \alpha_k \tau_j \tau_k = -\alpha^2 \mathbf{1}, \quad (6.11)$$

where use was made of Eq. (6.3), and  $\alpha$  is defined as

$$\alpha = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} . \quad (6.12)$$

From Eq. (6.11) it immediately follows that

$$(i\alpha_j \tau_j)^{2n} = (-)^n \alpha^{2n} \mathbf{1}, \quad (i\alpha_j \tau_j)^{2n+1} = (-)^n \alpha^{2n} (i\alpha_j \tau_j) , \quad (6.13)$$

so that we can write Eq. (6.10) as

$$\begin{aligned} X(\vec{\alpha}) &= \left\{ \sum_{n=0}^{\infty} \frac{(-)^n}{(2n)!} \left( \frac{\alpha}{2} \right)^{2n} \right\} \mathbf{1} + \left\{ \sum_{n=0}^{\infty} \frac{(-)^n}{(2n+1)!} \left( \frac{\alpha}{2} \right)^{2n+1} \right\} \left( \frac{i\alpha_j \tau_j}{\alpha} \right) \\ &= \cos \frac{\alpha}{2} \mathbf{1} + i \sin \frac{\alpha}{2} \frac{\alpha_j \tau_j}{\alpha} . \end{aligned} \quad (6.14)$$

It so happens that every  $2 \times 2$  matrix can be decomposed in the unit matrix  $\mathbf{1}$  and  $\tau_i$ :

$$X = c_0 \mathbf{1} + i c_i \tau_i . \quad (6.15)$$

If we furthermore use the product rule (6.3) and Eq. (6.5), and also

$$\text{Tr}(\mathbf{1}) = 2 , \quad (6.16)$$

the coefficients  $c_0$  and  $c_i$  can be determined for every  $2 \times 2$  matrix  $X$  :

$$c_0 = \frac{1}{2} \text{Tr}(X) ; \quad c_i = \frac{1}{2} \text{Tr}(X \tau_i) . \quad (6.17)$$

In our case, we read off the coefficients  $c_0$  and  $c_i$  directly from Eq. (6.14):

$$c_0 = \cos \frac{\alpha}{2}, \quad c_i = \frac{\alpha_i}{\alpha} \sin \frac{\alpha}{2} . \quad (6.18)$$

It is clear that all these coefficients are real. Furthermore, we simply establish:

$$c_0^2 + c_i^2 = 1 . \quad (6.19)$$

The expression (6.15) for  $X(\vec{\alpha})$  can now also be written in terms of two complex parameters  $a$  en  $b$  ,

$$X = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} , \quad (6.20)$$

with  $|a|^2 + |b|^2 = 1$  . Matrices of the form (6.20) with generic  $a$  and  $b$  obeying  $|a|^2 + |b|^2 = 1$  form the elements of the group  $SU(2)$  , the group of unitary  $2 \times 2$  matrices with determinant 1, because <sup>13</sup> they obey:

$$X^\dagger = X^{-1}, \quad \det X = 1 . \quad (6.21)$$

It should be clear that these matrices form a group: if  $X_1$  en  $X_2$  both obey (6.21) and (6.20), then also  $X_3 = X_1 X_2$  and so this matrix also is an element of the group. Furthermore, the unit matrix and the inverse matrix obey (6.20) en (6.21), so they also are in the group, while associativity for the multiplication is evident as well.

In chapter 3 we established that the rotations can be parameterized by vectors  $\vec{\alpha}$  that lie in a sphere with radius  $\alpha = \pi$  . The direction of  $\vec{\alpha}$  coincides with the axis of rotation, and its length  $\alpha$  equals the angle of rotation. Since rotations over  $+\pi$  and  $-\pi$  radians are equal, we established that

$$R(\vec{\alpha}) = R(-\vec{\alpha}) , \quad \text{if} \quad \alpha = \pi . \quad (6.22)$$

As we see in Eq. (6.14), the elements of  $SU(2)$  can be parameterized by the same vectors  $\vec{\alpha}$  . However, to parameterize all elements  $X(\vec{\alpha})$  , the radius of the sphere must be taken to be *twice as large*, that is, equal to  $2\pi$  . Again consider two vectors in opposite directions,  $\vec{\alpha}$  and  $\vec{\alpha}'$  , in this sphere, such that the lengths  $\alpha + \alpha' = 2\pi$  , so that they yield the same rotation,

$$R(\vec{\alpha}') = R(\vec{\alpha}) , \quad (6.23)$$

just because they rotate over the same axis with a difference of  $2\pi$  in the angles. *The two associated  $SU(2)$  elements,  $X(\vec{\alpha}')$  and  $X(\vec{\alpha})$  , however, are opposite to each other:*

$$X(\vec{\alpha}') = -X(\vec{\alpha}) . \quad (6.24)$$

---

<sup>13</sup>Similarly, the complex numbers with norm 1 form the group  $U(1)$  , which simply consists of all phase factors  $\exp i\alpha$  .

This follows from Eqs. (6.14), (6.18) and the fact that  $\cos \frac{\alpha'}{2} = -\cos \frac{\alpha}{2}$  en  $\sin \frac{\alpha'}{2} = \sin \frac{\alpha}{2}$ .

The above implies that, strictly speaking, the elements of  $SU(2)$  are not a representation of the three-dimensional rotation group, but a *projective* representation. After all, in the product of the rotations

$$R(\vec{\alpha}) R(\vec{\beta}) = R(\vec{\gamma}), \quad (6.25)$$

with  $\alpha, \beta$ , we would also have  $\gamma \leq \pi$ , but the product of the associated  $SU(2)$  matrices,

$$X(\vec{\alpha}) X(\vec{\beta}) = \pm X(\vec{\gamma}), \quad (6.26)$$

the value of  $\gamma$  depends on  $\alpha$  and  $\beta$  but its length can be either larger or smaller than  $\pi$ , so we may or may not have to include a minus sign in the equation<sup>14</sup> if we wish to restrict ourselves to vectors shorter than  $\pi$ . The group  $SU(2)$  does have the same structure constants, and thus the same group product structure, as the rotation group, but the latter only holds true in a small domain surrounding the unit element, and not exactly for the entire group.

A spinor  $\varphi^\alpha$  transforms as follows:

$$\varphi^\alpha \rightarrow \varphi^{\alpha'} = X^\alpha_\beta \varphi^\beta. \quad (6.27)$$

The complex conjugated vectors then transform as

$$\varphi_\alpha^* \rightarrow \varphi_\alpha^{*'} = (X^\alpha_\beta)^* \varphi_\beta^* = (X^\dagger)^\beta_\alpha \varphi_\beta^*. \quad (6.28)$$

Here, we introduced an important new notation: the indices are sometimes in a raised position (superscripts), and sometimes lowered (subscripts). This is done to indicate that spinors with superscripts, such as in (6.27), transform differently under a rotation than spinors with subscripts, such as (6.28). Upon complex conjugation, a superscript index becomes a subscript, and *vice versa*. Subsequently, we limit our summation convention to be applied only in those cases where one superscript index is identified with one subscript index:

$$\phi_\alpha \psi^\alpha \equiv \sum_{\alpha=1}^2 \phi_\alpha \psi^\alpha. \quad (6.29)$$

In contrast to the case of the rotation group, one cannot apply group-invariant summations with two superscript or two subscript indices, since

$$X^\alpha_{\alpha'} X^\beta_{\beta'} \delta^{\alpha'\beta'} = \sum_\gamma X^\alpha_\gamma X^\beta_\gamma \neq \delta^{\alpha\beta}, \quad (6.30)$$

because  $X$  in general is not orthogonal, but unitary. The only allowed Kronecker delta function is one with one superscript and one subscript index:  $\delta^\alpha_\beta$ . A summation such as in Eq.(6.29) is covariant:

$$\sum_{\alpha=1}^2 \phi'_\alpha \psi'^\alpha = (X^\alpha_\beta)^* X^\alpha_\gamma \phi_\beta \psi^\gamma = (X^\dagger X)^\beta_\gamma \phi_\beta \psi^\gamma = \delta^\beta_\gamma \phi_\beta \psi^\gamma = \sum_{\beta=1}^2 \phi_\beta \psi^\beta, \quad (6.31)$$

---

<sup>14</sup>On the other hand, we may state that the three-dimensional rotations are a representation of the group  $SU(2)$ .

where unitarity, according to the first of Eqs. (6.21), is used.

We do have two other *invariant tensors* however, to wit:  $\varepsilon_{\alpha\beta}$  and  $\varepsilon^{\alpha\beta}$ , which, as usual, are defined by

$$\varepsilon^{\alpha\beta} = e_{\alpha\beta} = -\varepsilon_{\beta\alpha}, \quad \varepsilon^{12} = \varepsilon_{12} = 1. \quad (6.32)$$

By observing that

$$X_{\alpha'}^{\alpha} X_{\beta'}^{\beta} \varepsilon^{\alpha'\beta'} = \det X \varepsilon^{\alpha\beta} = \varepsilon^{\alpha\beta}, \quad (6.33)$$

where the second of Eqs. (6.21) was used, we note that  $\varepsilon_{\alpha\beta}$  and  $\varepsilon^{\alpha\beta}$  after the transformation take the same form as before.

From this, one derives that the representation generated by the matrices  $X^*$  is *equivalent* to the original representation. With every *co-spinor*  $\varphi^{\alpha}$  we have a *contra-spinor*,

$$\psi_{\alpha} \stackrel{\text{def}}{=} \varepsilon_{\alpha\beta} \varphi^{\beta}, \quad (6.34)$$

transforming as in Eq. (6.28).

The fact that  $X$  and  $X^*$  are equivalent can also be demonstrated by writing  $\varepsilon_{\alpha\beta}$  as a matrix:

$$\begin{aligned} \varepsilon X \varepsilon^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^* & b^* \\ -b & a \end{pmatrix} \\ &= X^*, \end{aligned} \quad (6.35)$$

since  $\varepsilon^2 = -\mathbf{1}$ . From this, it follows that the two representations given by (6.27) and (6.28) are equivalent according to the definition given in Eq. (4.3).

Now, let us attempt to find *all* representations of the group  $SU(2)$ , rather than  $SO(3)$ . To this end, we let the  $SU(2)$  matrices act in an abstract vector space with *complex* coordinates <sup>15</sup>  $\varphi^{\alpha}$ , where  $\alpha = 1, 2$ . We consider all *analytic* functions  $f$  of these two coordinates. Perform the Taylor series expansion of such a function at the origin. At the  $N^{\text{th}}$  order, the Taylor expansion terms form homogeneous, symmetric polynomials in  $\varphi^{\alpha}$  of degree  $N$ . Obviously,  $N$  is a non negative integer. Since  $f$  is analytic, the complex conjugated spinors,  $\varphi_{\alpha}^*$  are not allowed to enter in these polynomials. Write

$$Y^{\alpha_1 \alpha_2 \dots \alpha_N} = \varphi^{\alpha_1} \varphi^{\alpha_2} \dots \varphi^{\alpha_N}. \quad (6.36)$$

Under  $SU(2)$ , these polynomials transform as follows:

$$Y^{\alpha_1 \alpha_2 \dots \alpha_N} \rightarrow Y^{\alpha'_1 \alpha'_2 \dots \alpha'_N} = X_{\alpha'_1}^{\alpha_1} X_{\alpha'_2}^{\alpha_2} \dots X_{\alpha'_N}^{\alpha_N} Y^{\alpha'_1 \alpha'_2 \dots \alpha'_N}. \quad (6.37)$$

In view of the above, we expect that the tensors  $Y^{\alpha_1 \alpha_2 \dots \alpha_N}$  (which, because of the symmetry under interchange of the indices, do not depend on the way these indices are

---

<sup>15</sup>The coordinates  $\varphi^{\alpha}$  are therefore slightly more difficult to interpret.

ordered), should transform as representations of  $SU(2)$ . Indeed, they are *irreducible* representations. The independent coefficients of these polynomials are completely characterized by specifying the number  $p_1$  of indices that are equal to 1 (the remaining indices, their number being  $p_2 = N - p_1$ , must be equal to 2), and so we find the number of independent coefficients in a polynomial of degree  $N$  to be

$$\sum_{p_1=0}^N = N + 1 . \quad (6.38)$$

Thus, here we have representations of dimension  $N + 1$ , for any non negative integer  $N$ .

Subsequently, we can write the  $SU(2)$  generators, acting on functions of the coordinates  $\varphi$ , as differential operators. This leads to

$$L_i^{SU(2)} = -\frac{1}{2}(\tau_i)^\alpha{}_\beta \varphi^\beta \frac{\partial}{\partial \varphi^\alpha} , \quad (6.39)$$

so that infinitesimal  $SU(2)$  transformations on functions  $f(\varphi)$  can be written as

$$\begin{aligned} f(\varphi) \rightarrow f'(\varphi) &= \left( \mathbf{1} - i\vec{\alpha} \cdot \vec{L}^{SU(2)} + O(\alpha^2) \right) f(\varphi) \\ &= f(\varphi) + \frac{i}{2} \alpha_j (\tau_j)^\alpha{}_\beta \varphi^\beta \frac{\partial f(\varphi)}{\partial \varphi^\alpha} + O(\alpha^2) . \end{aligned} \quad (6.40)$$

Note in passing that the index  $\alpha$  in  $\frac{\partial}{\partial \varphi^\alpha}$  is treated as a *subscript* index.

Making use of Eq. (6.39), we can now derive the Casimir operator  $(\vec{L}^{SU(2)})^2$  as a differential operator,

$$\begin{aligned} (L_i^{SU(2)})^2 &= \frac{1}{4} \sum_i \left( \tau_{i\beta}^\alpha \varphi^\beta \frac{\partial}{\partial \varphi^\alpha} \right) \left( \tau_{i\delta}^\gamma \varphi^\delta \frac{\partial}{\partial \varphi^\gamma} \right) \\ &= \frac{1}{4} (-\delta_\beta^\alpha \delta_\delta^\gamma + 2\delta_\delta^\alpha \delta_\beta^\gamma) \varphi^\beta \frac{\partial}{\partial \varphi^\alpha} \varphi^\delta \frac{\partial}{\partial \varphi^\gamma} \\ &= -\frac{1}{4} \varphi^\alpha \frac{\partial}{\partial \varphi^\alpha} \varphi^\gamma \frac{\partial}{\partial \varphi^\gamma} + \frac{1}{2} \varphi^\beta \frac{\partial}{\partial \varphi^\alpha} \varphi^\alpha \frac{\partial}{\partial \varphi^\beta} \\ &= \frac{1}{4} \left( \varphi^\alpha \frac{\partial}{\partial \varphi^\alpha} \right)^2 + \frac{1}{2} \left( \varphi^\alpha \frac{\partial}{\partial \varphi^\alpha} \right) . \end{aligned} \quad (6.41)$$

It is easy to see that the last two lines of Eq. (6.41) are equal by writing all derivatives to the right of the coordinates. The transition from the first to the second line is less trivial. There, use was made of the identity

$$\sum_i (\tau_i)^\alpha{}_\beta (\tau_i)^\gamma{}_\delta = -\delta_\beta^\alpha \delta_\delta^\gamma + 2\delta_\delta^\alpha \delta_\beta^\gamma . \quad (6.42)$$

A convenient way to derive this equation is by first multiplying it with an arbitrary matrix  $X_\delta^\gamma$ , after which one uses the decomposition rule (6.15) and Eq. (6.17) for this  $X$ . If now

the derivative of this equation is taken with respect to  $X_\delta^\gamma$ , we directly end up with the identity (6.42). Evidently, the validity of (6.42) can also be verified by choosing specific values for the indices  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$ .

Now, let the operator (6.41) act on the polynomials  $Y^{\alpha_1\alpha_2\cdots\alpha_{2s}}$ . Using the fact that

$$\left(\varphi^\alpha \frac{\partial}{\partial \varphi^\alpha}\right) Y^{\alpha_1\alpha_2\cdots\alpha_N} = N Y^{\alpha_1\alpha_2\cdots\alpha_{2s}}, \quad (6.43)$$

we find directly the result:

$$(L_i^{SU(2)})^2 Y^{\alpha_1\alpha_2\cdots\alpha_N} = \frac{1}{2}N(\frac{1}{2}N + 1) Y^{\alpha_1\alpha_2\cdots\alpha_N}. \quad (6.44)$$

Thus, we recognize the representations  $\ell$  of chapter 5, if we write  $\ell = s$ ,  $s = \frac{1}{2}N$ . We succeeded to (re)construct  $(2s + 1)$  dimensional representations of  $SU(2)$ , where  $s$  is an integer or an integer plus  $\frac{1}{2}$ . In these representations, the eigenvalue of the Casimir operator, according to Eq. (6.44), equals  $s(s + 1)$ . In Chapter 5, it was shown that this completes the set of all irreducible representations of  $SU(2)$ .

We expect that, for integral values of  $s$ , the representations coincide with the representation of the rotation group found in chapter 4. This indeed turns out to be the case. To see this, consider the tensors  $Y$  with an *even* number of indices. Then, arrange the factors  $\varphi^\alpha$  in pairs, and use in each pair the  $\varepsilon$  tensor to lower one of the superscript indices to obtain a subscript index:

$$Y^{\alpha\beta} = \varphi^\alpha \varphi^\beta; \quad \hat{Y}_\beta^\alpha = \varepsilon_{\beta\gamma} Y^{\alpha\gamma} = \varepsilon_{\beta\gamma} \varphi^\alpha \varphi^\gamma. \quad (6.45)$$

We read off easily that  $\text{Sp}(\hat{Y}) = 0$ , so that, according to the decomposition (6.15),  $\hat{Y}$  can be written as

$$\hat{Y} = \frac{1}{2} \sum_i x_i \tau_i; \quad x_i = \hat{Y}_\beta^\alpha (\tau_i)^\beta_\alpha. \quad (6.46)$$

Under  $SU(2)$ , the quantities  $x_i$  transform as

$$x_i \rightarrow x'_i = X^\alpha_{\alpha'} (X^{-1})^{\beta'}_\beta Y^{\alpha'}_{\beta'} (\tau_i)^\beta_\alpha, \quad (6.47)$$

where use was made of the transformation rules for superscript and subscript indices. And now we prove that

$$X^{-1}(\vec{\alpha}) \tau_i X(\vec{\alpha}) = R(\vec{\alpha})_{ij} \tau_j, \quad (6.48)$$

so that the tensors  $x_i$  actually transform exactly like the coordinates  $x_i$  in chapter 4. We verify the validity of the transformation rule (6.48) for infinitesimal transformations. One then has

$$\begin{aligned} X^{-1}(\vec{\alpha}) \tau_i X(\vec{\alpha}) &\approx \left( \mathbf{1} - \frac{i}{2} \alpha_j \tau_j + O(\alpha^2) \right) \tau_i \left( \mathbf{1} + \frac{i}{2} \alpha_k \tau_k + O(\alpha^2) \right) \\ &= \left( \tau_i + \frac{i}{2} \alpha_j [\tau_i, \tau_j] + O(\alpha^2) \right) \\ &= \tau_i + \epsilon_{ijk} \tau_j \alpha_k + O(\alpha^2), \end{aligned} \quad (6.49)$$

which indeed takes the same form as infinitesimal rotations of the coordinates  $x_i$ .

The rotation operator (6.39) is an exact analogue of the generator of rotations in  $x$  space:

$$L_i = -i\varepsilon_{ijk}x_j \frac{\partial}{\partial x_k} , \quad (6.50)$$

which we obtain if we apply an infinitesimal rotation (3.11) to a function  $\psi(\vec{r})$ :

$$\psi(\vec{r}) \rightarrow \psi(\vec{r} + \vec{r} \times \vec{\alpha}) = (\mathbf{1} + i\alpha_k L_k)\psi(\vec{r}) . \quad (6.51)$$

## 7. Spin and angular distributions

In the preceding chapters we saw how wave functions of some rotationally invariant system can be classified according to representations of the rotation group. These representations can be decomposed into irreducible ones, which in turn are characterized by an integer or half-odd integer  $\ell$  or  $s$ . We also know that this is a feature of the hydrogen atom, where the wave functions depend on an integer  $\ell$ , defining the orbital angular momentum of the electron. The generators,  $L_k$ , as defined in the preceding chapters (in particular in (3.13) and (6.50)), actually correspond to the angular momentum operators. This follows from the fact that the quantum mechanical operator associated with momentum, acting on wave functions  $\psi(\vec{r})$  has the form

$$\vec{p} = \frac{\hbar}{i} \frac{\partial}{\partial \vec{r}}. \quad (7.1)$$

The operator for angular momentum  $\vec{L} \equiv \vec{r} \times \vec{p}$ , takes the same form as the generators (6.50), apart from a factor  $\hbar$ . According to Eq. (5.9) it therefore follows that the total angular momentum of the states is given by the eigenvalue equation

$$\vec{L}^2 | \psi \rangle = \ell(\ell + 1) \hbar^2 | \psi \rangle. \quad (7.2)$$

However, in addition to orbital angular momentum, particles can have *intrinsic* angular momentum, a contribution to the angular momentum from the particle's internal structure besides the motion of its center of gravity. In principle, this can be understood by assuming that the particle has finite dimensions, so that its rotation about its center of gravity could be responsible for this extra contribution to the angular momentum, also called *spin*. This feature is known from classical mechanics (an example is a rotating top). This way of looking at intrinsic angular momentum is not free from troubles, but this does not concern us for the moment. Our starting point is that intrinsic angular momentum fits naturally within the formalism introduced in the previous chapters. We leave it to experiments to decide whether and to what extent Nature made use of this possibility. For particles with spin, it is no longer obvious that only representations occur with integral values of  $\ell$ . Indeed, it turns out that particles with half-odd integer spin occur in Nature, such as the electron, the proton, and the neutron.

From representation theory, it follows that a particle with spin  $s$  can be in  $2s + 1$  mutually independent states.<sup>16</sup> This is why the wave function has  $2s + 1$  different components. Upon a rotation, these different states are transformed into one another, as dictated by the associated representation of the rotation group. The fact that particles occur in different spin states, of course has implications for experiment. Many of the experiments that demonstrate the existence of spin, and the related properties, make use of the *magnetic moment* induced by spin. The well-known Stern-Gerlach experiments can serve as an example. An other measurement from which the existence of different spin

---

<sup>16</sup>An exception must be made for *massless* particles, such as the photon, which always move with the velocity of light. For massless particles, at most two different states are possible; spin is defined differently in that case.

states can be deduced, is that of the angular distribution in scattering experiments. In this chapter, we shall demonstrate this by describing an idealized experiment to observe scattering of protons against a carbon nucleus.

Protons have spin  $s = \frac{1}{2}$ , so that we are actually dealing with two species of protons, to be described by a wave function with two components. In practice, we write such a wave function as a spinor depending on the position in space, and on time. In total, one thus needs two independent wave functions describing the two possible proton states:  $\psi = \psi_{\uparrow} + \psi_{\downarrow}$ ,

$$\psi_{\uparrow}(\vec{r}) = f_{\uparrow}(\vec{r}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \psi_{\downarrow}(\vec{r}) = f_{\downarrow}(\vec{r}) \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (7.3)$$

The above states are eigenstates of the generator  $L_3^{SU(2)}$ , which is associated to rotations about the  $z$  axis. The first state is said to be the state with spin oriented along the positive  $z$  axis having the value  $S_z = \frac{1}{2}\hbar$ , the second state is the one with its spin in the opposite direction,  $S_z = -\frac{1}{2}\hbar$ . Upon a rotation, these two states may transform into one another. For example, a rotation about the  $x$  axis of  $\pi$  radians acts on the spinors as a matrix (see Eq. (6.14))

$$X = i\tau_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad (7.4)$$

so that our two states are being interchanged.

In this chapter, we consider the situation of a proton that is being scattered against a carbon nucleus ( $C^{12}$ ). Suppose that the proton comes in along the positive  $x$  axis, hitting a carbon nucleus at the origin of our coordinate frame. Carbon nuclei are spherically symmetric (in a good approximation), and we can compare the scattering of protons with spin ‘‘up’’ (that is,  $S_z = +\frac{1}{2}\hbar$ ) with the scattering of protons with spin ‘‘down’’ ( $S_z = -\frac{1}{2}\hbar$ ). It now turns out that one species of protons prefers to deviate towards the negative  $y$  axis, and the other towards the positive  $y$  axis. This is established in the *angular distribution*, which describes the probability for the proton to scatter into an angle  $\theta$ , which, in this case, is defined to be the angle between the  $z$  axis and the plane of scattering (see Figure 3). It turns out that the angular distribution strongly depends on the spin of the arriving particle (it is assumed that the particle against which we scatter, in this case the carbon nucleus, has no spin, and is therefore spherically symmetric); the larger the spin of the particle that enters, the more complicated the angular distribution may be. Here it will be shown that the angular distribution is nearly completely determined by the demand of rotational invariance, apart from a few adjustable constants.

Let us now consider the quantum states before and after the collision.<sup>17</sup> Before the

---

<sup>17</sup>The states of the particles entering are now characterized by writing the wave function at  $t \rightarrow -\infty$ , that is, the asymptotic states long before the scattering takes place. Similarly, the states of the scattered particles are specified by the wave function at  $t \rightarrow +\infty$ , that is, long after the collision took place. States specified this way are called ‘in’ states and ‘out’ states. To compute the inner products, we must use the wave functions at the same moment in time. These considerations are of great importance conceptually, but play no role here.

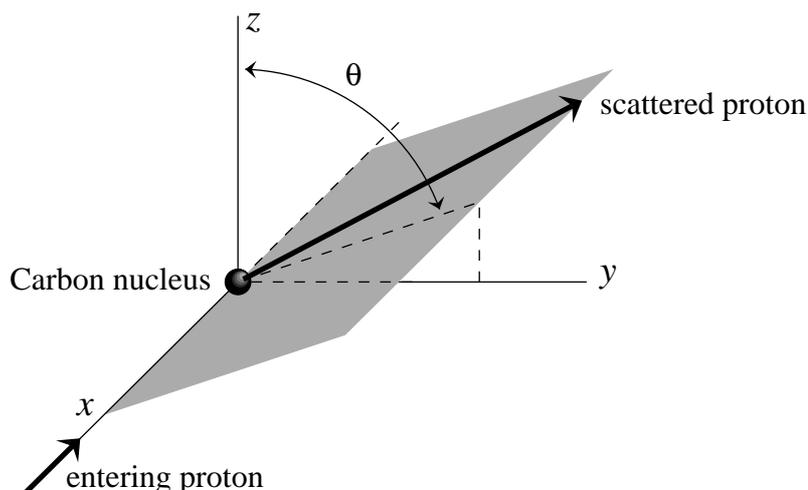


Figure 3: Proton scattering against a carbon nucleus. The protons enter along the  $x$  axis and the plane of scattering, being the plane defined by the trajectories of the proton going in and the proton going out.

collision, we are dealing with an initial proton, whose momentum is determined by the experimental conditions, while its spin state may be chosen at will; the carbon nucleus is in its ground state. Thus, the ‘in’ states can be described as  $|\uparrow\rangle$  and  $|\downarrow\rangle$ , corresponding to the proton wave functions (7.3). After the collision, the possible states are much more difficult to describe. Momentum and spin of the proton have changed, and also the carbon nucleus may in general no longer be at rest, and possibly in an excited state. However, apart from the scattering angle  $\theta$ , defined by the plane formed by the ‘in’ vector and the ‘out’ vector, the experiment is unable to measure all these details concerning the carbon-proton system after the collision, or in any case, we are not interested in these at the moment. The possible out-states are denoted by  $|\psi_{\{\rho\}}(\theta)\rangle$ , where  $\{\rho\}$  summarizes all those variables in the final state that we decided *not* to measure.

When the proton that enters along the  $x$ -axis has spin “up”, the probability that the proton going out is in a state characterized by  $\{\rho\}$  and the angle  $\theta$ , is taken to be  $|\langle\psi_{\{\rho\}}(\theta)|\uparrow\rangle|^2$ . We are not interested in  $\rho$ ; the chance to find a scattering angle  $\theta$ , regardless what  $\rho$  is, will be given by <sup>18</sup>

$$f_1(\theta) = \sum_{\{\rho\}} \left| \langle\psi_{\{\rho\}}(\theta)|\uparrow\rangle \right|^2. \quad (7.5)$$

In a similar way, we find the probability for a proton to scatter under an angle  $\theta$  when it entered with spin “down”, to be equal to

$$f_2(\theta) = \sum_{\{\rho\}} \left| \langle\psi_{\{\rho\}}(\theta)|\downarrow\rangle \right|^2. \quad (7.6)$$

---

<sup>18</sup>With this, we mean that the chance to measure a scattering angle between  $\theta$  and  $\theta + d\theta$  is equal to  $f_1(\theta) d\theta$ . For us it is only of secondary importance how to perform the summation over the variables  $\rho$ .

Upon a rotation over  $\pi$  radians about the  $x$ -axis, the angle  $\theta$  turns into  $\theta + \pi$ , while the two spin states (7.3) of the initial proton are interchanged. This is why invariance under rotations implies that one must have

$$f_1(\theta) = f_2(\theta + \pi). \quad (7.7)$$

We can say more, however. Consider an initial proton with spin characterized by

$$|\psi(\mathbf{a})\rangle = a_1 |\uparrow\rangle + a_2 |\downarrow\rangle. \quad (7.8)$$

A procedure similar to Eqs. (7.5) and (7.6) allows us to derive that the distribution over the angles  $\theta$  is given by

$$\begin{aligned} f(\theta) &= \sum_{\{\rho\}} \left| \langle \psi_{\{\rho\}}(\theta) | \psi(\mathbf{a}) \rangle \right|^2 \\ &= \sum_{\{\rho\}} \left| a_1 \langle \psi_{\{\rho\}}(\theta) | \uparrow \rangle + a_2 \langle \psi_{\{\rho\}}(\theta) | \downarrow \rangle \right|^2 \\ &= f_1(\theta) |a_1|^2 + f_2(\theta) |a_2|^2 + f_{12}(\theta) a_1 a_2^* + f_{12}^*(\theta) a_1^* a_2, \end{aligned} \quad (7.9)$$

where  $f_1$  en  $f_2$  are defined in (7.5) and (7.6), while the mixing term  $f_{12}$  is given by

$$f_{12}(\theta) = \sum_{\{\rho\}} \langle \psi_{\{\rho\}}(\theta) | \uparrow \rangle \langle \psi_{\{\rho\}}(\theta) | \downarrow \rangle^*. \quad (7.10)$$

Eq. (7.9) can also be written in matrix notation,

$$f(\theta) = \sum_{\alpha, \beta=1,2} F(\theta)_{\alpha\beta} a_\beta^* a_\alpha, \quad (7.11)$$

where the  $2 \times 2$  matrix  $F(\theta)$ , defined by

$$F(\theta) = \begin{pmatrix} f_1(\theta) & f_{12}^*(\theta) \\ f_{12}(\theta) & f_2(\theta) \end{pmatrix}, \quad (7.12)$$

is hermitian and positive definite.

Now consider a rotation over an angle  $\phi$  around the  $x$ -axis. To get the signs correct<sup>19</sup>, remember the definition (2.17) of a representation. Comparing Eq. (3.12), we see that  $R(\phi, 0, 0)$  is the rotation matrix

$$R(\phi) = e^{i\phi L_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix}, \quad (7.13)$$

---

<sup>19</sup>Versions of these notes prior to 25/06/07 were not very precise at this point; we modified the notation here.

which rotates the coordinates in the positive  $\theta$  direction in Fig. 3. Having  $\psi_i(Rx) = X_{ij}(\phi)\psi_j(x)$ , where  $\psi(x)$  stands for  $\langle\psi_{\{\rho\}}(\theta)|\psi\rangle$ , we find

$$\langle\psi_{\{\rho\}}(\theta + \phi)|\psi_i\rangle = \psi_i(Rx) = X_{ij}\psi_j(x) = X_{ij}\langle\psi_{\{\rho\}}(\theta)|\psi_j\rangle, \quad (7.14)$$

and  $X_{ij}$  is the matrix that transforms the states  $|\uparrow\rangle$  and  $|\downarrow\rangle$ :

$$X(\phi) = e^{i\phi\tau_1/2} = \begin{pmatrix} \cos \frac{\phi}{2} & i \sin \frac{\phi}{2} \\ i \sin \frac{\phi}{2} & \cos \frac{\phi}{2} \end{pmatrix}. \quad (7.15)$$

Writing Eq. (7.8) as  $|\psi\rangle = a_i|\psi_i\rangle$ , we find that Eq. (7.14) can be written as

$$\langle\psi_{\{\rho\}}(\theta + \phi)|\psi\rangle = a_i X_{ij} \langle\psi_{\{\rho\}}(\theta)|\psi_j\rangle. \quad (7.16)$$

Consequently, with our definitions (7.10)–(7.12) for the matrix  $F(\theta)$ , we have

$$F(\theta + \phi) = X^\dagger(\phi) F(\theta) X(\phi). \quad (7.17)$$

This equation determines the  $\theta$  dependence of the distribution. Just take  $\theta = 0$  in (7.17), so that this equation can be written as (we use  $X(\phi)^\dagger = X(\phi)^{-1}$ )

$$F(\phi) = X(\phi)^{-1} F(0) X(\phi). \quad (7.18)$$

Elaborating this matrix equation for the individual components yields (we replace  $\phi$  now back to  $\theta$ )

$$f_1(\theta) = f_1 \cos^2 \frac{\theta}{2} + f_2 \sin^2 \frac{\theta}{2} + i(f_{12} - f_{12}^*) \cos \frac{\theta}{2} \sin \frac{\theta}{2}, \quad (7.19)$$

$$f_2(\theta) = f_2 \cos^2 \frac{\theta}{2} + f_1 \sin^2 \frac{\theta}{2} - i(f_{12} - f_{12}^*) \cos \frac{\theta}{2} \sin \frac{\theta}{2}, \quad (7.20)$$

$$f_{12}(\theta) = f_{12} \cos^2 \frac{\theta}{2} + f_{12}^* \sin^2 \frac{\theta}{2} + i(f_1 - f_2) \cos \frac{\theta}{2} \sin \frac{\theta}{2}, \quad (7.21)$$

The  $\theta$  dependence of the angular distributions thus follows completely from rotation invariance. The functions  $f_1(\theta)$  en  $f_2(\theta)$  have exactly one maximum and one minimum in the interval  $-\pi \leq \theta \leq \pi$ , and thus they yield exactly one preferred direction for scattering, in agreement with what was described earlier.

The angular distributions can be predicted even more precisely by making use of yet an other symmetry. To a very good precision, Nature is also invariant under *mirror reflections* in space. Under a reflection  $P$  with respect to the origin, all space coordinates are given an extra minus sign,

$$\vec{r} \xrightarrow{P} -\vec{r}. \quad (7.22)$$

Two reflections in succession yield the identity, so that what we have here is a finite group consisting of two elements, the space reflection  $P$  and the identity  $\mathbb{I}$ . By adding the reflection to the rotation group  $SO(3)$ , the latter is expanded to  $O(3)$ , see the beginning of Chapter 3.

The generators of the rotation group do not change upon a reflection. This directly follows from the fact that a reflection with respect to the origin commutes with rotations, as can easily be checked. From this it follows that such a reflection acting on irreducible representations of the rotation group, must be proportional to the identity. This is why the spin of a state does not change under reflections, and the state itself can at most receive a minus sign.

Consider once more the scattering of protons against carbon nuclei, and the effect of the  $O(3)$  matrix

$$P_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \quad (7.23)$$

After this matrix, which is also a reflection since its determinant is  $-1$ , we encounter the same experimental situation, because the momentum of the initial proton did not change. Only the angle  $\theta$  is now replaced by  $\pi - \theta$ . Since, upon this substitution,  $f_1(\theta)$  en  $f_2(\theta)$  may not change, one must have

$$f_1(\theta) = f_1(\pi - \theta), \quad f_2(\theta) = f_2(\pi - \theta) . \quad (7.24)$$

This implies that  $f_1 = f_2$ .

Observe that  $P_z$  can be written as the product of the reflection  $P = -\mathbf{1}$  and a rotation over  $\pi$  radians about the  $z$ -axis. The reflection  $P$  does not affect the spin states, but the rotation corresponds to

$$e^{\frac{1}{2}i\pi\tau_3} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\tau_3 . \quad (7.25)$$

Consequently, we have (analogously to Eq. (7.17)) that  $\tau_3 F(\pi - \theta) \tau_3 = F(\theta)$ , which now also implies that  $f_{12}^* = -f_{12}$ .

## 8. Isospin

As explained in Chapter 1, elementary particles with the same spin and similar mass values, can be grouped in so-called *multiplets*. For example,

$$\begin{aligned}
 \text{nucleon-doublet :} & \quad (p, n), \\
 \text{pion-triplet :} & \quad (\pi^+, \pi^0, \pi^-), \\
 \text{Delta-quadruplet :} & \quad (\Delta^{++}, \Delta^+, \Delta^0, \Delta^-), \\
 & \quad \text{etc.}
 \end{aligned} \tag{8.1}$$

This structure can be explained by assuming that Nature is approximately invariant under what will be referred to as *isospin rotations*. Isospin rotations form a group, whose elements can be regarded as rotations in some “internal” three-dimensional space. The states that correspond to the isospin multiplets transform according to (irreducible) representations of this group. Thus, the notion of isospin is analogous to the usual notion of spin, which is connected to rotations in ordinary 3-space.

The force that binds the nucleons in an atomic nucleus, the so-called strong force, is invariant under isospin rotations. However, the electro magnetic forces are not. This is why that electric charges within one multiplet are not all the same. Conservation of electric charge *is* exact. This is associated with conservation of isospin-“angular momentum” in the 3 direction, since electric charge is described by the Gell-Mann-Nishijima relation

$$Q = I_3 + \frac{1}{2}Y, \tag{8.2}$$

where  $Y$  is the “hyper charge”, which has the same value for all states within one multiplet. We readily determine that  $Y = 1$  for the nucleon and  $\Delta$  multiplets, and 0 for the pion multiplet.

Now let us consider the decay of a  $\Delta^{++}$ ,

$$\Delta^{++} \longrightarrow p \pi^+. \tag{8.3}$$

For the particles involved in this decay, we have

$$I_3 |\Delta^{++}\rangle = \frac{3}{2} |\Delta^{++}\rangle, \quad I_3 |p\rangle = \frac{1}{2} |p\rangle, \quad I_3 |\pi^+\rangle = |\pi^+\rangle, \tag{8.4}$$

so that

$$I_3 \left( |p\rangle |\pi^+\rangle \right) = \left( I_3 |p\rangle \right) |\pi^+\rangle + |p\rangle \left( I_3 |\pi^+\rangle \right) = \frac{3}{2} |p\rangle |\pi^+\rangle. \tag{8.5}$$

The above equation follows from the fact that the generators  $I_i$  for a representation consisting of the products of states, is given by the sum of the generators acting on the separate states. Similar equations therefore hold for  $I_+$  and  $I_-$  as well.

$\Delta$  particles predominantly decay into one nucleon and one pion. Because of charge conservation,  $\Delta^{++}$  can only decay as in Eq. (8.3). As for the other  $\Delta$  particles, the

situation is not so simple. Let us begin by enumerating the six different nucleon-pion states, all being eigenstates of the  $I_3$  -operator,

$$\begin{aligned}
I_3 \left( |p\rangle |\pi^+\rangle \right) &= \frac{3}{2} |p\rangle |\pi^+\rangle , \\
I_3 \left( |p\rangle |\pi^0\rangle \right) &= \frac{1}{2} |p\rangle |\pi^0\rangle , \\
I_3 \left( |p\rangle |\pi^-\rangle \right) &= -\frac{1}{2} |p\rangle |\pi^-\rangle , \\
I_3 \left( |n\rangle |\pi^+\rangle \right) &= \frac{1}{2} |n\rangle |\pi^+\rangle , \\
I_3 \left( |n\rangle |\pi^0\rangle \right) &= -\frac{1}{2} |n\rangle |\pi^0\rangle , \\
I_3 \left( |n\rangle |\pi^-\rangle \right) &= -\frac{3}{2} |n\rangle |\pi^-\rangle .
\end{aligned} \tag{8.6}$$

The fact that  $\Delta$  particles can only decay into special linear combinations of these states can be explained by the fact that the product states (8.6) split up into two different irreducible representations of the isospin group. This follows immediately from an inspection of the eigenvalues of  $I_3$ : the eigenvalues  $\pm\frac{3}{2}$  occur only once each, while the eigenvalues  $\pm\frac{1}{2}$  occur twice each. The corresponding states must therefore form one irreducible representation with “total isospin”  $\frac{3}{2}$ , and one with “total isospin”  $\frac{1}{2}$ . Thus, we can write the following decomposition of the product representation (analogously to Eq. (4.13)),

$$\mathbf{2} \otimes \mathbf{3} = \mathbf{2} \oplus \mathbf{4} \tag{8.7}$$

where  $\mathbf{2}$ ,  $\mathbf{3}$ , and  $\mathbf{4}$  indicate the 2-, 3- and 4-dimensional representations, associated to isospin  $\frac{1}{2}$ , 1, and  $\frac{3}{2}$ , respectively.

If indeed these decay processes are invariant under isospin rotations, then the decays of the other  $\Delta$  particles can be predicted from what is known about the  $\Delta^{++}$  decay. The operator  $I_-$  turns the  $\Delta^{++}$  state into

$$I_- |\Delta^{++}\rangle = I_- \left| \frac{3}{2}, \frac{3}{2} \right\rangle = \sqrt{3} \left| \frac{3}{2}, \frac{1}{2} \right\rangle = \sqrt{3} |\Delta^+\rangle , \tag{8.8}$$

Where we used Eq. (5.22). For the nucleon pion final state, into which  $\Delta^{++}$  decays, we find that it transforms as

$$\begin{aligned}
I_- \left( |p\rangle |\pi^+\rangle \right) &= \left( I_- |p\rangle \right) |\pi^+\rangle + |p\rangle \left( I_- |\pi^+\rangle \right) \\
&= |n\rangle |\pi^+\rangle + \sqrt{2} |p\rangle |\pi^0\rangle .
\end{aligned} \tag{8.9}$$

Isospin invariance should now imply that  $|\Delta^+\rangle$ , which is obtained from an isospin rotation acting on  $|\Delta^{++}\rangle$ , should decay into the state obtained by applying the same rotation to the state  $|p\rangle |\pi^+\rangle$ ; this leads to the state that was obtained in Eq. (8.9), after normalization:

$$|\psi\rangle = \frac{1}{\sqrt{3}} \left( |n\rangle |\pi^+\rangle + \sqrt{2} |p\rangle |\pi^0\rangle \right) . \tag{8.10}$$

Thus, isospin invariance implies that  $\Delta^+$  must decay in the state (8.10), and we find that the decay probabilities for the two possible decay modes of  $\Delta^+$  obey

$$\Gamma(\Delta^+ \rightarrow n \pi^+) : \Gamma(\Delta^+ \rightarrow p \pi^0) = 1 : 2 , \tag{8.11}$$

as is observed in the experiments.

To study the decay modes of  $\Delta^0$ , we apply  $I_-$  once more,

$$\begin{aligned} I_- |\Delta^+\rangle &= 2 |\Delta^0\rangle, \\ I_- \left( |n\rangle |\pi^+\rangle \right) &= \sqrt{2} |n\rangle |\pi^0\rangle, \\ I_- \left( |p\rangle |\pi^0\rangle \right) &= |n\rangle |\pi^0\rangle + \sqrt{2} |p\rangle |\pi^-\rangle, \end{aligned} \quad (8.12)$$

so that

$$I_- \left( \frac{|n\rangle |\pi^+\rangle + \sqrt{2} |p\rangle |\pi^0\rangle}{\sqrt{3}} \right) = 2 \left( \frac{\sqrt{2} |n\rangle |\pi^0\rangle + |p\rangle |\pi^-\rangle}{\sqrt{3}} \right). \quad (8.13)$$

One concludes that the decay probabilities for the two possible decay modes of  $\Delta^0$  are related according to

$$\Gamma(\Delta^0 \rightarrow n \pi^0) : \Gamma(\Delta^0 \rightarrow p \pi^-) = 2 : 1. \quad (8.14)$$

Thus, we see in this example how the ratio 2 : 1 of decay rates of elementary particles can be a consequence of invariance under the transformations forming a Lie group.

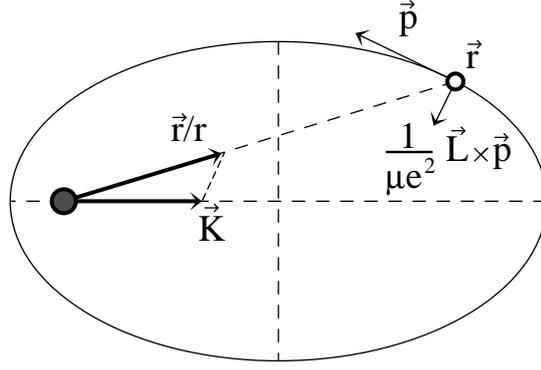


Figure 4: Construction of the so-called Runge-Lenz vector  $\vec{K}$ .

## 9. The Hydrogen Atom

In this chapter we show how group theory can be used to derive the entire energy spectrum of the hydrogen atom. The Hamiltonian of the hydrogen atom has the form

$$H = \frac{\vec{p}^2}{2\mu} - \frac{e^2}{r}, \quad (9.1)$$

where the first term represents the kinetic energy (with  $\mu$  the reduced mass of the electron proton system), and the second term is the contribution of the Coulomb potential. First consider the *classical* equations of motion following from Eq. (9.1):

$$\dot{\vec{p}} = -e^2 \frac{\vec{r}}{r^3}, \quad (9.2)$$

where the momentum  $\vec{p}$  is given by

$$\vec{p} = \mu \dot{\vec{r}}. \quad (9.3)$$

The hydrogen atom is known to feature a number of conserved quantities, that is, quantities that stay constant in time. The best known examples are the *energy*, given by Eq. (9.1), and the 3 components of the angular momentum vector,

$$\vec{L} = \vec{r} \times \vec{p}, \quad \text{or:} \quad L_i = \varepsilon_{ijk} r_j p_k. \quad (9.4)$$

Indeed, one easily derives from Eqs. (9.2) and (9.3) that

$$\frac{d}{dt} H = \frac{\vec{p} \cdot \dot{\vec{p}}}{\mu} + \frac{e^2}{r^3} \vec{r} \cdot \dot{\vec{r}} = 0, \quad (9.5)$$

$$\frac{d}{dt} \vec{L} = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} = 0. \quad (9.6)$$

The latter result follows from the fact that, according to Eq. (9.3),  $\dot{\vec{r}}$  and  $\vec{p}$  are parallel, and according to Eq. (9.2), also  $\vec{r}$  and  $\dot{\vec{p}}$  are parallel. There exists however yet another

vector that stays constant. This is the so-called Runge-Lenz vector, defined by

$$\vec{K} = \frac{1}{\mu e^2} \vec{L} \times \vec{p} + \frac{\vec{r}}{r} . \quad (9.7)$$

The fact that  $\vec{K}$  is conserved follows from Eqs. (9.2) en (9.3), but the proof is a little more involved. First, we show that

$$\begin{aligned} \frac{d}{dt} \vec{K} &= \frac{1}{\mu e^2} \vec{L} \times \dot{\vec{p}} + \frac{\dot{\vec{r}}}{r} - (\dot{\vec{r}} \cdot \vec{r}) \frac{\vec{r}}{r^3} \\ &= -\frac{1}{\mu} \frac{\vec{L} \times \vec{r}}{r^3} + \frac{r^2 \vec{p} - (\vec{p} \cdot \vec{r}) \vec{r}}{\mu r^3} , \end{aligned} \quad (9.8)$$

where in the first line Eq. (9.6) was used, and in the second line Eqs.(9.2) and (9.3). Next, we show that Eq. (9.8) is equal to zero by substituting Eq. (9.4), so that

$$\frac{d}{dt} \vec{K} = \frac{1}{\mu r^3} \{ -(\vec{r} \times \vec{p}) \times \vec{r} + r^2 \vec{p} - (\vec{r} \cdot \vec{p}) \vec{r} \} . \quad (9.9)$$

Using the relation

$$(\vec{a} \times \vec{b}) \times \vec{c} = (\vec{c} \cdot \vec{a}) \vec{b} - (\vec{c} \cdot \vec{b}) \vec{a} , \quad (9.10)$$

we subsequently show that Eq. (9.9) is equal to zero, so that  $\vec{K}$  is indeed a constant vector.

The Runge-Lenz vector enjoys a number of other interesting properties. Using the definitions and the elementary relations from vector analysis, we find

$$\vec{K} \cdot \vec{L} = 0 , \quad (9.11a)$$

$$\begin{aligned} \vec{K}^2 &= \left( \frac{1}{\mu e^2} \right)^2 (\vec{L} \times \vec{p})^2 + \frac{2}{\mu e^2 r} (\vec{L} \times \vec{p}) \cdot \vec{r} + 1 \\ &= \left( \frac{1}{\mu e^2} \right)^2 \vec{L}^2 \vec{p}^2 - \frac{2}{\mu e^2 r} \vec{L}^2 + 1 \\ &= \frac{2H}{\mu e^4} \vec{L}^2 + 1 . \end{aligned} \quad (9.11b)$$

For the elliptical orbit of a classical particle, the Runge-Lenz vector  $\vec{K}$  in fact points into the direction of the *semi major axis of the ellipse*, see the illustration in Figure 4. The length of  $\vec{K}$  is the *eccentricity*  $\varepsilon$  of the ellipse.

In quantum mechanics, the vectors  $\vec{r}$  en  $\vec{p}$  are operators, obeying Heisenberg's commutation relations,

$$\begin{aligned} [p_i, r_j] &= -i\hbar \delta_{ij} , \\ [r_i, r_j] &= [p_i, p_j] = 0 . \end{aligned} \quad (9.12)$$

In the coordinate representation, we describe states in terms of wave functions of the coordinates  $r_i$ . The coordinate operator then acts trivially on the wave function, multiplying it by the value of the coordinate, whereas the momentum operators are given by

$$\vec{p} = \frac{\hbar}{i} \frac{\partial}{\partial \vec{r}} . \quad (9.13)$$

Naturally, also  $H$ ,  $\vec{L}$  and  $\vec{K}$  are operators now, but the definition of  $\vec{K}$  presents us with a new problem, since  $\vec{p}$  and  $\vec{L}$  do not commute,

$$\begin{aligned} [L_i, p_j] &= i\hbar \epsilon_{ijk} p_k , \\ [L_i, r_j] &= i\hbar \epsilon_{ijk} r_k . \end{aligned} \quad (9.14)$$

which follows directly from Eqs. (9.4) and (9.12) using the commutation relation

$$[A, BC] = [A, B]C + B[A, C] . \quad (9.15)$$

In view of Eq. (9.14), it is not obvious in which order the operators  $\vec{L}$  en  $\vec{p}$  have to be placed in the definition of Eq. (9.7) for  $\vec{K}$ .<sup>20</sup> We now propose the following choice for the quantum mechanical definition of  $\vec{K}$ :

$$\vec{K} = \frac{1}{2\mu e^2} \left( \vec{L} \times \vec{p} - \vec{p} \times \vec{L} \right) + \frac{\vec{r}}{r} . \quad (9.16)$$

It is clear that this definition would boil down to Eq. (9.7) if  $\vec{L}$  and  $\vec{p}$  would commute. With this particular choice, the operator  $\vec{K}$  is hermitian.

In quantum mechanics, conserved quantities commute with the Hamiltonian. Let us first check this for the operators  $\vec{L}$ . Using the commutation relations (9.14), and Eq. (9.15), we readily derive that

$$[L_i, H] = 0 , \quad (9.17)$$

$$[L_i, L_j] = i\hbar \epsilon_{ijk} L_k . \quad (9.18)$$

Eq. (9.17) indicates that  $L_i$  is conserved. Eq. (9.18) shows that angular momentum operators obey the same commutation relations as the operators in Eq. (3.28) for the three-dimensional rotation group<sup>21</sup>.

By way of illustration, we give the explicit derivation of Eq. (9.18):

$$\begin{aligned} [L_i, L_j] &= \epsilon_{jkl} [L_i, r_k p_l] \\ &= \epsilon_{jkl} ([L_i, r_k] p_l + r_k [L_i, p_l]) \\ &= i\hbar \epsilon_{jkl} (\epsilon_{ikm} r_m p_l + \epsilon_{ilm} r_k p_m) \\ &= i\hbar (\epsilon_{jlk} \epsilon_{imk} + \epsilon_{jmk} \epsilon_{ikl}) r_m p_l \\ &= i\hbar \epsilon_{ijk} \epsilon_{kml} r_m p_l , \end{aligned} \quad (9.19)$$

<sup>20</sup>The reader may wish to check that this problem does not arise in the definitions (9.1) and (9.4) of  $H$  and  $\vec{L}$ .

<sup>21</sup>Apart from the factors  $\hbar$ ; the operators  $L_i$  in this chapter include the constant  $\hbar$  as commonly used in quantum mechanics.

Where successively Eqs. (9.4), (9.15) and (9.14) were used; furthermore, indices were renamed, and Eq. (3.33) for  $\epsilon$  tensors was employed.

In a similar fashion we can determine commutators containing the operators  $\vec{K}$ . First, by using Eq. (9.13), we verify that

$$\left[\vec{p}, \frac{1}{r}\right] = \frac{\hbar}{i} \left( \frac{\partial}{\partial \vec{r}} \frac{1}{r} - \frac{1}{r} \frac{\partial}{\partial \vec{r}} \right) = i\hbar \frac{\vec{r}}{r^3}. \quad (9.20)$$

Subsequently, we derive that

$$[L_i, K_j] = i\hbar \epsilon_{ijk} K_k \quad (9.21)$$

(a property that also follows from the fact that the  $K_i$  form a vector),

$$[K_i, p_j] = \frac{i\hbar}{\mu e^2} (p_i p_j - p^2 \delta_{ij}) + \frac{i\hbar}{r^3} (r^2 \delta_{ij} - r_i r_j), \quad (9.22)$$

$$\left[K_i, \frac{1}{r}\right] = \frac{i\hbar}{2\mu e^2} \epsilon_{ijk} \frac{L_j r_k + r_k L_j}{r^3}. \quad (9.23)$$

Using eqs. (9.22) and (9.23), it is now not difficult to prove that  $\vec{K}$  commutes with the Hamiltonian,

$$[\vec{K}, H] = 0. \quad (9.24)$$

Finally, we can conclude from the Jacobi identity,

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0, \quad (9.25)$$

that also  $[K_i, K_j]$  must commute with the Hamiltonian, and is therefore conserved. A computation shows that this commutator does not give us a new conserved quantity, but it is a product of two conserved quantities that we already had,<sup>22</sup>

$$[K_i, K_j] = i\hbar \epsilon_{ijk} L_k \left( \frac{-2H}{\mu e^4} \right). \quad (9.26)$$

Since  $L_k$  and  $H$  commute, the order of these operators is immaterial.

Herewith, all relevant commutators have been determined. Next, we must establish whether the classical relations (9.11) remain valid for the case that  $\vec{K}$  and  $\vec{L}$  are operators. For the first equation, we find

$$\vec{K} \cdot \vec{L} = \vec{L} \cdot \vec{K} = \frac{1}{2\mu e^2} \left\{ \vec{L} \cdot (\vec{L} \times \vec{p}) - \vec{L} \cdot (\vec{p} \times \vec{L}) \right\} + \frac{\vec{L} \cdot \vec{r}}{r}. \quad (9.27)$$

---

<sup>22</sup>On general grounds one can prove that the hydrogen atom cannot have more than 5 *independent* conserved quantities. For instance, one can choose the components of the angular momentum and the two coordinates of the Runge-Lenz vector that are orthogonal to  $\vec{L}$ .

Since  $\vec{L} \cdot \vec{r} = 0$  (this continues to hold true in quantum mechanics as one easily checks), only the first two terms contribute, proportional to

$$\epsilon_{ijk} L_i (L_j p_k - p_j L_k) = \epsilon_{ijk} (2 L_i L_j p_k - i\hbar \epsilon_{jkl} L_i p_l) , \quad (9.28)$$

where use was made of Eq. (9.14) to write  $p_i$  at the right of the operators  $L_i$ . Using Eq. (9.18) we write the first term proportional to  $\epsilon_{ijk} \epsilon_{ijl} L_l p_k \propto \vec{L} \cdot \vec{p} = 0$ ; the second term vanishes for the same reason, so that we find the same result as before,

$$\vec{K} \cdot \vec{L} = 0 . \quad (9.29)$$

Eq. (9.11b) no longer holds in the quantum mechanical case, and the exchange of two operators leads to extra terms. The correct result, which requires some more calculations, takes the following form,

$$\vec{K}^2 = \frac{2H}{\mu e^4} (\vec{L}^2 + \hbar^2) + 1 . \quad (9.30)$$

In the limit  $\hbar \rightarrow 0$  this result agrees with (9.11b).

The above results now enable us to determine the energy spectrum of the hydrogen atom. Consider the eigenstates of  $H$ , obeying

$$H | \psi \rangle = E | \psi \rangle , \quad (9.31)$$

where the energy  $E$  will be assumed to be negative,  $E < 0$ , so that we are dealing with bound states. Since the operators  $L_i$  and  $K_i$  commute with the Hamiltonian, all states  $L_i | \psi \rangle$  and  $K_i | \psi \rangle$  also solve the Schrödinger equation (9.31). Within the set (9.31), the operators  $L_i$  and  $K_i$  may be seen to act as the generators of a Lie group, because their commutators are linear combinations of the  $L_i$  and  $K_i$  themselves. Indeed, they generate the Lie group of all transformations that leave the energy  $E$  invariant. The Hamiltonian is a Casimir operator.

In order to analyze the mathematical structure of this Lie group, define the following two linear combinations of the operators  $\vec{L}$  and  $\vec{K}$ ,<sup>23</sup>

$$\vec{L}^\pm = \frac{1}{2} \left( \vec{L} \pm \sqrt{\frac{\mu e^4}{-2E}} \vec{K} \right) . \quad (9.32)$$

These operators commute with the Hamiltonian,

$$[L_i^\pm, H] = 0 . \quad (9.33)$$

---

<sup>23</sup>The fact that  $E$  is negative, ensures that the argument of the square root in these definitions is positive.

Note that  $\vec{L}^+$  and  $\vec{L}^-$  each are vectors, out of which the ladder operators  $L_\pm^+$  and  $L_\pm^-$  of Chapter 5 can be constructed.

Using Eqs. (9.18), (9.21), and (9.26), we prove the following commutation relations

$$[L_i^+, L_j^+] = i\hbar \epsilon_{ijk} L_k^+, \quad (9.34)$$

$$[L_i^-, L_j^-] = i\hbar \epsilon_{ijk} L_k^-, \quad (9.35)$$

$$[L_i^+, L_j^-] = 0. \quad (9.36)$$

In view of these commutation relations, we can regard  $\vec{L}^+$  and  $\vec{L}^-$  as the generators of two separate rotation groups, which mutually commute, and they act upon the states according to

$$|\psi\rangle \rightarrow |\psi'\rangle = e^{i\vec{\alpha}\cdot\vec{L}^+} e^{i\vec{\beta}\cdot\vec{L}^-} |\psi\rangle, \quad (9.37)$$

where the parameters  $\vec{\alpha}$  and  $\vec{\beta}$  describe the two groups generated by  $\vec{L}^+$  en  $\vec{L}^-$ .<sup>24</sup>

If  $|\psi\rangle$  obeys the Schrödinger equation (9.31), then also the transformed states  $|\psi'\rangle$  obey this equation, with the same eigenvalues  $E$ . All these states are to be characterized as representations of the two rotation groups (this product group is often denoted as  $SU(2) \otimes SU(2)$ ). The states should therefore transform according to a  $(2s_+ + 1)$  dimensional (irreducible) representation of the first rotation group. If now we let an element of the second group act on this representation, we find an identical representation of the first group (with this, we mean that each of the states of the original representation stays the same apart from a universal factor). Thus, the second group maps entire representations of the first group into entire representations. Obviously, the converse also holds: the first group maps representations of the second group into the same representations. One therefore is dealing with the tensor product of a representation of the first group and a representation of the second group. Assume we have a  $(2s_+ + 1)$  dimensional representation of the first, and a  $(2s_- + 1)$  dimensional representation of the second group, where  $s_{\pm} = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ . Then we find  $(2s_+ + 1)(2s_- + 1)$  states, to be indicated as  $|m_+, m_-\rangle$ , where  $m_{\pm}$  denotes the eigenvalues of the components  $L_3^{\pm}$ , obeying  $-s_{\pm} \leq m_{\pm} \leq s_{\pm}$ . These states obey

$$\begin{aligned} (\vec{L}^+)^2 |m_+, m_-\rangle &= s_+(s_+ + 1) \hbar^2 |m_+, m_-\rangle, \\ (\vec{L}^-)^2 |m_+, m_-\rangle &= s_-(s_- + 1) \hbar^2 |m_+, m_-\rangle, \end{aligned} \quad (9.38)$$

However, we also have an extra condition for the generators  $\vec{L}^{\pm}$ , following from Eq. (9.29):

$$(\vec{L}^+ + \vec{L}^-) \cdot (\vec{L}^+ - \vec{L}^-) \propto \vec{L} \cdot \vec{K} = 0, \quad (9.39)$$

or

$$(\vec{L}^+)^2 = (\vec{L}^-)^2. \quad (9.40)$$

---

<sup>24</sup>We point out once more that the two groups mutually commute. the two exponentiated operators in Eq. (9.37) can therefore be exchanged.

From this, we conclude that

$$s_+ = s_- = \frac{1}{2}(n-1), \quad (9.41)$$

with  $n = 2s_{\pm} + 1 = 1, 2, \dots$ , so that we are dealing with  $n^2$  states  $|m_+, m_-\rangle$ .

Finally, we can express the Hamiltonian, which is the energy  $E$ , in terms of the operators  $\vec{L}^{\pm}$  using eq. (9.30). We then find

$$-\left(\vec{L}^+ - \vec{L}^-\right)^2 = \frac{\mu e^4}{2E} K^2 = \left(\vec{L}^+ + \vec{L}^-\right)^2 + \hbar^2 + \frac{\mu e^4}{2E}, \quad (9.42)$$

from which

$$-2\left(\vec{L}^+\right)^2 - 2\left(\vec{L}^-\right)^2 = -2(n-1)\left(\frac{1}{2}(n-1) + 1\right)\hbar^2 = \hbar^2 + \frac{\mu e^4}{2E}. \quad (9.43)$$

This implies that the bound states are  $n^2$  fold degenerate, with energies

$$E_n = -\frac{\mu e^4}{2\hbar^2 n^2}. \quad (9.44)$$

The angular momentum operators are given by the operators  $\vec{L}^+ + \vec{L}^-$ . This linear combination acts the same way on both representations in the tensor product defining the states  $|m_+, m_-\rangle$ . Upon rotations in space, generated by the operators  $\vec{L}^+ + \vec{L}^-$ , the states  $|m_+, m_-\rangle$  therefore transform as a tensor product of two representations of the usual rotation group. Such product representations are no longer irreducible, as we saw earlier. From the product properties of representations we can determine which values of the angular momentum occur at a given  $n$ ; this we deduce from the representations that occur in the product of two  $n$  dimensional representations,

$$\mathbf{n} \otimes \mathbf{n} = \mathbf{2n-1} \oplus \mathbf{2n-3} \oplus \dots \oplus \mathbf{3} \oplus \mathbf{1}, \quad (9.45)$$

where  $\mathbf{n}$  denotes the irreducible  $n$ -dimensional representation of the rotation group (being a state with spin  $\frac{1}{2}(n-1)$ ). We do not prove this decomposition laws here; for  $n=3$  this was done for Eq. (4.13), and a similar product was determined in Eq. (8.7).

In conclusion, we can say that the bound states of the hydrogen atom can be determined by the *principal quantum number*  $n$ , which takes the values  $n = 1, 2, \dots$ . For given  $n$ , we have  $n^2$  states with angular momenta  $\ell = 0, 1, \dots, n-1$ , and energy  $E_n$  given by Eq. (9.44).

*References* : O. Klein, Z. für Physik **22** (1924) 109; W. Lenz, Z. für Physik **24** (1924) 197; W. Pauli, Z. für Physik **36** (1926) 336.

isospin-multiplet	spin	parity	isospin	$S$	mass (MeV)
pions, $\pi^\pm, \pi^0$	0	−	1	0	139.6, 135
kaons, $K^0, K^+$	0	−	$\frac{1}{2}$	1	497.7, 493.7
anti kaons, $K^-, \bar{K}^0$	0	−	$\frac{1}{2}$	−1	493.7, 497.7
eta, $\eta^0$	0	−	0	0	547.3
eta', $\eta'$	0	−	0	0	$\approx 960$
nucleons, $n, p$	$\frac{1}{2}$	+	$\frac{1}{2}$	0	939.5, 938.3
delta, $\Delta^-, \Delta^0, \Delta^+, \Delta^{++}$	$\frac{3}{2}$	+	$\frac{3}{2}$	0	$\approx 1235$
sigma, $\Sigma^-, \Sigma^0, \Sigma^+$	$\frac{1}{2}$	+	1	−1	1197.4, 1192.6, 1189.4
sigma*, $\Sigma^{*-}, \Sigma^{*0}, \Sigma^{*+}$	$\frac{3}{2}$	+	1	−1	$\approx 1385$
lambda, $\Lambda^0$	$\frac{1}{2}$	+	0	−1	1115.7
ksi, $\Xi^-, \Xi^0$	$\frac{1}{2}$	+	$\frac{1}{2}$	−2	1321.3, 1315
ksi*, $\Xi^{*-}, \Xi^{*0}$	$\frac{3}{2}$	+	$\frac{1}{2}$	−2	$\approx 1530$
omega, $\Omega^-$	$\frac{3}{2}$	+	0	−3	1672.4

Table 2: Various isospin-multiplets and their masses.

## 10. The group $SU(3)$

The Lie group  $SU(3)$  has played an important role in the realization of quark theory. In studying the *hadrons* (the strongly interacting subatomic particles) it was noted that they appear to fit in a scheme of multiplets that is more complicated than the isospin multiplets that we have seen. Various isospin multiplets appeared to fit in larger frames.

When it was realized that this might again be attributed to one or more Lie groups, various possibilities were tried out. This led to  $SU(3)$ . Today we know why exactly this group applies. Deeply hidden inside these subatomic particles are their constituents, the *quarks*. There are three species of these. We call them  $u$ ,  $d$  and  $s$ . Of these,  $u$  and  $d$  form an isospin doublet. So they are in a  $I = \frac{1}{2}$  representation of  $SU(2)$ . The third quark,  $s$ , is an isospin singlet. We write the quark wave functions as

$$|\psi\rangle = \begin{pmatrix} u \\ d \\ s \end{pmatrix}, \quad (10.46)$$

and we assume the physical properties of these particles to be (approximately) invariant under the *unitary* transformations of these 3-vectors. Therefore, we regard (10.46) as the *fundamental representation* of  $SU(3)$ , where 3 stands for the dimension of the vectors,  $U$  stands for “unitary”, and  $S$  for “special”: we restrict ourselves to matrices with determinant 1.

If we restrict ourselves to those transformations that transform  $u$  and  $d$  into one another, but leave  $s$  invariant, then we have the subgroup  $SU(2)$  of  $SU(3)$ . In  $SU(2)$  we saw that if you try to form new representations starting from products of representations with integral  $\ell$ , you only obtain representations where  $\ell$  is again an integer. These are

all representations that map the element

$$-\mathbb{I} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (10.47)$$

onto the unit matrix  $+\mathbb{I}$ .

Something similar happens to  $SU(3)$ . This group has three special elements that *commute with all elements of the same group* (we call the subgroup formed by these special elements the *center* of the group):

$$\mathbb{I}; \quad e^{\frac{2\pi i}{3}} \mathbb{I} = \begin{pmatrix} e^{\frac{2\pi i}{3}} & & \\ & e^{\frac{2\pi i}{3}} & \\ & & e^{\frac{2\pi i}{3}} \end{pmatrix}; \quad e^{-\frac{2\pi i}{3}} \mathbb{I} = \begin{pmatrix} e^{-\frac{2\pi i}{3}} & & \\ & e^{-\frac{2\pi i}{3}} & \\ & & e^{-\frac{2\pi i}{3}} \end{pmatrix}. \quad (10.48)$$

We now distinguish two types of representations of  $SU(3)$ : the *regular* and the *exotic* representations. Regular representations are the ones that map all three elements (10.48) of the center onto the unit matrix. All other representations are exotic. The elementary representation itself is exotic. The representations formed by the  $3 \times 3$  matrices  $U$  of the  $SU(3)$  transformations, the *adjoint representation*, is regular, because they transform according to

$$U \rightarrow U' = X U X^{-1}, \quad (10.49)$$

and if we take  $X$  to be one of the center elements, we see that  $U$  does not change.

Writing  $U = e^{iT}$ , we see that unitarity of  $U$  follows if we demand that  $T$  be hermitian:  $T = T^\dagger$ , and the fact that  $\det(U) = 1$  requires that we should demand  $\text{Tr}(T) = 0$ . The  $3 \times 3$  matrices have 9 complex elements. Hermiticity of  $T$  restricts this number to 9 *real* parameters. From  $\text{Tr}(T) = 0$  it follows that 8 of these are independent. Therefore, the adjoint representation spans an 8 dimensional real space. Just as in Chapter 6, Eq. (6.45), we write this representation as  $Y_\beta^\alpha$ , but now the indices  $\alpha, \beta$  can take three values. Selecting the elements where  $\alpha$  and  $\beta$  both take the values 1 or 2, we reobtain the three-dimensional representation  $\mathbf{3}$ , with  $I = 1$ , of  $SU(2)$ . If  $\alpha = 1, 2$  and  $\beta = 3$ , we have the complex, fundamental  $\mathbf{2}$  representation of  $SU(2)$ , and we find the same back if  $\alpha = 3$  and  $\beta = 1, 2$ . Finally, there is a  $\mathbf{1}$  representation of  $SU(2)$  on the diagonal. Since the  $\mathbf{2}$  representation is complex while the others are real, we count the complex one twice, but we can also say that it occurs twice, once in  $Y_3^\alpha$  and once in  $Y_\alpha^3$ :

$$\mathbf{8}_{SU(3)} = (\mathbf{3} + \mathbf{2} + \mathbf{2} + \mathbf{1})_{SU(2)}. \quad (10.50)$$

The other representations of  $SU(3)$  are found by multiplying symmetric or skew-symmetric combinations of the fundamental representation. This way, “tensors”  $Y_{\beta_1 \beta_2 \dots}^{\alpha_1 \alpha_2 \dots}$  are obtained, and from these we extract the irreducible combinations, which can be done by symmetrization or skew symmetrization, or by multiplication with *invariant tensors*. If we only want to use the fundamental indices (see later), one finds that there are only three invariant tensors:  $\delta_\beta^\alpha$ ,  $\varepsilon_{\alpha\beta\gamma}$  and  $\varepsilon^{\alpha\beta\gamma}$ . The latter two are invariant, just like the  $\varepsilon_{\alpha\beta}$  and  $\varepsilon^{\alpha\beta}$  in  $SU(2)$ , because of the constraint  $\det(X) = 1$  imposed on the transformation

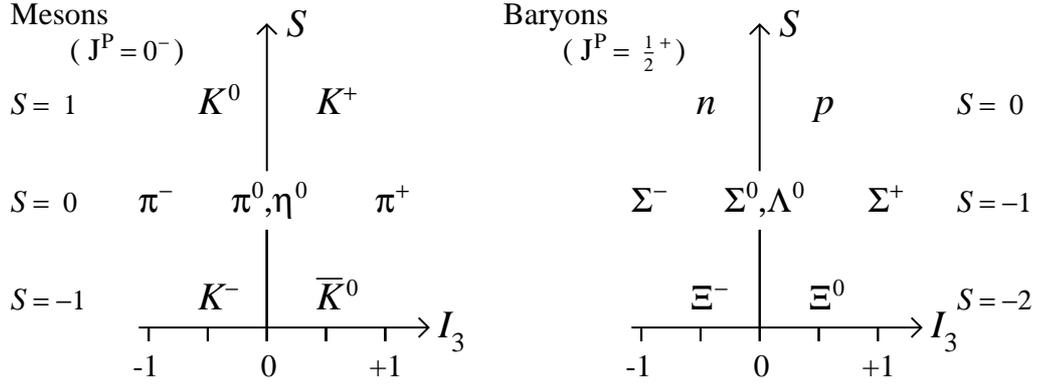


Figure 5: The meson octet with spin 0 and parity  $-$ , and the baryon-octet with spin  $\frac{1}{2}$  and parity  $+$

matrices  $X$  in  $SU(3)$ . We can use the  $\varepsilon_{\alpha\beta\gamma}$  tensors to turn all indices into superscripts (or all subscripts), so it suffices to look for all representations with only superscripts.

Unlike the situation in  $SU(2)$ , the representations generated by  $\phi^\alpha$  and  $\phi_\alpha$  are not equivalent. We call these two representations  $\mathbf{3}$  and  $\bar{\mathbf{3}}$ . The  $\bar{\mathbf{3}}$  can also be written as an skew-symmetric  $Y^{\alpha\beta}$ . The symmetric representation  $Y^{\alpha\beta}$  is a complex  $\mathbf{6}$  representation. The first non exotic representation, encountered after the  $\mathbf{8}$ , is a  $\mathbf{10}$ , formed by the completely symmetric tensors  $Y^{\alpha\beta\gamma}$ . This *decuplet* is built from  $SU(2)$  representations as follows:

$$\mathbf{10}_{SU(3)} = (\mathbf{4} + \mathbf{3} + \mathbf{2} + \mathbf{1})_{SU(2)} . \quad (10.51)$$

The set of all regular (non-exotic) representations can be found as follows: the number of superscript indices minus the number of subscript indices must be a multiple of 3.

In hadron physics, particles of a certain type can be arranged according to isospin  $I_3$  and a new quantum number  $S$ , called “strangeness”. The patterns obtained have been sketched in Figures 5 and 6.

We recognize two octets and a decuplet. Writing the meson octet as  $M^\alpha_\beta$  and the baryon decuplet as  $B^{\alpha\beta\gamma}$ , we see that strangeness  $S$  may correspond to the number of subscripts with the value 3, minus the number of superscripts taking the value 3. Now we know that we are dealing here with wave functions of the type  $|\psi^\alpha \cdot \psi_\beta\rangle$  where  $|\psi^\alpha\rangle$  indicate quarks of type  $\alpha$  and  $|\psi_\beta\rangle$  anti quarks of type  $\beta$ . Here,  $\alpha = 1$  and  $2$  form the isodoublet  $u$  en  $d$ , and  $\alpha = 3$  is the  $s$ -quark, that has been assigned strangeness  $-1$ . A ninth meson, with  $M^\alpha_\beta = \delta^\alpha_\beta$  also exists. This is the  $\eta'$ , whose mass deviates more from the others. The baryons are all written as  $B^{\alpha\beta\gamma}$ . The octet, which can be written as

$$B^{\alpha\beta\gamma} = \varepsilon^{\alpha\beta\kappa} B^\gamma_\kappa ; \quad B^\alpha_\alpha = 0 , \quad (10.52)$$

can also be defined by

$$B^{\alpha\beta\gamma} = -B^{\beta\alpha\gamma} ; \quad \varepsilon_{\alpha\beta\gamma} B^{\alpha\beta\gamma} = 0 . \quad (10.53)$$

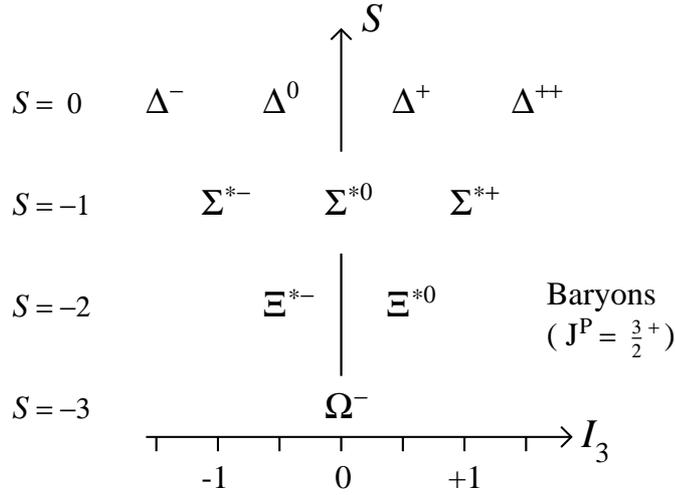


Figure 6: The baryon decuplet with spin  $\frac{3}{2}$  and parity  $P = +$ .

$SU(3)$  symmetry is broken more strongly than isospin, because the masses of these particles in different isospin multiplets differ much more than the masses within one isospin multiplet (see Table 2). Nevertheless this symmetry is of great importance. M. Gell-Mann discovered the decuplet when it was not yet complete, which enabled him to predict the missing particle  $\Omega^-$ . A striking feature is that all representations of  $SU(3)$  that we encounter among the hadrons in Nature are regular representations, the exotic representations are apparently forbidden.

We can link the baryon decuplet with the baryon octet. Consider the basis of ‘quark states’, where now also ordinary spin is taken into account. Every quark may now have  $S_z = +1$  or  $-1$ , so instead of 3 we now have 6 quark states. If we were allowed to transform all six quark states into one another, we get the Lie group  $SU(6)$ . The symmetric representation  $|\psi^{ABC}\rangle$  of  $SU(6)$  has  $6 \cdot 7 \cdot 8/3! = 56$  elements. The  $SU(6)$  indices each consist of one  $SU(3)$  index and one  $SU(2)$  index for spin, which we will call  $\sigma$ :

$$A = (\alpha, \sigma_1), \quad B = (\beta, \sigma_2), \quad C = (\gamma, \sigma_3). \quad (10.54)$$

Taking all states symmetric in  $\alpha, \beta$ , and  $\gamma$ , we easily establish that they also have to be symmetric under interchange of the ordinary spin indices  $\sigma_1, \sigma_2$  and  $\sigma_3$ . So, these states have spin  $\frac{3}{2}$ , exactly as in our baryonic decuplet. Every particle in the decuplet has 4 spin states, so here we have  $10 \cdot 4 = 40$  basis elements. Orthogonally to those, we have  $56 - 40 = 16$  elements left, exactly an octet with spin  $\frac{1}{2}$ , because each member of the octet has 2 spin states.

$SU(6)$  is even less well conserved than  $SU(3)$ , as exemplified by the fact that the masses of the baryon decuplet are considerably larger than those of the baryon octet.

Two questions have not yet been answered: 1) Why do the hadrons only occur in regular (non-exotic)  $SU(3)$  representations, and 2) How can it be that quarks with spin

$\frac{1}{2}$  are in a *symmetric* wave function; having spin  $\frac{1}{2}$  one would have expected that quarks are fermions, which should occur only in completely *skew-symmetric* wave functions, in view of Pauli's exclusion principle. The answer to both questions is that quarks carry yet another, 'internal' quantum number, called *color*, which again can take three values. The color theory can explain the nature of the forces between the quarks. Attractive forces between quarks with different values for the color index are so great that these colors must neutralize each other completely. This theory will not be further explained here, but it is known to imply that all hadrons must form a color singlet representation. The wave function must therefore be completely skew-symmetric under exchanges of the color indices. In agreement with Pauli's principle, the entire wave functions of the hadrons are completely skew-symmetric under the interchange of two quarks. This leaves us with completely symmetric wave functions when only two  $SU(6)$  indices are interchanged.

To describe the generators of an  $SU(3)$  transformation, we use the hermitian matrices  $\lambda_a$ ,  $\alpha = 1, \dots, 8$ :

$$X = \exp(\frac{1}{2}i\theta^a \lambda_a) . \quad (10.55)$$

$\lambda_a$ , called the *Gell-Mann matrices*, are a generalization of the  $SU(2)$  matrices  $\tau_i$ :

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & \cdot \\ 1 & 0 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & \cdot \\ i & 0 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & \cdot \\ 0 & -1 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & \cdot & 1 \\ \cdot & \cdot & \cdot \\ 1 & \cdot & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & \cdot & -i \\ \cdot & \cdot & \cdot \\ i & \cdot & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & 0 & 1 \\ \cdot & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & 0 & -i \\ \cdot & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (10.56)$$

The  $\lambda_a$  are normalized according to

$$\text{Tr}(\lambda_a \lambda_b) = 2\delta_{ab} , \quad (10.57)$$

and they obey the commutation relations

$$[\frac{1}{2}\lambda_a, \frac{1}{2}\lambda_b] = if_{abc}(\frac{1}{2}\lambda_c) . \quad (10.58)$$

The list of the structure constants  $f_{abc}$  is

$$\begin{aligned} f_{123} &= 1 \\ f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} &= \frac{1}{2} \\ f_{458} = f_{678} &= \frac{1}{2}\sqrt{3} ; \end{aligned} \quad (10.59)$$

the remaining ones follow from the complete antisymmetry of these coefficients:

$$f_{abc} = f_{bca} = -f_{bac} , \quad \text{etc.} \quad (10.60)$$

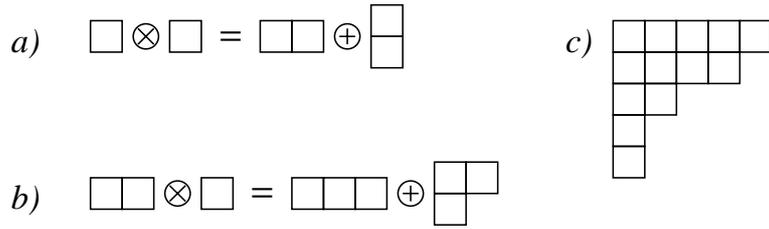


Figure 7: Young tableaux

$f_{abc}$  en  $\delta_{ab}$  are not the only invariant tensors of the  $\mathbf{8}$  representation. We can also consider the *anticommutators* of the  $\lambda$ 's, to find:

$$\left\{ \frac{1}{2}\lambda_a, \frac{1}{2}\lambda_b \right\} = \frac{1}{3}\delta_{ab} + d_{abc}\left(\frac{1}{2}\lambda_c\right). \quad (10.61)$$

The anticommutator is defined as  $\{A, B\} = AB + BA$ . The  $d_{abc}$  are completely symmetric:

$$d_{abc} = d_{bca} = d_{bac}, \quad \text{etc.} \quad (10.62)$$

and they can also be listed:

$$\begin{aligned} d_{118} = d_{228} = d_{338} = -d_{888} &= \frac{1}{\sqrt{3}} \\ d_{146} = d_{157} = d_{256} = d_{344} = d_{355} &= \frac{1}{2} \\ d_{247} = d_{366} = d_{377} &= -\frac{1}{2} \\ d_{448} = d_{558} = d_{668} = d_{778} &= -\frac{1}{2\sqrt{3}} \end{aligned} \quad (10.63)$$

The tensors  $\delta_{ab}$ ,  $f_{abc}$  en  $d_{abc}$  are the only independent invariant tensors in the  $\mathbf{8}$  representation.

## 11. Representations of $SU(N)$ ; Young tableaux

We have already seen a glimpse of  $SU(6)$ . Larger Lie groups also occur in physics. Take  $SU(N)$  for any value of  $N$ . General methods exist to find all their representations. First, we decide to replace all subscript indices by superscripts using the invariant tensor  $\varepsilon^{\alpha_1\alpha_2\cdots\alpha_N}$ .

The fundamental representation, corresponding to the  $N$  dimensional ‘spinor’  $\phi^\alpha$  in  $SU(N)$ , is indicated by a box. The product of two of these representations is depicted as two boxes, see Figure 7a. If the representation is symmetric, we picture the boxes next to each other horizontally; the skew-symmetric component is depicted by the vertical arrangement. In Fig. 7b, we see how larger representations can be constructed. Every box represents an index. If we have a complicated representation, such as in Fig. 7c, then the associated tensor is symmetric under all permutations of the indices within one horizontal row of boxes. In the vertical direction the rule is somewhat more complicated; it is completely determined by the demand that all allowed representations must be independent,

so that two tensors that are described by different Young tableaux are always orthogonal to one another. It then turns out that we arrive at the following rule:

- i* All independent, irreducible representations of  $SU(N)$  can be distinguished by the fact that they have different Young tableaux, and every Young tableau represents one irreducible representation, provided that:
- ii* Every horizontal row in a Young tableau may not have more boxes than the row above it, and every column may not have more boxes than the column left of it.
- iii* There cannot be more than  $N - 1$  rows, so the columns may not have more than  $N - 1$  boxes.

A typical example is the Young tableau of Fig. 7c.

## 12. Beyond these notes

There is a large amount of literature on Lie groups. The problem of classifying *all* inequivalent Lie groups has been solved (W. Killing and E. Cartan). One first defines the *simple* Lie groups, which are the Lie groups that cannot be written as the product of two non-trivial mutually commuting Lie groups. Then it was found that there are four infinite series of these:

The groups  $A_n = SU(n + 1)$  ;  $n \geq 1$  ,

The groups  $B_n = SO(2n + 1)$  ;  $n \geq 2$  ,

The groups  $C_n = Sp(2n)$  ,  $n \geq 3$  ,

The groups  $D_n = SO(2n)$  ,  $n \geq 4$  ;

$n$  may take values smaller than the limits indicated, but these groups are not independent of the others given, for instance: the algebra of  $SO(3)$  coincides with that of  $SU(2)$ , as was described in these notes. We write  $SO(3) \cong SU(2)$ . Similarly, one has  $SO(4) \cong SO(3) \otimes SO(3)$  and  $SO(6) \cong SU(4)$ .

The groups  $Sp(2n)$  are the so-called *symplectic* groups.

Besides these regular series, there are five more simple Lie groups that are not in one of these series, called the *exceptional groups*,  $E_6$ ,  $E_7$ ,  $E_8$ ,  $F_4$ , and  $G_2$ . The  $E_8$  algebra is the largest and most complicated of these exceptional cases. This group is so large that it could easily encompass all known particles and fields, and since it also plays a role in super string theories, it receives a lot of interest from physicists.

The Lie algebra can also be extended to algebras where, at various places, the commutators are replaced by *anticommutators*. One then gets the *graded* Lie groups. These play a role in supersymmetric particle theories.

## Appendix A

### Summary of the properties of matrices

A square matrix of dimension  $n$  is a table of  $n \times n$  elements  $a_{ij}$ :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

The matrix elements  $a_{ij}$  are (possibly complex) numbers. Suppose we have an other  $n \times n$  matrix  $B$  with elements  $b_{ij}$ .

*Definition:* The sum of two matrices  $A$  and  $B$  is a matrix  $C$  with matrix elements given by

$$c_{ij} = a_{ij} + b_{ij}. \quad (\text{A.1})$$

Their difference  $C'$  is given by the matrix elements  $c'_{ij} = a_{ij} - b_{ij}$ . We write:

$$C = A + B, \quad \text{or} \quad C' = A - B.$$

From Eq. (A.1) it follows that the summation is commutative:  $A + B = B + A$ .

*Definition:* The product of the matrices  $A$  and  $B$  is a matrix  $C$  with matrix elements  $c_{ij}$  given by

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}. \quad (\text{A.2})$$

We write:

$$C = AB. \quad (\text{A.3})$$

Explanation: Eq. (A.2) defines the matrix elements  $c_{ij}$  if  $a_{ij}$  and  $b_{ij}$  are given. Since  $a_{ij}$  and  $b_{ij}$  are ordinary numbers, it does not matter in which order they are written. This does not hold, however, for products of matrices, where in general  $C = AB \neq BA$ . Note that  $C' = BA$  is a matrix whose elements are given by

$$c'_{ij} = \sum_k b_{ik} a_{kj}, \quad (\text{A.4})$$

which differs from Eq. (A.2). We say that matrix multiplication is *non commutative*. It is *associative*:  $A(BC) = (AB)C$ .

*Definition:* The commutator of two matrices  $A$  and  $B$  is given by

$$[A, B] \equiv AB - BA. \quad (\text{A.5})$$

We say that  $A$  and  $B$  commute if the commutator is equal to zero.

*Definition:* The product of a (possibly complex) number  $x$  with a matrix  $A$  is a matrix  $B$  with elements:

$$b_{ij} = x a_{ij} . \quad (\text{A.6})$$

Notation:

$$B = xA . \quad (\text{A.7})$$

From Eq. (A.6), it follows that

$$B^2 = x^2 A^2 ,$$

since

$$\sum_k b_{ik} b_{kj} = x^2 \sum_k a_{ik} a_{kj} .$$

*Definition:* The unit matrix  $\mathbf{1}$  is defined as

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} ,$$

and gives  $\mathbf{1} A = A \mathbf{1} = A$  for every matrix  $A$ .

For every  $n \times n$  matrix we can define its *determinant*, being a (possibly complex) number. It is defined recursively by taking out one of its rows, say the  $i^{\text{th}}$  row:

$$\text{Det}(A) = \sum_{k=1}^n a_{ik} (-1)^{k+i} \text{Det}(A(i, k)) , \quad (\text{A.8})$$

where the  $(n-1) \times (n-1)$  matrix  $A(i, k)$  is obtained from  $A$  by removing the  $i^{\text{th}}$  row and the  $k^{\text{th}}$  column. The determinant of a  $0 \times 0$  matrix is 1.

Many properties of determinants are easy to derive: if  $C = AB$  then

$$\text{Det}(C) = \text{Det}(AB) = \text{Det}(A) \text{Det}(B) . \quad (\text{A.9})$$

If  $C = xA$  with  $x$  a number and  $A$  an  $n \times n$  matrix,

$$\text{Det}(C) = \text{Det}(xA) = x^n \text{Det}(A) . \quad (\text{A.10})$$

Furthermore: the inverse of a matrix  $A$  is a matrix  $B$  such that  $BA = AB = \mathbf{1}$ . Notation:  $B = A^{-1}$ . The matrix elements of  $B$  are (note:  $A(j, i)$ , not  $A(i, j)$ ):

$$b_{ij} = \frac{(-1)^{i+j} \text{Det}(A(j, i))}{\text{Det}(A)} . \quad (\text{A.11})$$

from this, it follows that a matrix  $A$  has an inverse if  $\text{Det}(A) \neq 0$ . An other definition of the determinant makes use of the  $\epsilon$  symbols, defined by

$$\begin{aligned} \epsilon_{i_1 \dots i_n} &= 0 && \text{if in } i_1 \dots i_n \text{ two or more of the indices are equal,} \\ \epsilon_{i_1 \dots i_n} &= \pm 1 && \text{if in } i_1 \dots i_n \text{ no two indices are equal, in which case we have} \\ \epsilon_{i_1 \dots i_n} &= 1 && \text{if } i_1 \dots i_n \text{ are an even permutation of } 123 \dots n, \\ \epsilon_{i_1 \dots i_n} &= -1 && \text{if } i_1 \dots i_n \text{ are an odd permutation of } 123 \dots n. \text{ In particular,} \\ \epsilon_{123 \dots n} &= +1. \end{aligned}$$

The determinant of a matrix  $A$  is then given by

$$A_{i_1 j_1} A_{i_2 j_2} \cdots A_{i_n j_n} \epsilon_{j_1 \dots j_n} = \text{Det}(A) \epsilon_{i_1 \dots i_n}.$$

Finally, for a given matrix  $A$ , the transpose  $\tilde{A}$  of a matrix  $A$  is defined by the matrix elements

$$\tilde{a}_{ij} = a_{ji},$$

and if  $A^{-1}$  is the inverse of a matrix  $A$ , one has

$$\widetilde{AB} = \tilde{B} \tilde{A}, \quad (AB)^{-1} = B^{-1} A^{-1}. \quad (\text{A.12})$$

The first equation follows from the definition (A.2) of the matrix product, and the second formula follows from  $\mathbf{1} = (AB)(AB)^{-1} = AB B^{-1} A^{-1} = \mathbf{1}$ .

## Appendix B

### Differentiation of matrices

Matrix elements can be functions of one or more variables. These functions may be differentiable. In this case, one can define the derivative of a matrix.

*Definition:* The derivative of a matrix  $A$  is a matrix whose elements are the derivatives of the matrix elements of  $A$ .

Thus, the matrix  $A'$  with

$$A'(x) = \frac{dA(x)}{dx}, \quad (\text{B.1})$$

has the matrix elements

$$a'_{ij}(x) = \frac{da_{ij}(x)}{dx}. \quad (\text{B.2})$$

In general, the derivative of a matrix will not commute with the matrix itself. Apart from this fact, one may use all formulae that are familiar from differentiation theory, and they follow from the definition (B.1). For example,

$$\frac{d(A+B)}{dx} = \frac{dA}{dx} + \frac{dB}{dx}, \quad (\text{B.3})$$

and

$$\frac{d(AB)}{dx} = \frac{dA}{dx}B + A\frac{dB}{dx} . \quad (\text{B.4})$$

The latter result follows from the chain rule

$$\frac{d}{dx} \sum_k a_{ik} b_{kj} = \sum_k a'_{ik} b_{kj} + \sum_k a_{ik} b'_{kj} . \quad (\text{B.5})$$

Similarly,

$$\frac{dA^n}{dx} = A^{n-1}A' + A^{n-2}A'A + \cdots + A'A^{n-1} . \quad (\text{B.6})$$

Note that this is *not* always equal to  $nA^{n-1}A'$ , because  $A$  and  $A'$  do not necessarily commute.

The derivative of the inverse of a matrix  $A$  can be found the following way. Since  $A^{-1}A = \mathbf{1}$ , we have

$$0 = \frac{d}{dx}(A^{-1}A) = \frac{dA^{-1}}{dx}A + A^{-1}\frac{dA}{dx} ,$$

and consequently,

$$\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1} . \quad (\text{B.7})$$

## Appendix C

### Functions of matrices

Functions based on additions and multiplications can readily be defined for matrices. But many more functions are used. In particular, the function  $e^x$  is of importance. There are two possible definitions.

Let a matrix  $A$  be given. The matrix  $e^A$  is then defined by

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k , \quad (\text{C.1})$$

where  $A^0 = \mathbf{1}$ . A second definition is

$$e^A = \lim_{m \rightarrow \infty} \left( \mathbf{1} + \frac{1}{m}A \right)^m . \quad (\text{C.2})$$

The proof that the two definitions (C.1) and (C.2) coincide is just like it goes for ordinary exponentials, by applying the binomial expansion of Eq. (C.2).

Eq. (C.2) allows us to derive an important relation:

$$\text{Det}(e^A) = \text{Lim}_{m \rightarrow \infty} \left\{ \text{Det}\left(\mathbf{1} + \frac{1}{m}A\right) \right\}^m . \quad (\text{C.3})$$

We compute the determinant between brackets by ignoring contributions of order  $1/m^2$  or smaller. For instance, first consider a  $2 \times 2$  matrix. There:

$$\mathbf{1} + \frac{1}{m}A = \begin{pmatrix} 1 + \frac{a_{11}}{m} & \frac{a_{12}}{m} \\ \frac{a_{21}}{m} & 1 + \frac{a_{22}}{m} \end{pmatrix} .$$

If, in the calculation of the determinant, we ignore the contributions of order  $1/m^2$  and higher, the off-diagonal elements can be ignored (in a determinant, there are no terms containing only one off-diagonal element as a factor). In this approximation, the determinant is simply the product of the diagonal elements:

$$\begin{aligned} \text{Det} \left( \mathbf{1} + \frac{1}{m}A \right) &\approx \left( 1 + \frac{a_{11}}{m} \right) \left( 1 + \frac{a_{22}}{m} \right) \\ &\approx 1 + \frac{1}{m} (a_{11} + a_{22}) + \mathcal{O} \left( \frac{1}{m^2} \right) . \end{aligned}$$

In general, one finds:

$$\text{Det} \left( \mathbf{1} + \frac{1}{m}A \right) = 1 + \frac{1}{m} \text{Tr} (A) + \mathcal{O} \left( \frac{1}{m^2} \right) ,$$

where  $\text{Tr} (A)$  is the *trace* of the matrix  $A$ , that is, the sum of its diagonal elements:

$$\text{Tr} (A) = \sum_k a_{kk} .$$

Substituting this in Eq. (C.3), gives

$$\text{Det}(e^A) = \lim_{m \rightarrow \infty} \left\{ 1 + \frac{1}{m} \text{Tr} (A) + \mathcal{O} \left( \frac{1}{m^2} \right) \right\}^m = e^{\text{Tr}(A)} ,$$

and thus we arrive at:

$$\text{Det}(e^A) = e^{\text{Tr}(A)} . \tag{C.4}$$

## Appendix D

### The Campbell-Baker-Hausdorff formula

Let there be given two matrices  $A$  and  $B$ . The *CBH* formula tells us that there exists a matrix  $C$  such that

$$e^A e^B = e^C , \tag{D.1}$$

where  $C = A + B + \frac{1}{2}[A, B] +$  repeated commutators of  $A$  en  $B$ . Repeated commutators are expressions of the following type:

$$\begin{aligned} [[A, B], B] &= [A, B]B - B[A, B] \\ &= ABB - BAB - BAB + BBA , \end{aligned}$$

$$[[A, [A, B]], B], \quad [[[A, B], B], B], \quad \text{etc.}$$

The numeric factors that come with the coefficients of these repeated commutators are not so easy to find. For our purposes, it suffices to establish that  $C$  indeed is of the form (D.1). We shall prove this and also derive the coefficients of the next order, containing either commutators with two  $B$ 's and one  $A$ , or two  $A$ 's and one  $B$ .

To begin with, we introduce a variable  $x$ , and consider

$$e^{xA}e^{xB} = e^{C(x)}. \quad (\text{D.2})$$

So, the  $x$  dependence at the l.h.s. of (D.2) is simple and explicit. However,  $C(x)$  will be a complicated function of  $x$ . In fact, the series expansion of  $C(x)$  with respect to  $x$  will at the same time be the expansion of  $C$  in powers of  $A$  and  $B$ . Therefore, it will be useful to try to identify the function  $C(x)$ . This we do by differentiating Eq. (D.2) left and right with respect to  $x$ . This gives us an equation from which  $C(x)$  can be solved, yielding a power series. First, we need to prove some more identities.

Consider

$$H(y) = e^{-yF}G e^{yF}, \quad (\text{D.3})$$

where  $y$  is a variable and  $F$  and  $G$  are matrices not depending on  $y$ .  $H$  can be written in the form of a power expansion:

$$H = H_0 + yH_1 + \frac{y^2}{2!}H_2 + \frac{y^3}{3!}H_3 + \dots. \quad (\text{D.4})$$

Here, the factors  $2!$ ,  $3!$ ,  $\dots$  have been included for later convenience. we shall compute the matrices  $H_n$ . Differentiation of Eq. (D.3) to  $y$  yields:

$$\frac{dH(y)}{dy} = \left( \frac{de^{-yF}}{dy} \right) G e^{yF} + e^{-yF} G \frac{de^{yF}}{dy}. \quad (\text{D.5})$$

Using

$$\frac{d}{dy}e^{yF} = F e^{yF} = e^{yF} F, \quad (\text{D.6})$$

we find

$$\frac{dH(y)}{dy} = -F e^{-yF} G e^{yF} + e^{-yF} G e^{yF} F = [H, F], \quad (\text{D.7})$$

and substituting this in the power series (D.4),

$$H_1 + y H_2 + \frac{y^2}{2!} H_3 + \dots = [H_0, F] + y[H_1, F] + \frac{y^2}{2!}[H_2, F] + \dots. \quad (\text{D.8})$$

Since  $H_0 = G$ , identifying the coefficients for equal powers of  $y$  gives us:

$$H_n = [H_{n-1}, F] = [[[\dots [G, F], F], F] \dots F]. \quad (\text{D.9})$$

Substituting this in Eq. (D.4) gives for  $y = 1$  :

$$e^{-F} G e^F = G + [G, F] + \frac{1}{2!} [[G, F], F] + \frac{1}{3!} [[[G, F], F], F] + \dots . \quad (\text{D.10})$$

Let us introduce a shorthand notation:

$$H_n = \{G, F^n\} , \quad (\text{D.11})$$

and, even shorter, we rewrite Eq. (D.10) as

$$e^{-F} G e^F = \{G, e^F\} , \quad (\text{D.12})$$

with which we mean that  $e^F$  must first be replaced by its power series expansion. Let now  $F$  be some function of a variable  $x$  . Consider

$$e^{-yF(x)} \frac{d}{dx} e^{yF(x)} . \quad (\text{D.13})$$

In a way similar to the above, one derives

$$\begin{aligned} e^{-F(x)} \frac{d}{dx} e^{F(x)} &= F' + \frac{1}{2!} [F', F] + \frac{1}{3!} [[F', F], F] + \dots \\ &= \left\{ F', \frac{e^F - 1}{F} \right\} . \end{aligned} \quad (\text{D.14})$$

In fact, Eq. (D.14) is what one gets if in Eq. (D.10),  $G$  is replaced by  $\frac{d}{dx}$  , because in that case,

$$[G, F] = \frac{d}{dx} F - F \frac{d}{dx} \quad (\text{D.15})$$

Here, it was assumed that  $d/dx$  acts to the right on whatever else one has to the right of the commutator  $[G, F]$  .

According to the chain rule,

$$\frac{d}{dx} F = \frac{dF}{dx} + F \frac{d}{dx} = F' + F \frac{d}{dx} , \quad (\text{D.16})$$

so that

$$[G, F] = F' . \quad (\text{D.17})$$

From Eq. (D.14), it follows that

$$\frac{d}{dx} e^{F(x)} = e^{F(x)} \left\{ F', \frac{e^F - 1}{F} \right\} . \quad (\text{D.18})$$

We now return to Eq. (D.2). Differentiation with respect to  $x$  gives, in view of Eq. (D.18):

$$e^{xA} A e^{xB} + e^{xA} B e^{xB} = e^{C(x)} \left\{ C', \frac{e^C - 1}{C} \right\} , \quad (\text{D.19})$$

or

$$e^{xA} e^{xB} (e^{-xB} A e^{xB} + e^{-xB} B e^{xB}) = e^{C(x)} \left\{ C', \frac{e^C - 1}{C} \right\} . \quad (\text{D.20})$$

The factors outside the brackets are the same left and right, in view of Eq. (D.2). Because  $Be^{xB} = e^{xB}B$ , the second term within the brackets is nothing but  $B$ . To the first term, we can apply Eq. (D.12). Thus, we obtain:

$$\{C', \frac{e^C - 1}{C}\} = \{A, e^{xB}\} + B . \quad (\text{D.21})$$

From this equation,  $C$  can be solved as a power expansion. The strength of our abstract notation shows if we temporarily call the r.h.s. of this equation  $H$ :

$$\begin{aligned} C' &= \{H, \frac{C}{e^C - 1}\} = \{H, 1 - \frac{C}{2} + \frac{C^2}{12} + \dots\} \\ &= H - \frac{1}{2}[H, C] + \frac{1}{12}[[H, C], C] + \dots , \end{aligned} \quad (\text{D.22})$$

where we used the power expansion of  $\frac{C}{e^C - 1} = \frac{1}{1 + \frac{1}{2}C + \frac{1}{6}C^2 + \dots} = 1 - \frac{1}{2}C + \frac{1}{12}C^2 \dots$ .

Let us write

$$\begin{aligned} C &= C_0 + xC_1 + \frac{x^2}{2!}C_2 + \frac{x^3}{3!}C_3 + \dots , \\ C' &= C_1 + xC_2 + \frac{x^2}{2!}C_3 + \dots , \\ H &= H_0 + xH_1 + \frac{x^2}{2!}H_2 = A + B + x[A, B] + \frac{x^2}{2!}[[A, B], B] + \dots . \end{aligned} \quad (\text{D.23})$$

We can then compare equal powers of  $x$  in Eqs. (D.22) and (D.23). First of all:

$$C_0 = C(0) = 0 . \quad (\text{D.24})$$

The terms independent of  $x$  give

$$C_1 = H_0 = A + B , \quad (\text{D.25})$$

and the terms linear in  $x$ :

$$C_2 = H_1 - \frac{1}{2}[H_0, C_1] = [A, B] . \quad (\text{D.26})$$

Continuing this way, all other coefficients  $C_n$  can be found:

$$\begin{aligned} \frac{1}{2}C_3 &= \frac{1}{2}H_2 - \frac{1}{4}[H_0, C_2] - \frac{1}{2}[H_1, C_1] , \\ C_3 &= [[A, B], B] - \frac{1}{2}[A + B, [A, B]] - [[A, B], A + B] \\ &= \frac{1}{2}[A[A, B]] + \frac{1}{2}[[A, B], B] , \end{aligned} \quad (\text{D.27})$$

(since  $[B, [A, B]] = -[[A, B], B]$ ).

All further coefficients  $C_i$  obtained this way, can always be written in terms of multiple commutators containing preceding  $C_j$ ,  $j < i$ . Since these  $C_j$  only consist of multiple commutators, so do  $C_i$ , by induction.

Substituting Eqs. (D.24), (D.25), (D.26), and (D.27) in the series expansion (D.23), yields, for  $x = 1$ :

$$e^A e^B = e^C \quad (\text{D.28})$$

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] + \frac{1}{12}[[A, B], B] + \dots \quad (\text{D.29})$$

It is a nice challenge to further streamline the method described here for deriving the series of commutators for  $C$ .

## Appendix E

### Complex inner product, unitary and hermitian matrices

Let an  $n$ -dimensional linear space be given with a basis  $e_1, \dots, e_n$ , which is orthogonal when regarding some well defined inner product. Every vector in this space can be written as a linear combination of the  $e_1, \dots, e_n$ . In general, the coefficients are complex numbers.

Consider now two vectors  $\alpha$  and  $\beta$ ,

$$\begin{aligned} \alpha &= \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n, \\ \beta &= \beta_1 e_1 + \beta_2 e_2 + \dots + \beta_n e_n. \end{aligned}$$

The coefficients  $\alpha_i$  and  $\beta_i$  are just numbers. The inner product  $(\alpha, \beta)$  is a complex number given by

$$(\alpha, \beta) = \alpha_1^* \beta_1 + \alpha_2^* \beta_2 + \dots + \alpha_n^* \beta_n. \quad (\text{E.1})$$

Notice that we have

$$(\beta, \alpha) = (\alpha, \beta)^* \quad (\text{E.2})$$

One has the following axioms for the inner product:

$$(\alpha, \alpha) \geq 0, \quad \text{the equal sign only if } \alpha = 0, \quad (\text{E.3})$$

$$(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma), \quad (\text{E.4})$$

$$(\alpha, x\beta) = x(\alpha, \beta), \quad x = \text{number}, \quad (\text{E.5})$$

$$(x\alpha, \beta) = x^*(\alpha, \beta). \quad (\text{E.6})$$

Here,  $\beta + \gamma$  is a vector with components  $\beta_i + \gamma_i$ , and  $x\beta$  is a vector with components  $x\beta_i$ .

One can turn this argument around: if the above properties hold, then Eq. (E.1) implies the orthogonality of the basis vectors:

$$(e_i, e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} . \quad (\text{E.7})$$

Let now be given a matrix  $A$ . this matrix defines a mapping of the space into itself:

$$\alpha' = A\alpha , \quad (\text{E.8})$$

which is symbolic for

$$\alpha'_i = \sum_k a_{ik} \alpha_k . \quad (\text{E.9})$$

So, for every vector  $\alpha$ , the matrix  $A$  defines an image  $\alpha'$ . Which are the conditions to be imposed on  $A$  such that the inner product is invariant, in the sense that  $(\alpha, \beta) = (\alpha', \beta')$  for any pair  $\alpha, \beta$  ?

By expansion in components, one easily derives:

$$\begin{aligned} (\alpha', \beta') &= \sum_k \alpha'_k{}^* \beta'_k = \sum_{klm} (a_{kl} \alpha_l)^* a_{km} \beta_m \\ &= \sum_{k,l,m} \alpha_l^* a_{kl}^* a_{km} \beta_m . \end{aligned}$$

If now  $\tilde{A}$  is the transpose of  $A$  then

$$(\tilde{a})_{lk} = a_{kl} \quad (\text{E.10})$$

and if  $A^\dagger$  is the conjugate transpose of  $A$ :

$$(a^\dagger)_{lk} = a_{kl}^* \quad \text{and if } B = A^\dagger A , \quad (\text{E.11})$$

then evidently

$$\begin{aligned} b_{lm} &= \sum_k a_{lk}^\dagger a_{km} \\ &= \sum_k a_{kl}^* a_{km} \end{aligned}$$

So:

$$(\alpha', \beta') = \sum_{\ell,m} \alpha_\ell^* b_{\ell m} \beta_m \quad (\text{E.12})$$

with  $\beta = A^\dagger A$ . Let now  $\beta''$  be the vector obtained from  $\beta$  by applying  $B$  to it:

$$\beta'' = B\beta \quad (\text{E.13})$$

that is, according to Eq. (E.9)

$$\beta'_i = \sum_m b_{im} \beta_m \quad (\text{E.14})$$

and Eq. (E.12) is simply the inner product  $(\alpha, \beta'')$ , so  $(\alpha', \beta') = (\alpha, \beta'')$ . This must be equal to  $(\alpha, \beta)$  for all choices of  $\alpha$  and  $\beta$ . This can only be the case if  $\beta'' = \beta$  for all  $\beta$ , and therefore,  $B$  must be the unit matrix.

we conclude that the inner product is invariant under the mapping with a matrix  $A$  provided one has

$$A^\dagger A = \mathbf{1} . \quad (\text{E.15})$$

or equivalently,  $A^\dagger = A^{-1}$ . A matrix obeying this condition is called *unitary*. In real spaces, one has  $A^\dagger = \tilde{A}$ , and so  $A^{-1} = \tilde{A}$ . Such a matrix is *orthogonal*.

A brief symbolic display of the above derivation reads

$$(\alpha', \beta') = (A\alpha, A\beta) = (\alpha, A^\dagger A\beta) . \quad (\text{E.16})$$

Note that for inner products, we have

$$(A\alpha, \beta) = (\alpha, A^\dagger \beta) , \quad (\text{E.17})$$

which holds for any matrix  $A$ .

An eigenvector of a matrix  $A$  is a vector with the property

$$A\alpha = \lambda\alpha , \quad (\text{E.18})$$

where  $\lambda$  is a real or complex number; *i.e.*

$$\sum_k a_{ik} \alpha_k = \lambda \alpha_i . \quad (\text{E.19})$$

$\lambda$  is the eigenvalue of  $A$  associated to the eigenvector  $\alpha$ . Which are the conditions on  $A$  that guarantee all  $\lambda$  to be real?

Consider the inner product  $(\alpha, A\alpha) = \lambda(\alpha, \alpha)$ . Now,  $(\alpha, \alpha)$  is real and larger than zero if  $\alpha$  is not zero. Real  $\lambda$  apparently implies that  $(\alpha, A\alpha)$  is real. That means

$$(\alpha, A\alpha) = (\alpha, A\alpha)^* = (A\alpha, \alpha) = (\alpha, A^\dagger \alpha) . \quad (\text{E.20})$$

This clearly holds if  $A^\dagger = A$ . A matrix obeying  $A^\dagger = A$  is called hermitian. Therefore:

$$A = \text{hermitian, } i.e., A^\dagger = A \longrightarrow \text{only real eigenvalues.} \quad (\text{E.21})$$

The converse is not necessarily true: there are many non-hermitian matrices with only real eigenvalues.