Symmetries in Quantum Mechanics

Lecture by Frank Krauss

TU–Dresden, 2000-2001

Contents

1 Introduction : The role of symmetries

Symmetries play a fundamental role in modern physics. Their significance manifests itself in a twofold way :

1. In general, we will like to construct Lagrange or Hamilton functions of systems under investigation. Essential input for this is the form of the system, the form of the interactions and the like. For more transparent treatment, we will look for more general structures, invariances and symmetries. a good example for this strategy can be found in particle physics.

   There, symmetries in the form of *Gauge Symmetries* mark a central ingredient in any current attempt to construct a theory on the level of elementary particles and their interactions. In addition, *Lorentz–invariance* defines the structure of space–time, and some discrete symmetries like invariance under the combined application of C (charge–conjugation), P (parity,i.e. the mirroring from "left–handed" to "right–handed" coordinate systems) and T (time reversal, $t \to -t$) pose severe constraints on the form of interactions. In this framework, it might be interesting to mention, that there is another symmetry called *Supersymmetry*, that connects the particles of different statistics, namely bosons (e.g. photons) with fermions (electrons, for example, obeying some exclusion principle). This additional symmetry is the best candidate these days for some new physics on the level of elementary particles. In fact it is widely employed to overcome some conceptual shortcomings of the so–called Standard Model.

2. The symmetry properties of systems under investigation give clear hindsights on their dynamics. Thus some of the basic techniques in the construction of solutions for a large class of problems rely on the investigation of symmetries.

Especially the second point will be in the focus of this lecture. But before we turn our attention to the formulation of symmetry properties in terms of groups and to quantum mechanics, we will highlight the role of symmetries in classical physics.

1.1 Some elementary examples

The most elegant way to see the impact of symmetries in classical mechanics is via *Noethers theorem*, formalizing the connection between symmetries and conservation laws. However, before properly phrasing and proving it, let us start with some more elementary examples.

- Homogeneity in space :
  The assumption of a homogenous space is reflected by the statement, that the Lagrange function $\mathcal{L}$ is invariant under translations $\vec{r} \to \vec{r}' = \vec{r} + \delta\vec{r}$,

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\vec{r}_i}\delta\vec{r} = 0\,. \tag{1}$$

The relation above means, that *any* coordinate vector $\vec{r}_i$ is shifted by some common $\delta\vec{r}$, i.e. the whole system is displaced. The Euler–Lagrange Equations of Motion (E.o.M.) are

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i} - \frac{\partial\mathcal{L}}{\partial q_i} = 0\,, \tag{2}$$

and taking the ordinary space coordinates r as generalized coordinates yields q for the x–component

$$\sum_i \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{x}_i} \equiv \frac{dP_x}{dt} = 0 \tag{3}$$

or $P_x = \text{const.}$. Thus, homogeneity in space along any direction x enforces the conservation of the total momentum along the same axis.

Similarly, if the potential $\mathcal{V}$ within the system depends only on the spatial distances of any two particles i and j, $\mathcal{V} = \mathcal{V}(|\vec{r}_i - \vec{r}_j|)$, then the overall momentum $\vec{P}$ is a conserved quantity as well.

Note, that this conservation law enables us to define a center of mass frame and the analysis of our system in terms of relative coordinates only.

- Isotropy in space :
  Isotropy in space means, that there is no preferred spatial direction, i.e. that systems under these conditions do not behave differently when rotated as a whole. More formally, the Lagrange function is invariant under rotations. Let us consider now infinitesimal rotations $\delta\vec{\phi} = (\delta\phi_x, \delta\phi_y, \delta\phi_z)^T$. Applying such a rotation on some vector $\vec{r}$, we find, that this vector is displaced by

$$\delta\vec{r} = \delta\vec{\phi}\times\vec{r}\,. \tag{4}$$

Velocities $\vec{v}$ are rotated identically,

$$\delta\vec{v} = \delta\vec{\phi}\times\vec{v}\,. \tag{5}$$

and we find for the change of the Lagrange function under infinitesimal rotations

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\vec{r}_i}\delta\vec{r}_i + \frac{\partial\mathcal{L}}{\partial\vec{v}_i}\delta\vec{v}_i \equiv 0\,, \tag{6}$$

where summation over identical indices is understood (Einstein convention). However, the canonical momenta are defined as

$$\vec{\pi}_i = \frac{\partial\mathcal{L}}{\partial\vec{v}_i} \tag{7}$$

and, according to the E.o.M. they fulfill

$$\dot{\vec{\pi}}_i = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\vec{v}_i} = \frac{\partial\mathcal{L}}{\partial\vec{r}_i}\,. \tag{8}$$

Plugging this into Eq.(6)and employing $\vec{a}\cdot(\vec{b}\times\vec{c}) = \vec{b}\cdot(\vec{c}\times\vec{a})$, we find

$$\begin{aligned} 0 &\equiv \dot{\vec{\pi}}_i \cdot \delta\vec{r}_i + \vec{\pi}_i \cdot \delta\vec{v}_i \\ &= \delta\vec{\phi}\cdot\left(\vec{r}_i \times \dot{\vec{\pi}}_i + \dot{\vec{r}}_i \times \vec{\pi}_i\right) = \delta\vec{\phi}\cdot\frac{d}{dt}\left(\vec{r}_i \times \vec{\pi}_i\right)\,, \end{aligned} \tag{9}$$

or, in other words, that the classical total angular momentum $\vec{L} = \vec{r}_i \times \vec{\pi}_i$ is a conserved quantity,

$$\frac{d\vec{L}}{dt} = \frac{d}{dt}(\vec{r}_i \times \vec{\pi}_i) = 0\,. \tag{10}$$

- Homogeneity in time :
  Homogeneity with respect to time translates into an invariance of the Equations of Motion under the time translation $t \to t' = t + \delta t$. This is nothing else, but the statement, that there is no explicit dependence of the Lagrange–function on the time, i.e. $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i)$. From this

$$\frac{d\mathcal{L}}{dt} = \frac{\partial\mathcal{L}}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{d\dot{q}_i}{dt} \tag{11}$$

results, and applying the E.o.M. yields

$$\frac{d\mathcal{L}}{dt} = \frac{d}{dt}\left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) \tag{12}$$

or

$$0 = \frac{d}{dt}\left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L}\right) \equiv \frac{d\mathcal{H}}{dt}\,. \tag{13}$$

Identifying the Hamilton–function $\mathcal{H}$ with the energy of the system connects invariance under time translations with energy conservation.

1.2 Noethers theorem

After contemplating these "classical examples" of symmetries in classical mechanics, we're in the position to turn our attention now to the more formal statement of Noethers theorem :

The invariance of the Euler–Lagrange Equations of Motion under a coordinate transformation $\{t, q_i\} \to \{t'(t), q'_i(q_j, t)\}$, implies the existence of a conserved quantity.

We are now going to prove this theorem by constructing the conserved quantity explicitly. In the course of this proof we will first constrain the form of transformations of the Lagrange function under coordinate transformations above, which keep the action and the E.o.M. invariant. Then we deduce the form of an infinitesimal transformation of the Lagrange function which allows us to construct the conserved quantity.
Reconsider the coordinate transformation above,

$$\{t, q_i\} \to \{t' = t'(t) = t + \delta t(t), q'_i = q'_i(q_j, t) = q_i + \delta q_i(q_j, t)\}\,, \tag{14}$$

linking the primed and the unprimed systems with – accordingly primed or unprimed – Lagrange functions $\mathcal{L}'$ and $\mathcal{L}$.
The primed velocities are then given by

$$\dot{q}'_i = \frac{dq'_i}{dt'} = \frac{dq'_i}{dt}\frac{dt}{dt'} = \left(\dot{q}_i + \frac{d}{dt}\delta q_i\right)\frac{1}{1 + \frac{d}{dt}\delta t}\,. \tag{15}$$

Concentrating on infinitesimal and invertible transformations only (this is in the spirit of our discussion), this allows us to rewrite

$$\delta\dot{q}_i = \dot{q}'_i - \dot{q}_i = \left(\dot{q}_i + \frac{d}{dt}\delta q_i\right)\left(1 - \frac{d}{dt}\delta t\right) - \dot{q}_i = \frac{d}{dt}\delta q_i - \dot{q}_i\frac{d}{dt}\delta t\,. \tag{16}$$

Note, that here and in the following we ignore terms of the order $\mathcal{O}(\delta^2)$.
Of course, the action should remain invariant under the coordinate transformation above, i.e.

$$S(t_1,t_2) = \int\limits_{t_1}^{t_2} dt \mathcal{L}(q_i(t),\dot{q}_i(t),t) = \int\limits_{t'_1(t_1)}^{t'_2(t_2)} dt' \mathcal{L}'(q'_i(t'),\dot{q}'_i(t'),t') = S'(t_1,t_2)\,. \quad (17)$$

Evidently, this is fulfilled, if the Lagrange function transforms like

$$\mathcal{L}'(q'_i,\dot{q}'_i,t') = \mathcal{L}(q_i(q'_j,t),\dot{q}_i(q_j,\dot{q}_j,t),t(t'))\frac{dt}{dt'}\,. \quad (18)$$

If the resulting E.o.M. remain form-invariant, the coordinate transformation is dubbed *Symmetry Transformation*.
However, the condition above is fulfilled, if the two Lagrange functions $\mathcal{L}$ and $\mathcal{L}'$ differ by a total derivative in time only,

$$\mathcal{L}'(q'_i,\dot{q}'_i,t') = \mathcal{L}(q'_i,\dot{q}'_i,t') + \frac{d}{dt'}\Lambda(q'_i,t')\,. \quad (19)$$

After some rewriting of differentials, we find, that $d\Lambda/dt'$ indeed fulfills the E.o.M.,

$$\begin{aligned}
&\frac{d}{dt'}\frac{\partial}{\partial\dot{q}'_i}\frac{d\Lambda}{dt'} - \frac{\partial}{\partial q'_i}\frac{d\Lambda}{dt'} \\
&= \frac{d}{dt'}\frac{\partial}{\partial\dot{q}'_i}\left(\frac{\partial\Lambda}{\partial q'_j}\dot{q}'_j + \frac{\partial\Lambda}{\partial t'}\right) - \frac{\partial}{\partial q'_i}\left(\frac{\partial\Lambda}{\partial q'_j}\dot{q}'_j + \frac{\partial\Lambda}{\partial t'}\right) \\
&= \frac{d}{dt'}\frac{\partial\Lambda}{\partial q'_i} - \left(\frac{\partial^2\Lambda}{\partial q'_i\partial q'_j}\dot{q}'_j + \frac{\partial^2\Lambda}{\partial t'\partial q'_i}\right) \\
&= \frac{\partial^2\Lambda}{\partial q'_i\partial q'_j}\dot{q}'_j + \frac{\partial^2\Lambda}{\partial q'_i\partial t'} - \left(\frac{\partial^2\Lambda}{\partial q'_i\partial q'_j}\dot{q}'_j + \frac{\partial^2\Lambda}{\partial q'_i\partial t'}\right) \equiv 0\,. \quad (20)
\end{aligned}$$

In the course of the manipulations above, we have used, that $d\Lambda/dt'$ does not depend on $\dot{q}'_i$.
We can now conclude, that $\mathcal{L}$ and $\mathcal{L}'$ lead to the very same E.o.M.. Plugging Eq.(18) into Eq.(19) we find after returning to the old coordinates

$$\mathcal{L}(q_i,\dot{q}_i,t) \;=\; \mathcal{L}(q'_i(q_j,t),\dot{q}'_i(q_j,\dot{q}_j,t),t'(t))\frac{dt'}{dt} + \frac{d}{dt}\Lambda(q'_i(q_j,t),t'(t))\,. \quad (21)$$

Using the definition of the coordinate transformations above, we can prepare for the treatment of infinitesimal transformations via

$$
\begin{aligned}
&\mathcal{L}(q_i,\dot q_i,t)-\mathcal{L}(q_i'(q_j,t),\dot q_i'(q_j,\dot q_j,t),t'(t)) \\
&= \mathcal{L}(q_i'(q_j,t),\dot q_i'(q_j,\dot q_j,t),t'(t))\frac{d\delta t}{dt}+\frac{d}{dt}\Lambda(q_i'(q_j,t),t'(t))\,, \qquad (22)
\end{aligned}
$$

where we have employed again, that $dt'/dt = 1 + d\delta t/dt$. The Equation above can now be rewritten as

$$
\begin{aligned}
-\delta\mathcal{L} &= \mathcal{L}(q_i,\dot q_i,t)-\mathcal{L}(q_i+\delta q_i,\dot q_i+\delta\dot q_i,t+\delta t) \\
&= \mathcal{L}(q_i,\dot q_i,t)\frac{d\delta t}{dt}+\frac{d}{dt}\Lambda(q_i+\delta q_i,t+\delta t)+\mathcal{O}(\delta^2)\,. \qquad (23)
\end{aligned}
$$

Choosing $\delta q_i = \delta t = 0$ we find immediately, that $d\Lambda/dt$ vanishes trivially with $\delta\mathcal{L}$, see for instance Eq.21, since $q' = q$, $\dot q' = \dot q$, and $t' = t$. This reasoning leaves us with

$$
\begin{aligned}
-\delta\mathcal{L} &= -\frac{\partial\mathcal{L}}{\partial q_i}\delta q_i-\frac{\partial\mathcal{L}}{\partial\dot q_i}\delta\dot q_i-\frac{\partial\mathcal{L}}{\partial t}\delta t \\
&= \mathcal{L}\frac{d\delta t}{dt}+\frac{d}{dt}\delta\Lambda\,. \qquad (24)
\end{aligned}
$$

Using the expression for $\delta\dot q_i$, Eq.(16) we have

$$
\left(\frac{\partial\mathcal{L}}{\partial q_i}+\frac{\partial\mathcal{L}}{\partial\dot q_i}\frac{d}{dt}\right)\delta q_i+\frac{\partial\mathcal{L}}{\partial t}\delta t+\left(\mathcal{L}-\frac{\partial\mathcal{L}}{\partial\dot q_i}\dot q_i\right)\frac{d\delta t}{dt}=-\frac{d\delta\Lambda}{dt} \qquad (25)
$$

Finally the E.o.M. can be employed once more, and we arrive after some more analysis at

$$
\frac{d}{dt}\left[\frac{\partial\mathcal{L}}{\partial\dot q_i}\delta q_i+\left(\mathcal{L}-\frac{\partial\mathcal{L}}{\partial\dot q_i}\dot q_i\right)\delta t+\delta\Lambda\right]=0\,. \qquad (26)
$$

We see, that the expression in the square brackets is a conserved quantity, i.e. an integral of the E.o.M..
This result can be checked by rewriting the left hand side of this equation as

$$
\begin{aligned}
l.h.s. &= \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot q_i}\delta q_i+\frac{\partial\mathcal{L}}{\partial q_i}\frac{d\delta q_i}{dt}+\frac{\partial\mathcal{L}}{\partial q_i}\dot q_i\delta t+\frac{\partial\mathcal{L}}{\partial\dot q_i}\ddot q_i\delta t+\frac{\partial\mathcal{L}}{\partial t}\delta t \\
&\quad -\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot q_i}\dot q_i\delta t-\frac{\partial\mathcal{L}}{\partial\dot q_i}\ddot q_i\delta t+\left(\mathcal{L}-\frac{\partial\mathcal{L}}{\partial\dot q_i}\dot q_i\right)\frac{d\delta t}{dt}+\frac{d\delta\Lambda}{dt} \\
&= \frac{\partial\mathcal{L}}{\partial q_i}\delta q_i+\frac{\partial\mathcal{L}}{\partial q_i}\frac{d\delta q_i}{dt}+\frac{\partial\mathcal{L}}{\partial t}\delta t+\left(\mathcal{L}-\frac{\partial\mathcal{L}}{\partial\dot q_i}\dot q_i\right)\frac{d\delta t}{dt}+\frac{d\delta\Lambda}{dt} \qquad (27)
\end{aligned}
$$

and we recover Eq.(25) again.

We will highlight the impact of Eqs.(25) and (26) with two more examples.

1. Invariance under Galilei transformations :
   The Galilei transformations are

$$\vec{r} \to \vec{r}' = \vec{r} + \vec{v}_0 t \ , \ \ \vec{v} \to \vec{v}' = \vec{v} + \vec{v}_0 \ , \ \ t \to t' = t \, . \tag{28}$$

   For the sake of some simplification we choose now $\vec{v}_0 = (0, 0, v_3)^T$. Since $\delta t = 0$, Eq.(25) is simplified to

$$\left(\frac{\partial \mathcal{L}}{\partial x_3} + \frac{\partial \mathcal{L}}{\partial \dot{x}_3} \frac{d}{dt} \right) v_3 t = - \frac{d\Lambda}{dt} \tag{29}$$

   and, furthermore,

$$\left(- \frac{\partial \mathcal{V}}{\partial x_3} t + m \dot{x}_3 \right) v_3 = - \frac{\partial \Lambda}{\partial x_i} \delta \dot{x}_i - \frac{\partial \Lambda}{\partial t} \, , \tag{30}$$

   where $\mathcal{V}$ is the (only space–dependent) potential. Comparing coefficients shows, that with some function f

$$\delta \Lambda = - m x_3 v_3 + f(t) \text{ and } \ \frac{df(t)}{dt} = \frac{\partial \mathcal{V}}{\partial x_3} v_3 t \, , \tag{31}$$

   ergo $\partial \mathcal{V} / \partial x_3$ =const., and

$$\delta \Lambda = \left(- m x_3 + \frac{1}{2} \frac{\partial \mathcal{V}}{\partial x_3} t^2 \right) v_3 \, . \tag{32}$$

   Adding invariance under translations to our cocktail, i.e. $\partial \mathcal{V} / \partial x_3 = 0$, then we have, returning to vectors

$$\vec{x} - \frac{\vec{p}}{m} t = \text{const.} \tag{33}$$

   as conserved quantity, i.e. the system moves with constant speed.

2. Particle in a homogenous electric field :
   Let's start with the Lagrange function for a particle in an electromagnetic field as described by the scalar potential Φ and the vector potential A, (note, that we use *natural units*, i.e. $\hbar = c = 1$.)

$$L = \frac{m\dot{r}^2}{2} + q\vec{A}\dot{\vec{r}} - q\Phi \,. \tag{34}$$

The canonical momentum is in this case

$$\vec{p} = m\dot{\vec{r}} + q\vec{A} \tag{35}$$

and the Lorentz force reads

$$\frac{\partial\mathcal{L}}{\partial\vec{r}} = q\frac{\partial\vec{A}\dot{\vec{r}}}{\partial\vec{r}} - q\frac{\partial\phi}{\vec{r}} \,. \tag{36}$$

However, we want to check now, whether the canonical momentum is conserved in different gauges of the electromagnetic field. For this purpose we consider two frames leading to the very same field configuration, namely

$$\begin{array}{lllcllll} \Phi & = & -\vec{E}\vec{r} & , & \vec{A} & = & \vec{0} \\ \Phi' & = & 0 & , & \vec{A}' & = & -\vec{E}t \end{array}$$

and find as E.o.M.

$$\frac{d}{dt}\vec{p} - q\vec{E} = \frac{d}{dt}(\vec{p} - q\vec{E}t) = 0 \tag{37}$$

and

$$\frac{d}{dt}\vec{p}' = 0 \,, \tag{38}$$

respectively. This contradiction seems at a first glance to indicate some major disaster, namely the dependence of conserved quantities on the – in principle unobservable – gauge of an electromagnetic field. However, closer inspection shows, that the canonical momentum $\vec{p}$ and $\vec{p}'$ look differently, namely

$$\vec{p} = m\dot{\vec{r}} \text{ and } \vec{p}' = m\dot{\vec{r}} - q\vec{E}t \tag{39}$$

and we recover the expected conservation of the same quantity in the two frames. Or stated in other words, the definition of the canonical momentum is gauge–variant.

♣ In a magnetic field, similar reasoning holds true.

2 Finite Groups and their Representations

2.1 Finite Groups

2.1.1 Definition of a group

A group consists of a set $\mathcal{G}$, on which an operation "$\cdot$" ("multiplication") is defined. This operation exhibits the following properties

1. *Closure* : If g_i and g_j are elements of $\mathcal{G}$, then also $g_k = g_i \cdot g_j$ is a member of $\mathcal{G}$.

2. *Identity Element* (Existence of a "1") : There is exactly one identity element g_1 in $\mathcal{G}$, that maps all elements g_i exactly on themselves, $g_i = g_i \cdot g_1 = g_1 \cdot g_i$.

3. *Inverse Element* : For every element g_i of $\mathcal{G}$ there is exactly one inverse element g_i^{-1} with the property $g_i^{-1} \cdot g_i = g_i \cdot g_i^{-1} = g_1$.

4. *Associativity* : $g_i \cdot (g_j \cdot g_k) = (g_i \cdot g_j) \cdot g_k$ is fulfilled.

The property of commutativity, i.e. the validity of $g_i \cdot g_j \stackrel{?}{=} g_j \cdot g_i$, decides, whether a group as defined above is *Abelian* or *non–Abelian*.
The number of elements of a group is called its *order*.
A group with a finite number of elements is called – surprise, surprise – finite, infinite groups may be either discrete or continuous, depending on whether the elements of the group are denumerably or not. Examples for the latter two cases are the set of integers versus the set of real numbers.
Some examples of groups include :

- $\{1, -1, i, -i\}$ together with the ordinary multiplication (note here, that the subset $\{-1, 1\}$ with the multiplication forms a *subgroup*),
- The integer numbers with the addition as operation,
- The set $\{1, 2, \ldots, p-1\}$ forms a group under multiplication modulo p, if p is a prime number.
- the set of all permutations of three objects a, b, and c, sitting on positions 1, 2, and 3. The group of permutations contains the six operations

( )	do nothing, $(a,b,c) \to (a,b,c)$
(12)	exchange objects on positions 1 and 2, $(a,b,c) \to (b,a,c)$
(13)	exchange objects on positions 1 and 3, $(a,b,c) \to (c,b,a)$
(23)	exchange objects on positions 2 and 3, $(a,b,c) \to (a,c,b)$
(123)	cyclic permutation, and $(a,b,c) \to (c,a,b)$
(321)	anti-cyclic permutation $(a,b,c) \to (b,c,a)$.

The group properties of this example are easily checked. Closure can be tested directly by noting, that each sequence of two permutations can be understood as application of one permutation only. For instance, $(23) \cdot (12) = (321)$. The identity element is ( ), each element (ij) is its own inverse, and the two elements (ijk) applied consecutively yield ( ). To prove associativity is should be sufficient to notice, that after each step of $x \cdot y \cdot z$ an unambiguous "physical" state is reached, characterized by the sequence of the object on their positions.

- The group D_4 of symmetry transformations on a square with corners labeled clockwise from 1 to 4. This group consists of the following elements :

E	Identity element.
$\mathcal{R}_{90,180,270}$	Clockwise rotations around the center of the square, with angle as indicated.
$\mathcal{M}_{x,y}$	Reflections around the horizontal (x) and vertical (y) axis.
$\mathcal{M}_{13,24}$	Reflections around the axis (13) or (24), respectively.

Obviously, the "multiplication" of two such operations can be defined as their sequential application. Here and in the following it is understood, that in the product $A \cdot B$ first B and then A is executed, i.e. the order of operation is from the right to the left. This anticipates already the effect of such operations on state vectors, "kets", $|\Psi\rangle$.

2.1.2 Multiplication tables

Before proceeding further let us investigate in a little bit more depth the effect of the sequential application of the permutation operations of the example above. It proves useful to organize the products of two permutations in a *multiplication table*.

2nd op.	1st op					
	( )	(12)	(13)	(23)	(123)	(321)
( )	( )	(12)	(13)	(23)	(123)	(321)
(12)	(12)	( )	(321)	(123)	(23)	(13)
(13)	(13)	(123)	( )	(321)	(12)	(23)
(23)	(23)	(321)	(123)	( )	(13)	(12)
(321)	(321)	(23)	(12)	(13)	( )	(123)
(123)	(123)	(13)	(23)	(12)	(321)	( )

Note, that we have arranged the rows and columns in two different ways in order to have the neutral elements in the principal diagonal. The so–called *rearrangement theorem* states, that in such a multiplication table, each group element appears once and only once in each row or each column of the results.

To prove it, consider two *distinct* elements C and B, such that $B \cdot A = C \cdot A = D$ and multiply from the right with A^{-1}. A consequence of this theorem is, that for any function f of the group elements

$$\sum_{g_i \in \mathcal{G}} f(g_i) = \sum_{g_i \in \mathcal{G}} f(h \cdot g_i) \tag{40}$$

for finite groups. Here, h denotes an arbitrary fixed element of $\mathcal{G}$
Somewhat exaggerated we can now rephrase our definition of a group :
A group is nothing else but a multiplication table with the properties (1) to (4) of our definition above.
♣ Find the multiplication table for D_4.

2.1.3 Generators of finite groups

In general it is possible to generate all the elements of a *finite group* by starting from a certain limited set of elements subject to some relations. Consider now the smallest set of elements generating all other elements by repeated application of the operation belonging to the group. The elements of this minimal set are called the *generators* of the group. Note, that this set is not necessarily unique.
♣ Coming back to our standard example of the permutation group over three elements, it is quite easy to check, that for instance any two out of the three (ij) generate the whole group. We want to leave this as a problem for the interested reader.

2.1.4 Conjugate elements, similarity transformations, and classes

Two elements h_1 and h_2 of a group $\mathcal{G}$ are dubbed *conjugate elements*, if there is some element g with

$$g^{-1} h_1 g = h_2 \, . \tag{41}$$

The operation above is then called a *similarity transformation*. We can now split our group $\mathcal{G}$ into sets, so that each two elements of the same set are conjugate to each other, two elements of distinct sets are not conjugate. These sets are called *classes*.
As an example for the classification of group elements into classes, consider the group of symmetry transformations of a square. Here, we find the following class structure :

$$\{E\} \, , \quad \{\mathcal{R}_{90} \, , \, \mathcal{R}_{270}\} \, , \quad \{\mathcal{R}_{180}\} \, , \quad \{\mathcal{M}_x \, , \, \mathcal{M}_y\} \, , \quad \{\mathcal{M}_{13} \, , \, \mathcal{M}_{24}\} \, . \tag{42}$$

There are a few lections, that can be learned out of this.

1. The identity element always constitutes a class of its own.

2. Rotations with different magnitudes of angles belong to different classes. Rotations with a different sense, i.e. clockwise vs. anticlockwise, belong to the same class only, if there's a suitable reflection contained in the group, that takes a right–handed to a left–handed system. In our example, for instance $\mathcal{M}_x$ plays this role.

3. Rotations with different axes or reflections in different planes belong to the same class only, if the axes or planes can be mapped onto each other by some group element. In our examplatory group, this is the case with $\mathcal{M}_x$ and $\mathcal{M}_y$, but not with $\mathcal{M}_x$ and $\mathcal{M}_{13}$.

Two classes $\mathcal{C}_n$ and $\mathcal{C}_m$ with n and m elements, respectively, of a group $\mathcal{G}$ can be multiplied via

$$\mathcal{C}_n \cdot \mathcal{C}_m = \{\mathcal{C}_n^i \cdot \mathcal{C}_m^j\}\,, \tag{43}$$

where the $\mathcal{C}_n^i$, $\mathcal{C}_m^j$ are the elements of the corresponding class. It can be shown now, that this new set consists of classes by applying ($g \in \mathcal{G}$)

$$g^{-1}\mathcal{C}_n^i \cdot \mathcal{C}_m^j g = g^{-1}\mathcal{C}_n^i g \cdot g^{-1}\mathcal{C}_m^j g = \mathcal{C}_n^{i'} \cdot \mathcal{C}_m^{j'}\,. \tag{44}$$

2.1.5 Subgroups

A set $\mathcal{H}$ is said to be a *subgroup* of $\mathcal{G}$, if all members of $\mathcal{H}$ are elements of $\mathcal{G}$, and if $\mathcal{H}$ forms a group under the same operation as $\mathcal{G}$. As an example, consider the set $\{E\,,\ \mathcal{R}_{90}\,,\ \mathcal{R}_{180}\,,\ \mathcal{R}_{270}\}$ of transformations on our square. The two trivial subgroups of every group are the identity element and the group itself. Note, that elements of a class in a group do not necessarily belong to the same class in a subgroup.
If a group $\mathcal{H}$ of order h is a subgroup of $\mathcal{G}$ with order g, then g is an integral multiple of h. To prove this multiply any element of $\mathcal{H}$ with some element $h^{(1)}$ not contained in $\mathcal{H}$. The resulting set $h^{(1)} \cdot \mathcal{H}$ does not belong to $\mathcal{H}$, and we have obviously found h new elements of $\mathcal{G}$. Now repeat this procedure with another element $h^{(2)}$ not contained in $\mathcal{H}$ or $h^{(1)} \cdot \mathcal{H}$. The two sets $h^{(1)} \cdot \mathcal{H}$ and $h^{(2)} \cdot \mathcal{H}$ are disjoint. To see this, consider

$$h^{(1)} \cdot h_i = h^{(2)} \cdot h_j \implies h^{(1)} = h^{(2)} \cdot h_j \cdot h_i^{-1} = h^{(2)} \cdot h_k\;. \tag{45}$$

Evidently, this produces a contradiction with our assumptions above.

2.1.6 Direct Products

The *direct product* of two groups $\mathcal{H}_1$ and $\mathcal{H}_2$ of order h_1 and h_2, respectively, is defined as

$$\mathcal{G} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \{h_1^i \cdot h_2^j\} \ . \tag{46}$$

This definition holds true only, if both groups have only the identity element in common and if the elements of both groups commute mutually. As an example, consider

$$\{E\,,\ \mathcal{M}_x\} \otimes \{E\,,\ \mathcal{M}_y\} = \{E\,,\ \mathcal{M}_x\,,\ \mathcal{M}_y\,,\ \mathcal{R}_{180}\}\,. \tag{47}$$

Note, that taking direct products is the least involved way to enlarge groups. Consequently, it has been used in abundance. As a “classical example”, take the gauge group of the Standard Model of particle physics given by $SU(3) \otimes SU(2) \otimes U(1)$ for the strong, weak and electromagnetic interactions.

2.2 Representations of finite groups

2.2.1 Definitions

A mapping $\mathcal{D}$ of the group $\mathcal{G}$ with the property

$$\mathcal{D}(g_i) \cdot \mathcal{D}(g_j) = \mathcal{D}(g_i \cdot g_j) \tag{48}$$

for all elements g_i and g_j of $\mathcal{G}$ is called a *representation* of the group $\mathcal{G}$. More particularly, the name *representation* is reserved for a collection of square matrices T with the property above. The *dimension of the representation* is identified with the rank of the matrices.

So, returning once more to one of our favorite examples, we now give a representation of the permutation group of three elements in terms of square 3×3 matrices,

$$\begin{aligned}
\mathcal{D}(\ ) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \mathcal{D}(12) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\mathcal{D}(13) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \mathcal{D}(23) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\mathcal{D}(321) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} & \mathcal{D}(123) &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} .
\end{aligned}$$

In this particular representation, the operation "$\cdot$" can be identified with the multiplication of matrices, for instance $(12)\cdot(23) = (123)$ or

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} . \tag{49}$$

As a matter of fact, we see again, that this group is non–Abelian.
To construct a representation, suppose there is an n–dimensional vector–space with orthonormal basis $|\phi_i\rangle$, on which the group elements g_k act. Using the completeness relation

$$\mathbf{1} = \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i| \tag{50}$$

we can write the effect of any g_k on a basis vector as

$$g_k|\phi_i\rangle = \sum_{j=1}^{n} |\phi_j\rangle\langle\phi_j|T(g_k)|\phi_i\rangle \tag{51}$$

and simple matrix algebra yields

$$\sum_{j=1}^{n}\langle\phi_k|T(g_1)|\phi_j\rangle\langle\phi_j|T(g_2)|\phi_i\rangle = \langle\phi_k|T(g_1\cdot g_2)|\phi_i\rangle \; . \tag{52}$$

♣ We want to leave the proof of the equation above for the interested reader.

Since the matrices T do not need to be distinct, the set of the T does not necessarily form a group under matrix multiplication.
A very trivial case of a representation is the so called identity representation, in which all $T \equiv 1$. Therefore, whenever we refer to a "group" T in the following, we mean a set of distinct matrices forming a group under multiplication.

2.2.2 Equivalent representations

Two representations T_1 and T_2 of a group are called equivalent, if there is a nonsingular matrix S with

$$T_1 = S^{-1}T_2S \, . \tag{53}$$

In other words if the two representations can be mapped onto each other by a similarity transformation within the same vector space, in which both

representations (matrices) are defined and act, then the two representations are equivalent.
A very useful theorem on representations is, that each representation T of a finite group is equivalent to a representation by unitary matrices. To prove this, we define a hermitian matrix

$$H = \sum_{g_i \in G} T(g_i) T^\dagger(g_i) \tag{54}$$

and use the fact, that each hermitian matrix can be diagonalized by a unitary matrix,

$$U^{-1} H U = H_{\rm diag} \,. \tag{55}$$

The diagonal elements of $H_{\rm diag}$ are then the real eigenvalues of H.
Using the definition of H, we see, that

$$\begin{aligned} H_{\rm diag} &= \sum_{g_i \in G} U^{-1} T(g_i) U U^{-1} T^\dagger(g_i) U \\ &= \sum_{g_i \in G} T'(g_i) T'^\dagger(g_i) \,, \end{aligned} \tag{56}$$

and the kth diagonal element of $H_{\rm diag}$ is

$$\begin{aligned} [H_{\rm diag}]_{kk} &= \sum_{g_i \in G} \sum_j T'_{kj}(g_i) {T'_{jk}}^\dagger(g_i) \\ &= \sum_{g_i \in G} \sum_j |T_{kj}(g_i)|^2 > 0 \end{aligned} \tag{57}$$

Therefore, $H_{\rm diag}$ is positive definite, and, consequently, non–singular. Thus, powers of $H_{\rm diag}$ are given by a matrix with diagonal elements, that are the powers of the diagonal elements of $H_{\rm diag}$,

$$H^n_{\rm diag} = {\rm diag}\left\{ \left[H^{ii}_{\rm diag} \right]^n \right\} \,. \tag{58}$$

Unitary matrices $\tilde{T}(g_i)$ are then constructed via

$$\tilde{T}(g_i) = H^{-1/2}_{\rm diag} U^{-1} T(g_i) U H^{1/2}_{\rm diag} \,. \tag{59}$$

For the check of unitarity of the $\tilde{T}$, consider

$$
\begin{aligned}
\tilde{T}(g_i)\tilde{T}^\dagger(g_i) &= H_{\rm diag}^{-1/2} T'(g_i) H_{\rm diag} T'^\dagger(g_i) H_{\rm diag}^{-1/2} \\
&= H_{\rm diag}^{-1/2} T'(g_i) \left(\sum_{g_j \in G} T'(g_j) T'^\dagger(g_j) \right) T'^\dagger(g_i) H_{\rm diag}^{-1/2} \\
&= H_{\rm diag}^{-1/2} \left(\sum_{g_j \in G} T'(g_i g_j) T'^\dagger(g_i g_j) \right) H_{\rm diag}^{-1/2} \\
&= H_{\rm diag}^{-1/2} H_{\rm diag} H_{\rm diag}^{-1/2} = \mathbf{1}\,. \qquad (60)
\end{aligned}
$$

Note, that the sequence of the group elements in the "daggered" matrix $T'^\dagger(g_i g_j)$ stems from

$$
(AB)^\dagger = B^\dagger A^\dagger\,. \qquad (61)
$$

Put in other words, Eq.(60) states, that we can go from non–orthonormal basis vectors related to the T to orthonormal ones, related to the representation $\tilde{T}$.

2.2.3 Invariant subspaces and reducible representations

Let us assume we have a vector space L_n which is closed under a group $\mathcal{G}$, i.e. no transformation of L_n by any of the group members yields a vector outside of L_n. Any proper subspace L_m of L_n (any vector in L_m is in L_n, but the vectors of L_m do not exhaust L_n), which is invariant under $\mathcal{G}$, i.e. mapped onto itself, is called *an invariant subspace of L_n under $\mathcal{G}$*, and, conversely, L_n is called *reducible* under $\mathcal{G}$.
In this case, the representation $T = T(g_i)$ has, in a suitable basis, the form

$$
T(g_i) = \begin{pmatrix} t^{(1)}(g_i) & \mathbf{0} \\ x(g_i) & t^{(2)}(g_i) \end{pmatrix}\,, \qquad (62)
$$

where $t^{(1,2)}$ are square matrices of order m and $n-m$, respectively, x is a $(n-m)\times m$–matrix, and $\mathbf{0}$ is a $m \times (n-m)$–0–matrix. To prove this, consider the effect of such a matrix $T(g_i)$ on a basis vector $|\phi_\mu\rangle = (0,0,\dots,1_\mu,0,\dots,0)^T$ with an 1–entry in the μth column. Let us assume now, that the basis vectors of L_n are arranged in such a way, that the first m vectors form a basis of L_m. If now $\mu \le m$, we find immediately, that all components with $\mu' > m$ of $T(g_i)|\phi_\mu\rangle$ must be zero.
Multiplying now two such matrices $T(g_i)T(g_j) = T(g_k)$, we see, that the result has the same form. More specifically, we find

$$
\begin{aligned}
t^{(1)}(g_i)t^{(1)}(g_j) &= t^{(1)}(g_k) \\
t^{(2)}(g_i)t^{(2)}(g_j) &= t^{(2)}(g_k) \\
x(g_i)t^{(2)}(g_j) + t^{(1)}(g_i)x(g_j) &= x(g_k)\,. \qquad (63)
\end{aligned}
$$

A set of matrices $T(g_i)$ having these properties is called a *reducible representation*, and we see, that this is connected with the existence of a corresponding proper subspace L_m.
Using, in addition to this, unitary matrices $\tilde{T}$ as representation, they have an even simpler form, namely

$$\tilde{T}(g_i) = \begin{pmatrix} \tilde{t}^{(1)}(g_i) & \mathbf{0} \\ \mathbf{0} & \tilde{t}^{(2)}(g_i) \end{pmatrix}, \tag{64}$$

and we can repeat this procedure in a way, that finally the $\tilde{T}$ reach a maximal reduced block diagonal form

$$\tilde{T} = \begin{pmatrix} \tilde{t}^{(1)} & & & \\ & \tilde{t}^{(2)} & & \\ & & \ddots & \\ & & & \tilde{t}^{(n)} \end{pmatrix}. \tag{65}$$

The block diagonal elements $\tilde{t}^{(i)}$ are then called *irreducible* representations of the group and the $\tilde{T}$ is called *fully reduced.*

2.2.4 Schur's Lemmas

There are – in principle – two lemmas going under this name.

1. The first one states, that if $\tilde{t}^{(i)}$ is an irreducible representation of a group $\mathcal{G}$ and if a matrix P commutes with all the matrices $\tilde{t}^{(i)}(g_j)$, then P must be a multiple of the unity matrix.

   One possibility to prove this theorem is to notice, that if the dimension of the $\tilde{t}^{(i)}$ is n, then P is a square $n \times n$–matrix. Since we are talking about unitary matrices $\tilde{t}^{(i)}$, we find, that each of the $\tilde{t}^{(i)}(g_j)$ has a complete set of n eigenvectors. Since P commutes with all $\tilde{t}^{(i)}(g_j)$, it has to have n linearly independent eigenvectors, i.e. it can be transformed into diagonal form. Thus, writing

   $$P|x_k\rangle = c_k|x_k\rangle\,, \tag{66}$$

   where the $|x_k\rangle$ are the eigenvectors and the c_k the eigenvalues, we find from the commutativity, that

   $$P\tilde{t}^{(i)}(g_j)|x_k\rangle = c_k\tilde{t}^{(i)}(g_j)|x_k\rangle\,, \tag{67}$$

   or, that for all g_j $\tilde{t}^{(i)}(g_j)|x_k\rangle$ is an eigenvector of P with the same eigenvalue. But these eigenvectors generate again a subspace L_m

invariant under $\mathcal{G}$. So, either the representation is reducible, contrary to the hypothesis, or this L_m must be identical with L_n, the space on which P and the $\tilde{t}^{(i)}$ act. Then, of course, all the eigenvalues are equal to each other, leaving P as a multiple of the $\mathbf{1}$–matrix.

2. The second lemma states, that if $\tilde{t}^{(i)}$ and $\tilde{t}^{(j)}$ are two irreducible representations of a group $\mathcal{G}$ with dimensions m_i and m_j, and if a matrix M satisfies

$$\tilde{t}^{(i)}(g_k)M = M\tilde{t}^{(j)}(g_k) \quad \text{for all} \;\; g_k \in \mathcal{G}\,, \tag{68}$$

then M is either the $m_i \times m_j$–null matrix or $\tilde{t}^{(i)}$ and $\tilde{t}^{(j)}$ are equivalent.

To prove this lemma, let us first state, that two representations can only be equivalent, if their dimensions are identical. We will come back later to this. In the meantime we take the hermitian conjugate of both sides of the equation above and multiply from the right with M. We end up with

$$M^\dagger \tilde{t}^{(i)}(g_k^{-1})M = \tilde{t}^{(j)}(g_k^{-1})M^\dagger M \tag{69}$$

and using Eq.(68) we find

$$M^\dagger M \tilde{t}^{(j)}(g_k^{-1}) = \tilde{t}^{(j)}(g_k^{-1})M^\dagger M\,. \tag{70}$$

Invoking Schur's Lemma 1, we find, that, since $M^\dagger M$ commutes with an irreducible representation, it must be a constant matrix, i.e. a multiple of the unity matrix,

$$M^\dagger M = c\mathbf{1} \quad \Longrightarrow \quad \det(M^\dagger M) = \det(M^\dagger)\det(M) = c^{m_i}\,. \tag{71}$$

If the two representations have the same dimension and if $c \neq 0$, then $\det(M) \neq 0$ and M^{-1} exists and we are allowed to multiply Eq.(68) with M^{-1} from the right, showing that the two representations are in fact equivalent. For $c = 0$ we find, that M has to be the null matrix, since taking any (i,i) element of $M^\dagger M$ we are left with

$$0 = \sum_k M^\dagger_{ik} M_{ki} = \sum_k |M_{ki}|^2\,. \tag{72}$$

Similar reasoning can be applied for the case of different dimensions for the two representations. Assuming $m_i < m_j$, we supply a quadratic $m_i \times m_i$ matrix M with $m_j - m_i$ rows of zeros. This gives a zero determinant for the new matrix M'.

2.2.5 The orthogonality theorem

We are now in the position to construct the orthogonality theorem, which is in a central position in representation theory of groups. Let us start with some matrix M given by

$$M = \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}(g_k) X \tilde{t}^{(j)}(g_k^{-1}) \,, \tag{73}$$

where the $\tilde{t}$ are again irreducible representations of $\mathcal{G}$ with dimensions m_i and m_j, respectively, and X is some arbitrary $m_i \times m_j$–matrix. Multiplying Eq.(73) from the left with $\tilde{t}^{(i)}(g_l)$ we end up with

$$\begin{aligned}
\tilde{t}^{(i)}(g_l) M &= \tilde{t}^{(i)}(g_l) \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}(g_k) X \tilde{t}^{(j)}(g_k^{-1}) \\
&= \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}(g_l g_k) X \tilde{t}^{(j)}(g_k^{-1}) \\
&= \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}(g_l g_k) X \tilde{t}^{(j)}(g_k^{-1} g_l^{-1}) \tilde{t}^{(j)}(g_l) \\
&= M \tilde{t}^{(j)}(g_l)
\end{aligned} \tag{74}$$

and find with help of Schur's Lemma 2, that $M \equiv \mathbf{0}$. Now we take the (p, q) element of Eq.(73) and for the fun of it, we choose $X_{pq} = \delta_{pm}\delta_{qn}$, i.e. a matrix which has zero entries everywhere but in the (m, n) position. Inserting this into Eq.(73) we have

$$0 = \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{sm}(g_k) \tilde{t}^{(j)}_{nt}(g_k^{-1}) = \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{sm}(g_k) \tilde{t}^{(j)*}_{nt}(g_k) \,. \tag{75}$$

Next we construct a matrix N as

$$N = \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}(g_k) X \tilde{t}^{(i)}(g_k^{-1}) \,, \tag{76}$$

and, in analogy with Eq.(74) we can rewrite this as

$$\tilde{t}^{(i)}(g_l) N = N \tilde{t}^{(i)}(g_l) \,. \tag{77}$$

Schur's lemma 1 now states, that in this case N must be a multiple of the unity matrix, $N = a\mathbf{1}$. The same assignment for the matrix X now leaves us with

$$a\delta_{st} = \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{sm}(g_k) \tilde{t}^{(i)}_{nt}(g_k^{-1}) \,. \tag{78}$$

So far, this is not particularly helpful. However, taking the traces on both sides of Eq.(76), we find that

$$
\begin{aligned}
\mathrm{Tr}(N) = am_i &= \sum_{k=1}^{m_i} \sum_{g_k \in \mathcal{G}} \sum_{p,q} \tilde{t}^{(i)}_{kp}(g_k) X_{pq} \tilde{t}^{(i)}_{qk}(g_k^{-1}) \\
&= \sum_{p,q} X_{pq} \sum_{k=1}^{m_i} \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{qk}(g_k^{-1}) \tilde{t}^{(i)}_{kp}(g_k) \\
&= \sum_{p,q} X_{pq} \sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{qp}(\mathbf{1}) \\
&= g\mathrm{Tr}(X) = g\delta_{mn}\,, \qquad (79)
\end{aligned}
$$

where g is the order of $\mathcal{G}$, and we have made use of the form of X. This translates Eq.(78) into

$$
\sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{sm}(g_k) \tilde{t}^{(i)}_{nt}(g_k^{-1}) = \frac{g}{m_i} \delta_{st} \delta_{mn} \qquad (80)
$$

and enables us to write the *Orthogonality theorem* for irreducible representations in the following, compact form

$$
\sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{sm}(g_k) \tilde{t}^{(j)}_{nt}(g_k^{-1}) = \frac{g}{m_i} \delta_{ij} \delta_{st} \delta_{mn}\,. \qquad (81)
$$

This theorem has quite an elegant interpretation. Assume, that there are c distinct irreducible representations of a finite group $\mathcal{G}$ of order g, and let us think of the distinct $\tilde{t}^{(i)}_{lm}(g_k)$ as discrete functions of g_k with some g discrete points. Now, for every value of i, l, m $(1 \leq i \leq c,\ 1 \leq \{k,\, m\} \leq m_i)$ we have one such function. Thus, all together we have m_i^2 functions for all of the c allowed values of i. Ergo we find

$$
N_f = \sum_{i=1}^{c} m_i^2 \qquad (82)
$$

such functions, each defining a g–dimensional vector space, since in this function space each function can be specified by its g "components". The orthogonality theorem enables us to interpret the left hand side of Eq.(81) as a scalar product, and we see, that all the different functions $\tilde{t}^{(i)}_{lm}(g_k)$ are orthogonal to each other. The only thing we do not know yet, is whether they span the full space of the group. But we know, that the number of independent vectors in a vector space can not exceed its dimension, in our case – identifying the functions with vectors – we finally find

$$\sum_{i=1}^{c} m_i^2 \leq g\,. \tag{83}$$

It can be shown, that in fact the "="–sign holds here.
As an examplatory analogy, consider the two–dimensional space of the spin functions for a particle of spin $j = 1/2$. Obviously, there are two basis functions related to the up– and down–orientation of the spin with respect to an arbitrary axis, say the z axis. Thus, all other spin constellations can be expressed as linear combinations of the two functions

$$\chi^{\pm 1/2}(s_z)\,,$$

which are clearly ortogonal to each other. This reasoning can be easily extended to higher spin configurations.

2.3 Characters of a representation

2.3.1 Definition

As we have already learned, two representations $T_{1,2}$ of a group might be equivalent, i.e. Eq.(53) holds,

$$T_1 = S^{-1} T_2 S\,.$$

In other words, representations of a group in a vector space are not unique, but might be transformed into other representations via a similarity transformation S. However, there are quantities related to the matrices T, which remain invariant under such transformations and are therefore suitable to characterize the invariant properties of the representations. One example is the set of eigenvalues of the matrices, their use however leads to the construction of Casimir–operators. We will come back to this later. Another possibility is the use of the traces of the matrices. In fact, the traces

$$\chi(g_i \in \mathcal{G}) = \sum_{j}^{d} \tilde{T}_{jj}(g_i) \tag{84}$$

of the representation matrices $T(g_i)$ for the group elements g_i form another invariant under similarity transformations and are called the *characters of the representation* $T(\mathcal{G})$.
Their invariance is pretty evident, when one considers the cyclic property of the trace applied to Eq.(53).

In fact, it turns out, that the individual representations of elements of the same classes have identical characters. This is easy to prove, given the fact, that elements of the same class are connected via another similarity transformation. A nice feature of the characters of the irreducible representation is, that they follow the orthogonality theorem, Eq.(81), too. By just replacing indices s with m and t with n, we have

$$\sum_{g_k \in \mathcal{G}} \tilde{t}^{(i)}_{mm}(g_k)\tilde{t}^{(j)}_{nn}(g_k^{-1}) = \sum_{g_k \in \mathcal{G}} \chi^{(i)}(g_k)\chi^{(j)*}(g_k) = g\delta_{ij} \,. \tag{85}$$

We can now split our group $\mathcal{G}$ in n classes $\mathcal{C}_i$ with c_i elements and find

$$\sum_{c=1}^{n} c_c \chi^{(i)}(\mathcal{C}_c)\chi_c^{(j)*}(\mathcal{C}_c) = g\delta_{ij} \,. \tag{86}$$

This suggests, that we can think of the irreducible characters as a set of orthogonal vectors in some n dimensional space. From this we are lead to the conclusion, that there can not be more than n different irreducible representations. Or stated in another way, the number of irreducible representations is limited by the number of classes. However, the Equation above, (85), can be expressed alternatively as

$$\sum_{i} c_c \chi^{(i)}(\mathcal{C}_c)\chi_c^{(i)*}(\mathcal{C}_{c'}) = g\delta_{cc'} \,, \tag{87}$$

where the sum runs over all irreducible representations i and the arguments are the characters of the classes c and c', which are orthogonal, too.

2.3.2 Reducing representations

We will now turn to the solution of a problem, which may have arised in the turn of our definitions so far. Let us assume, that we have a representation of a group, for example our group D_4, which is in general reducible. The question now is how to reduce this representation ?
A first step is the observation, that the number of irreducible representations of a group equals the number of its classes. In the case of D_4 we have five classes, cf. Eq.42, consequently we expect five irreducible representations with dimension d_i. In accordance with Eq.83 we know, that

$$\sum_i d_i^2 = g \quad \stackrel{D_4}{\Longrightarrow} \quad 1+1+1+1+4=8 \tag{88}$$

which is the only possible solution for D_4 and integer dimensions of the representations.
Using Eqs. (85) and (87), we now construct a character table for D_4,

	$\mathcal{C}_1$ $\{E\}$	$\mathcal{C}_2$ $\{\mathcal{R}_{90,270}\}$	$\mathcal{C}_3$ $\{\mathcal{R}_{180}\}$	$\mathcal{C}_4$ $\{\mathcal{M}_{x,y}\}$	$\mathcal{C}_5$ $\{\mathcal{M}_{13,24}\}$
$\chi^{(1)}$	1	1	1	1	1
$\chi^{(2)}$	1	-1	1	1	-1
$\chi^{(3)}$	1	-1	1	-1	1
$\chi^{(4)}$	1	1	1	-1	-1
$\chi^{(5)}$	2	0	-2	0	0

where the classes are listed in columns and the representation in rows.
The construction of this table is pretty simple, starting from the (so–called unfaithful) identity representation with unity characters, we have the first row. Additionally, we observe, that in any representation the identity element $\mathcal{E}$ is suitably represented by the identity matrix of the corresponding dimension, leaving us with the first column. Since for one dimensional representations (rows 1 to 4), the characters themselves must represent the multiplication table, we can conclude, that for any element, whose square equals $\mathcal{E}$, i.e. the $\mathcal{M}$ and $\mathcal{R}_{180}$, the only allowed characters are ± 1. For the latter, $\mathcal{R}_{180}$, let us consider the product $\mathcal{M}_x \mathcal{M}_y$. We immediately conclude, that in the one–dimensional representations $\mathcal{R}_{180} \equiv 1$. Orthogonality of the four rows yields then immediately the result displayed above. The fifth row is constructed by observing, that for every square matrix with orthogonal rows, the columns are orthogonal, too. We would like to mention here, that similar reasoning holds true for the construction of the character table of any finite group. However, it might happen that relations like $\mathcal{M}_x^2 = 1$ might generalize to $\mathcal{M}_x^n = 1$, leaving us with complex characters.
We continue now in our journey to the irreducible representations of D_4. Since for one dimensional representations the characters equal the representations, we have found immediately the first four irreducible representations of dimension one and are left with the fifth of dimension $d = 2$, only. To construct this representation, let us start by introducing to basis vectors, which we choose conveniently to be the positive x and the y axes of the square. Simple inspection shows, that $\mathcal{R}_{90}$ acts on the base vectors as

$$\mathcal{R}_{90} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{89}$$

and we have found the representation of $\mathcal{R}_{90}$, which is traceless, as indicated by the character table. Use of the multiplication table and similar reasoning helps us to find all the matrices representing the other elements of our group.
♣ As an exercise, repeat the program outlined above for the symmetry group D_3 of a triangle.

3 Lie Groups and Lie Algebras

In principle, all the theorems and nice properties stated in the section above hold also true for discrete infinite groups, real trouble occurs only for continuous groups, and only then, some modifications are really needed. From a physicists point of view, the so–called Compact Lie Groups, which are basically continuous groups with some extra properties, are of special interest. But before getting into closer contact with these groups, we will briefly give some more examples of continuous groups :

- As already stated above, the set of all real numbers form a continuous group under addition.
- Linear transformations : $x \to x' = T(a,b)x = ax + b$ with a, b arbitrary real numbers and $a \neq 0$.
  $T(a,b) = T(a_1,b_1)T(a_2,b_2) = T(a_1a_2, a_1b_2 + b_1)$, ergo in this group the identity element is $\mathbf{1} = T(1,0)$ and the inverse is
  $T^{-1}(a,b) = T(1/a, -b/a)$.
- The set of all displacements D for a n–dimensional vector space, $\vec{x}' = D\vec{x} = \vec{x} + \vec{d}$. Here, the identity element is the null vector, $\vec{d} = \vec{0}$ and the inverse of a displacement is the corresponding mirrored vector, $-\vec{d}$.
- The set of all linear homogenous transformations in a n–dimensional vector space. For two dimensions, this is

  $$\vec{x}' = A\vec{x} \quad \text{with } \det(A) \neq 0\,. \tag{90}$$

  Obviously, this is a group with n^2–parameters and it is denoted by $GL(2)$. This group is known as the linear group.
- The group of rotations around a common axis in the two–dimensional plane with angle $\theta \in [0, 2\pi]$. The multiplication of two operations can be defined as their subsequent application, then the individual angles merely add to yield the total angle of the resulting rotation. The identity element is the zero rotation, the inverse element is the rotation with the negative angle. This group is denoted by $SO(2)$, and is represented by 2×2 real matrices with determinant 1.
- The set of all rotations around all axes passing through one fixed point in the three dimensional space. The group elements can be characterized by the three Euler–angles and is commonly known as $SO(3)$.

3.1 Lie Groups, the basics

3.1.1 Definition

Stated in an intuitive manner (suitable for physicists), a Lie group is a continuous group, i.e. the elements are labeled by some continuous real

parameter, where the multiplication law depends smoothly on these parameters. In other words, the multiplication law can be rewritten as a function of these parameters, the corresponding inverse elements exist and can be labeled as smooth functions of the parameters.
In a fashion better suited for mathematicians, this can be expressed formally as :

- Consider two elements $x_{1,2}$ of the group, which can be mapped in a one–to–one correspondence on some elements $P(x_{1,2})$ of a suitable r–dimensional real inner product space, their images. Assume, we have $x_3 = x_1 x_2$. This composition law is dubbed continuous in x_2, if for every $\epsilon > 0$ it is possible to find an $\delta_\epsilon > 0$, so that for all $||P(x) - P(x_2)|| < \delta_\epsilon$ the element labeled by $x_1 x$ is in the ϵ–neighborhood of $P(x_3)$, i.e. $||P(x_1 x) - P(x_3)|| < \epsilon$. More intuitively, small changes in the factors produce only small changes in the product, when mapped from the parameters to the group elements. Similarly, we can define continuity for the inverse, boiling down to small changes in an element produces small changes in the corresponding inverse. A group fulfilling these criteria is called *topological*.

- Let us now assume, that we can write the elements of a topological group as $x_{1,2} = x_{1,2}(a^{1,2}_{i=1\ldots r})$. Consider the elements $x_3 = x_1 x_2$ and $x_4 = x_1^{-1}$. The parameters $a^{3,4}$ can be expressed as functions of the $a^{1,2}$. A topological group is called an r–dimensional *Lie Group*, if there exists a neighborhood N of the identity element $\mathbf{0}$, such that the continuous parameters of a product of two elements in N or the inverse of an element in N are continuous differentiable functions of the parameters. I.e. if the corresponding $a_i^{3,4}$ are analytic functions of the $a_i^{1,2}$ provided all four elements are in N, the group is called a Lie group. Choosing $\mathbf{0} = x(0,0,\ldots,0)$ we find, that

  $$x(0,0,\ldots,\epsilon_j,\ldots,0) = x(0,0,\ldots,0) + i\epsilon_j I_j(0,0,\ldots,0) + \mathcal{O}(\epsilon^2) \tag{91}$$

  For obvious reasons, the r operators I_j are called the generators of the Lie group, since their multiple application spans the group.

  Here, a warning is in order. Basically, there are two relevant spaces, which must not be confused. First, there's the r–dimensional space of the group generators, and second, there's the (physical) Hilbert space, on which the elements of the Lie group act. Since, however, for a compact Lie group, we can always take the latter to be a finite dimensional space, we can think of the generators as a set of r finite dimensional hermitian matrices.

As indicated above, any group element, which can be derived from the identity element by continuous changes in the parameters can be re-expressed as

$$\exp(i\alpha_l I_l)\ , \tag{92}$$

where the summation over l is understood. The I_l are the linearly independent generators defined by

$$I_l = \lim_{\epsilon_l \to 0}\left[\frac{x(0,\ldots,\epsilon_l,\ldots,0) - x(0,0,\ldots,0)}{i\epsilon_l}\right] \tag{93}$$

and the α_l are real parameters. All elements of a Lie group within a subset containing the identity element can therefore be generated with help of the equation above, i.e. by assigning specific values α_l, thus the name "generators" for the I_l becomes more transparent, and it is immediately understood, that an r–dimensional Lie group has r generators.
There is a nice thing about the infinitesimal elements of a Lie group, namely, that they form an Abelian group. Consider the elements $x_{j,k} = x(0,0,\ldots,\epsilon_{j,k},\ldots,0)$, where the non–zero entries do not appear in the same position. It is easy to show then, that

$$x_j x_k = x_k x_j = \mathbf{0} + i(\epsilon_j I_J + \epsilon_k I_k) + \mathcal{O}(\epsilon^2)\,. \tag{94}$$

3.1.2 Representations of a Lie group

This leads us immediately to the issue of how to represent a Lie group. The answer to this is pretty straightforward and already indicated above : Lie groups are represented by finite dimensional matrices.
Such representations T are dubbed *continuous* if they fulfill

$$T(x) \to T(x') \quad \text{as} \quad x \to x'\ , \tag{95}$$

where we have characterized the group elements by the continuous parameters x and x'. Extending our previous results on finite groups, we find the following properties of representations of Lie groups :

1. For any representation of a Lie group we can find an equivalent representation with unitary matrices.

2. Any unitary representation can be cast into block diagonal form as displayed in Eq.(65).

3. Any irreducible representation is finite dimensional.

3.1.3 Generator algebra

The more interesting thing about the generators I_l is, that they define to a large extend the structure of the group. The basic point is, that they satisfy simple commutation relations

$$[I_a, I_b] = I_a I_b - I_b I_a = i f_{abc} I_c \,, \tag{96}$$

which can be understood considering the product

$$\begin{aligned} \exp(i\lambda I_a)\exp(i\lambda I_b)\exp(-i\lambda I_a)\exp(-i\lambda I_b) = \\ 1 + \lambda^2 [I_a, I_b] + \mathcal{O}(\lambda^3) \equiv \exp(i\kappa I_c) \end{aligned} \tag{97}$$

and comparing coefficients. In addition, the generators satisfy the Jacobi–identity, i.e.

$$[I_a, [I_b, I_c]] + \text{ cycl.} = 0 \,. \tag{98}$$

Let us remark here, that the commutator relation above, Eq.(96), replaces the multiplication table for finite groups in the sense, that it determines completely the structure of the Lie group. Thus, the constants f_{abc} are called *Structure Constants* of the group. They are a characteristic property of the group and not of the representation chosen. However, they are not unique, since the generators themselves are not unique. After a brief note on Lie algebras we will come back to this issue.
♣ As an exercise, we leave the proof of the Jacobi identity for the reader.

3.1.4 Lie Algebras

The r generators of a Lie group form the basis vectors of an r–dimensional vector space. The set of real linear combinations of the generators is called a Lie algebra. Such an algebra L has the following properties :

1. If $x, y \in L$, then $[x, y] \in L$.
2. $[x, y] = -[y, x]$.
3. The Jacobi identity holds, $[x, [y, z]]$+cycl.$= 0$.

Since the generators of a Lie group together with the commutation relation form a Lie algebra, it is easy to see, that the structure constants fulfill

$$f_{abc} = -f_{bac} \,. \tag{99}$$

Even more, Eq.(98) translates into

$$f_{bcd}f_{ade} + f_{abd}f_{cde} + f_{cad}f_{bde} = 0\,. \tag{100}$$

and defining a set of new matrices J via

$$(J_a)_{bc} = -if_{abc} \tag{101}$$

we find

$$[J_a, J_b] = if_{abc}J_c\,. \tag{102}$$

Stated in less formal words, the structure constants f_{abc} themselves generate a representation of the algebra, which is called the *adjoint representation*. In contrast, the representation generated by the I is called *fundamental representation*. The dimension of this adjoint representation equals the number of generators, since this is the space, on which it acts. This is exactly the number of real parameters needed to describe every group element.
A subalgebra is called *invariant subalgebra*, if the set of generators constituting it maps into itself under commutation. An algebra without non–trivial subalgebras (trivial are the full algebra and the empty set) is dubbed *simple*. *Abelian subalgebras* have generators, which commute with everything. Therefore they can be identified with simple $U(1)$ phase factors on the group elements. A nasty feature of these generators is, that they do not show up in the structure constants. That means, if I_a is such an $U(1)$–generator, $f_{abc} \equiv 0$ for all possible b and c. As we will see later, this feature complicates the issue of finding a suitable representation of the group. However, algebras without such annoyances are called *semi–simple*, their generators have at least one non–zero commutator with another generator. It can be shown, that they are constructed by putting together simple algebras.

3.1.5 Construction of representations

We would like to come back now to the non–uniqueness of the structure constants, which are dependent on the choice of the basis for the vector space of the generators. A suitable choice can be obtained via the adjoint representation, defined in Eq.(101). Consider the trace

$$Tr(J_aJ_b) = \sum_{d,e} f_{ade}f_{edb}\,. \tag{103}$$

J_aJ_b is a real, symmetric matrix. Therefore it can be diagonalized with real eigenvalues by appropriate choices of the generators I, i.e. by replacing the original I with suitable linear combinations, and, correspondingly, by linear combinations of the generators of the adjoint representation J. Thus,

$$Tr(J_a J_b) = k_a \delta_{ab} \,, \tag{104}$$

and since we are free to rescale the generators, we could choose all k_a in a way, that $|k_a| = 1$. In the framework of this lecture, we will be mainly concerned with compact, semi–simple Lie–groups, which are characterized by positive k_a, which we rescale to some common κ. Then we see, that in this particular basis, the f_{abc} are completely antisymmetric, because we can identify

$$f_{abc} = -\frac{i}{\kappa}\,\mathrm{Tr}([J_a, J_b]\, J_c) \tag{105}$$

and use the cyclic property of the trace. In this particular basis, the J_a are hermitian matrices.

3.1.6 Casimir operators

Finally, let us note, that we are able to construct operators, that commute with all generators of a Lie group. Such operators are dubbed *Casimir operators*, and the number of independent Casimir operators defines the *rank* of the group. For $SU(2)$, it can be shown, that the operator

$$J^2 = J_1^2 + J_2^2 + J_3^2 \tag{106}$$

commutes with all the J_a. For $SO(3)$ the only Casimir operator looks similarly. Since the Casimir operators can then be diagonalized simultaneously with the generators, their eigenvalues may be used to label the irreducible representations of the group.

3.2 Some first examples : Fun with $SO(2)$ and $SU(2)$

We are now going to give first brief examples of Lie groups and construct their representations. We chose the groups $SO(2)$ and the simplest non–Abelian Lie group, we can make of, namely $SU(2)$. Later in this section, we will see, that $SU(2)$ has some nice impact on nuclear and particle physics in the form of the isospin group.

3.2.1 $SO(2)$

The group $SO(2)$ describes the set of all rotations of a circle around the axis perpendicular to its plane through its center. Each element can be characterized by one angle $\theta \in [0, 2\pi[$. Integer multiples of 2π added to the interval can obviously be ignored, since they do not add any new physical information. The law of composition for group elements $T(\theta)$ is

$$T(\theta) \cdot T(\phi) = T(\phi) \cdot T(\theta) = \begin{cases} T(\theta + \phi) & \text{if } \theta + \phi < 2\pi \\ T(\theta + \phi - 2\pi) & \text{if } \theta + \phi \geq 2\pi \end{cases} \tag{107}$$

The identity element is $T(0)$, inverse elements are

$$T^{-1}(\phi) = T(2\pi - \phi) \,. \tag{108}$$

A representation of this group can be constructed by studying the effect of the group elements on some point (x, y) in the plane. Then we find

$$T(\phi) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} , \tag{109}$$

which is clearly an orthogonal 2×2–matrix with determinant 1.
Coming back to Eq.(107), we see, that this group is Abelian, and following our previous remarks, we are lead to some alternative representation via $U(1)$ factors, namely $\exp(ic\phi)$. These terms are identical with the characters χ of the group, too,

$$\chi(\phi) = \exp(ic\phi) \tag{110}$$

with the constraint, that c is some integer, due to the representation of the identity element and the law of composition. From this we can conclude, that we have some irreducible 1–dimensional representation of $SO(2)$ for every integer value m of c. The orthogonality theorem then reads

$$\int_0^{2\pi} d\phi \chi^{(m)*}(\phi) \chi^{(m')}(\phi) = 2\pi \delta_{mm'} \,. \tag{111}$$

Here we have used Eq.(85) with a suitable normalization and with the replacement of the summation over the – finite – number of group elements with an integration over the space parametrizing the – infinite – number of elements of a Lie group.
Due to its property as one–parameter group, $SO(2)$ has one generator, namely either

$$I = \lim_{\phi \to 0} \frac{1}{i\phi} \left[\exp(im\phi) - 1\right] = m \tag{112}$$

or

$$I = \lim_{\phi \to 0} \frac{1}{i\phi} \left[\begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{113}$$

depending on the representation chosen for the group elements. Having identified the result for the generator in the second choice with the Pauli–matrix σ_y, we can rewrite

$$T(\phi) = \exp(i\phi\sigma_y) . \tag{114}$$

We can find yet another generator with some more physical meaning by considering the effect of the $T(\phi)$ on some function $f = f(x, y)$ defined on the plane, the group acts on. Then we have

$$T(\phi)f(x, y) = f(x\cos\phi + y\sin\phi, -x\sin\phi + y\cos\phi) \tag{115}$$

and the generator can be found via

$$\begin{aligned} I \cdot f(x, y) &= \lim_{\phi \to 0} \frac{1}{i\phi} \left[f(x\cos\phi + y\sin\phi, -x\sin\phi + y\cos\phi) - f(x, y) \right] \\ &= -i \left(\frac{y\partial}{\partial x} - \frac{x\partial}{\partial y} \right) f(x, y) . \end{aligned} \tag{116}$$

Identifying the component of the angular momentum normal to the (x, y)–plane L_z with

$$L_z = i\hbar \left(\frac{y\partial}{\partial x} - \frac{x\partial}{\partial y} \right) = -i\hbar \frac{\partial}{\partial\phi} \tag{117}$$

we are left with

$$T(\phi) = \exp\left(-\frac{i\phi L_z}{\hbar} \right) . \tag{118}$$

3.2.2 $SU(2)$

The simplest non–Abelian Lie–algebra, we could make up, consists of three generators J_a with structure constants given by the completely antisymmetric tensor, $f_{abc} = \epsilon_{abc}$. Then the commutation relations read

$$[J_a, J_b] = i\epsilon_{abc}J_c \tag{119}$$

and we recover the angular momentum algebra. To construct the representations of the algebra (119), let us forget everything we know about the special role of the Casimir operator $\vec{L}^2$ in angular momentum or analogously of J^2 for $SU(2)$. Assume, that we have an N–dimensional irreducible representation of $SU(2)$, and that we can diagonalize one of the hermitian operators, say J_3. Now let us look for the α eigenstates $|j,\alpha\rangle$ of J_3 with maximal eigenvalue j.

$$J_3|j,\alpha\rangle = j|j,\alpha\rangle\ , \tag{120}$$

and reshuffle them in a way, that they're orthogonal,

$$\langle j,\beta|j,\alpha\rangle = \delta_{\alpha\beta}\ . \tag{121}$$

As usual we construct the non–hermitian ladder operators

$$J_\pm = \frac{1}{\sqrt{2}}(J_1 \pm iJ_2) \tag{122}$$

with

$$[J_3, J_\pm] = \pm J_\pm\ ,\ \ [J_+, J_-] = J_3\ . \tag{123}$$

By acting on eigenstates $|m\rangle$ of J_3, these ladder operators raise or lower the eigenvalue m of them by one unit,

$$J_3J_\pm|m\rangle = J_\pm J_3|m\rangle \pm J_\pm|m\rangle = (m\pm 1)J_\pm|m\rangle\ . \tag{124}$$

Consequently, when acting on the eigenstate with the highest eigenvalue, J_+ yields

$$J_+|j,\alpha\rangle = 0\ . \tag{125}$$

Conversely, with some appropriate normalization $N_j(\alpha)$ the states

$$J_-|j,\alpha\rangle = N_j(\alpha)|j-1,\alpha\rangle \tag{126}$$

are orthogonal,

$$\begin{aligned} N_j^*(\beta)N_j(\alpha)\langle j-1,\beta|j-1,\alpha\rangle &= \langle j,\beta|J_+J_-|j,\alpha\rangle \\ &= \langle j,\beta|[J_+,J_-]|j,\alpha\rangle \\ &= \langle j,\beta|J_3|j,\alpha\rangle = j\delta_{\alpha\beta}\ . \end{aligned} \tag{127}$$

When replacing the product with the commutator in the equation above, we have used the fact, that J_+ annihilates the state $|j,\alpha\rangle$. Thus, we have

$$|N_j(\alpha)|^2 = j \tag{128}$$

and we are free to choose phases of the $N_j(\alpha)$ in order to make them real and positive numbers,

$$N_j(\alpha) = \sqrt{j}\,. \tag{129}$$

Therefore,

$$\begin{aligned} J_+|j-1,\alpha\rangle &= \frac{1}{N_j} J_+ J_- |j,\alpha\rangle \\ &= \frac{1}{N_j}[J_+, J_-]|j,\alpha\rangle = N_j|j,\alpha\rangle \end{aligned} \tag{130}$$

We are free to repeat this procedure, and step by step we find orthonormal states $|j-k,\alpha\rangle$ satisfying

$$\begin{aligned} J_-|j-k,\alpha\rangle &= N_{j-k}|j-k-1,\alpha\rangle \\ J_+|j-k-1,\alpha\rangle &= N_{j-k}|j-k,\alpha\rangle \end{aligned} \tag{131}$$

with a recursion relation among the N_{j-k},

$$\begin{aligned} N_{j-k}^2 &= \langle j-k,\alpha|J_+J_-|j-k,\alpha\rangle \\ &= \langle j-k,\alpha|[J_+,J_-]|j-k,\alpha\rangle + \langle j-k,\alpha|J_-J_+|j-k,\alpha\rangle \\ &= j-k+N_{j-k+1}^2\,. \end{aligned} \tag{132}$$

Adding up these normalization factors from N_j^2 down to N_{j-k}^2 we find

$$N_{j-k}^2 = \frac{(k+1)(2j-k)}{2} \tag{133}$$

or, introducing $m = j-k$

$$N_m = \sqrt{\frac{(j-m+1)(j+m)}{2}}\,. \tag{134}$$

Climbing down the ladder, we will arrive sooner or later at some state $|j-l\rangle$ that gets annihilated by J_-, $J_-|j-l\rangle = 0$. But this requires

$$N_{j-l} = 0 \implies l = 2j \,. \tag{135}$$

Since l is an integer number, the j are half–integers, and the space breaks up into α subspaces of dimension $2j+1$ plus possibly a subspace which we might not have found yet. This is exactly, what we would have found taking into account our knowledge about the Casimir operator J^2 and its eigenvalues. On the subspaces labeled by α, however, the J's act in a block–diagonal form, since the orthogonality in the indices α and β, Eq.(127), remains conserved throughout our ladder procedure. But we have assumed, that we already worked in some irreducible representation, therefore there is only one α. We can use this, to break down any representation into its irreducible components. Starting with any representation we find the block–diagonal form, where the highest j–representation is explicitly singled out. We can repeat this procedure with the remaining block until we have decomposed the representation completely.
The eigenvalues of J_3 are called weights, and j is the highest weight. The irreducible representations are characterized by the highest weight j and they are called spin–j representations. In fact, they are associated with a particle at rest and spin angular momentum j, given by $J^2 = \hbar^2 j(j+1)$. The existence of hermitian angular momentum operators in a problem allows us to decompose the Hilbert space of the full problem into states $|j, m, \alpha\rangle$ in the spin–j representation, where m is the eigenvalue of J_3, j is related to the eigenvalues of J^2 and α stands for all other labels, which have to be introduced, in order to characterize the states. We can construct states satisfying

$$\langle m, j, \alpha | m', j' \alpha' \rangle = \delta_{mm'} \delta_{jj'} \delta_{\alpha\alpha'} \,. \tag{136}$$

The δ–function in the α is in analogy with what we have discussed above, the δ–function in m and m' results from the fact, that they are eigenvalues of an hermitian operator (which has different eigenvectors and eigenvalues). Even without employing the fact, that the operator J^2 is hermitian, too, we can show, that the δ–function in j and j' holds. Suppose, that we have $j' > j$ and consider

$$\langle j, j, \alpha | j, j', \beta \rangle = \sqrt{\frac{2}{(j'+j)(j'-j+1)}} \langle j, j, \alpha | J_- | j+1, j', \beta \rangle = 0 \tag{137}$$

Let us now construct explicitly irreducible representations of $SU(2)$:

- $\underline{j = 1/2 :}$
  The simplest non–trivial representation of this group has dimension 2, i.e. $j = 1/2$. The vectors, in which J_3 is already diagonal are then $|1/2\rangle = (1,0)^T$ and $|-1/2\rangle = (0,1)^T$. Hence, the ladder operators have the form

$$J_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \ J_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ . \tag{138}$$

Thus,

$$J_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ J_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \tag{139}$$

and by use of the commutator relations, we find

$$J_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{140}$$

Not unsurprisingly, we recover the common Pauli–matrices σ_i. So, the simplest, non–trivial representation of $SU(2)$ is generated by $\sigma_i/2$.

- $\underline{j = 1 :}$
For the spin–1 representation we have the following basis vectors : $|1\rangle = (1,0,0)^T$, $|0\rangle = (0,1,0)^T$ and $|-1\rangle = (0,0,1)^T$. The ladder operators are

$$J_+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \ , \ J_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \ . \tag{141}$$

and we find

$$\begin{aligned} J_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \ , \ J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ , \\ J_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \ . \end{aligned} \tag{142}$$

♣ As an exercise, construct the matrix representation for the spin–3/2 representation.
We should come back once more to our statement above, that the highest weight construction allows us to decompose any representation into simpler irreducible representations with decreasing weights. We can revert this procedure and construct arbitrary representations out of the simpler

irreducible representations. The first way is to build the direct sum of two representations $D_n\mathcal{G}$ and $D_m\mathcal{G}$ of a group $\mathcal{G}$ with dimension n and m, respectively. We find the block–diagonal matrices

$$D_n(\mathcal{G}) \oplus D_m(\mathcal{G}) = \begin{pmatrix} D_n(\mathcal{G}) & \mathbf{0} \\ \mathbf{0} & D_m(\mathcal{G}) \end{pmatrix} , \tag{143}$$

which is generated via

$$I_a^{D_n \oplus D_m} = \begin{pmatrix} I_a^{D_n} & \mathbf{0} \\ \mathbf{0} & I_a^{D_m} \end{pmatrix} . \tag{144}$$

Obviously the result is $n+m$–dimensional. Instead, we could also form the direct product $D_n \otimes D_m$. Denoting the basis vectors by $|n\rangle$ and $|m\rangle$ we are lead to write

$$(D_n \otimes D_m)(\mathcal{G})|n\rangle|m\rangle = D_n(\mathcal{G})|n\rangle D_m(\mathcal{G})|m\rangle . \tag{145}$$

In matrix language this translates into

$$[(D_n \otimes D_m)(\mathcal{G})]_{nm,n'm'} = [D_n(\mathcal{G})]_{nn'}[D_m(\mathcal{G})]_{mm'} . \tag{146}$$

The generators are then obtained by differentiating the equation above. They read

$$\begin{aligned} \left[I_a^{D_n \otimes D_m}\right]_{nm,n'm'} &= \left[I_a^{D_n}\right]_{nn'} \delta_{mm'} + \delta_{nn'} \left[I_a^{D_m}\right]_{mm'} \\ &= I_a^{D_n} \otimes \mathbf{1} + \mathbf{1} \otimes I_a^{D_m} = I_a^{D_n} + I_a^{D_m} . \end{aligned} \tag{147}$$

It should be noted here, that relations like the one in Eq.(147) apply for instance in the addition of angular momenta. The tricky point here is obviously, to keep track of what matrix acts on which space.
As a simple example, let us study the addition of two spins. In general, spins are described by suitable spin–j representations of $SU(2)$. For instance, for spin–1/2 particles, we can identify suitable basis vectors by

$$J_3|\uparrow\rangle = \frac{1}{2}|\uparrow\rangle , \ J_3|\downarrow\rangle = -\frac{1}{2}|\downarrow\rangle , \tag{148}$$

thus the equations

$$|\uparrow\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle , \quad |\downarrow\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle \tag{149}$$

identify the $|s, m\rangle$ quantum numbers related to the particular spin state. Similarly, spin–1 particles are identified by

$$| \uparrow\rangle = |1, 1\rangle \ , \quad | \leftrightarrow\rangle = |1, 0\rangle \ , \quad | \downarrow\rangle = |1, -1\rangle \ . \tag{150}$$

To add two spins, for instance two electrons with orbital angular momentum suppressed, we may write

$$S = S_1 + S_2 = S_1 \otimes \mathbf{1} + \mathbf{1} \otimes S_2 \tag{151}$$

for the spin operators of the compound system in terms of the two single spin operators. Since S_1 and S_2 act on different electrons, we have

$$[S_{1i}, S_{2j}] = 0 \, , \tag{152}$$

i.e. they commute. Of course, commutators within the space of the individual electrons, the spin operators obey the usual commutation laws, see Eq.(119),

$$[S_{1i}, S_{1j}] = i\epsilon_{ijk} S_{1k} \ , \quad [S_{2i}, S_{2j}] = i\epsilon_{ijk} S_{2k} \, .$$

As a direct consequence, the total spin operator obeys the same commutation equations,

$$[S_i, S_j] = i\epsilon_{ijk} S_k \, .$$

The eigenvalues of the significant operators are then given by

$$\begin{array}{lcllcl} S^2 & = & (S_1 + S_2)^2 & : & s(s+1) \\ S_3 & = & S_{13} + S_{23} & : & m \\ S_{13} & & & : & m_1 \\ S_{12} & & & : & m_2 \end{array} \tag{153}$$

We can immediately construct two alternative representations, based on two different choices of the basis vectors for the compound system.

1. The so–called $\{m_1, m_2\}$ representation based on

$$| \uparrow\uparrow\rangle \, , \quad | \uparrow\downarrow\rangle \, , \quad | \downarrow\uparrow\rangle \, , \quad | \downarrow\downarrow\rangle \, . \tag{154}$$

2. The $\{s, m\}$ or – in this case – triplet–singlet representation (or $(1, 0)$ representation) based on

$$
\begin{aligned}
|1,1\rangle &= |\uparrow\uparrow\rangle\,, \\
|1,-1\rangle &= |\downarrow\downarrow\rangle\,, \\
|1,0\rangle &= \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) \\
|0,0\rangle &= \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right)\,.
\end{aligned}
\tag{155}
$$

The spin assignments are not to complicated to understand, obviously,

$$
S_3|1,1\rangle = |1,1\rangle\,, \tag{156}
$$

which coincides with physical intuition. Applying the ladder operator S_- on this state, and keeping track of the proper normalization for this operator, we find

$$
S_-|\uparrow\uparrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) = |1,0\rangle\,. \tag{157}
$$

Similarly, we obtain $|1,-1\rangle$. The singlet state $|1,0\rangle$ is then obtained by the condition, that the four states should form an orthonormal basis. Let us note, that the normalizations for these basis vectors are called *Clebsch–Gordan coefficients*. They relate the $\{m_1, m_2\}$ with the $\{s, m\}$ representation. Later, we will work out recursion relations for the general cases of spin additions.

♣ In quite a similar way, we can work out the $\{s, m\}$ representation for the addition of a spin–1/2 particle and a spin–1 particle into a compound system.

3.3 Physics with $SU(2)$: Isospin

3.3.1 Motivation

As a first example of the impact of group theory on quantum physics we would like to study the isospin group. Isospin is an internal symmetry with no obvious connection to symmetries related to the structure of space and time. In fact, it is related to the change of particle identities and consequently, it falls in the domain of particle physics. However, it was first discussed and applied in nuclear physics, and this is the place we want to introduce it. Actually, it was the big similarity between protons and neutrons, that lead to the discovery, that these particles can be understood as two different states of the same particle the nucleon. The properties of the nucleons are displayed in Table 1. The idea is now, that the state of a nucleon can be characterized by

	mass [MeV=mc^2]	spin	isospin	lifetime	magn. moment [$\mu_N^2 = e\hbar/(2m_pc)$]
p	938.213	1/2	1/2	stable	2.793
n	939.507	1/2	$-1/2$	918 ± 14s	-1.913

Table 1: Properties of the nucleons.

its coordinates and its momentum, by its spin, and finally by its isospin, which is $\tau = \pm 1/2$. When studying, how the protons and neutrons stick together in a nucleus, we can think of three forces at work and ignore gravity. These forces are

- the strong force, sticking together the nucleons,
- electromagnetism, causing the mutual repulsion of the protons,
- and the weak force, responsible for β–decays.

Let us note, that the last two forces are responsible for the differences in the properties of the proton and neutron and then let us forget about them for the rest of this section. We will assume in the following, that the strong force is independent of the nucleon type, i.e. it is the same for pp, pn and nn interactions. Let us look now at any nucleus and let us assume, that we replace all n's by p's. Since the forces are the same, we should find the same ground–state energy for the nucleus, shouldn't we ? This is only nearly right, since the n's are identical particles, thus we have to sum over all replacements $n \to p$. Let us come back to this later, and concentrate first on the operators involved so far. They are :

- $$\tau_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{158}$$

  with eigenvalues $\pm 1/2$ and eigenvectors $|p\rangle = (1,0)^T$ and $|n\rangle = (0,1)^T$.

- $$\tau_+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \tau_- = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{159}$$

  which, up to some normalization, transforms protons into neutrons and vice versa.

But we already know these operators ! Linear combinations of the $\tau_\pm$ together with τ_3 lead us to the spin–1/2 representation of our friend $SU(2)$. So, in a more abstract way, we can define nucleon states according to

$$|N\rangle = |I, I_3\rangle \ , \tag{160}$$

with the absolute value of the isospin and the value of its 3–component. Note, that this is in close analogy with the j and m in our treatment above. Hence,

$$|p\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle \ , \ \ |n\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \ . \tag{161}$$

3.3.2 Isospin as bookkeeping device

We would like to come back to the issue of keeping track of all possible replacements $n \to p$. The best way not to get lost in this debris, is to introduce so–called creation and annihilation operators, which link Hilbert spaces with different particle numbers. Now, let $P_\alpha^\dagger$ (P_α) be the operator, that creates (annihilates) a proton with quantum numbers α. But wait ! Since protons are fermions, they should better obey the Pauli exclusion principle. Thus, the operators $P_\alpha^\dagger$ and P_α have to behave as

$$P_\alpha^\dagger P_\alpha^\dagger = P_\alpha P_\alpha = 0 \, , \ \ P_\alpha^\dagger P_\alpha = 1 \, , \tag{162}$$

where the second relation holds, because the first term is clearly 1 and the second term vanishes when acting on a state including one proton in state α. In more compact form, we can write

$$\{P_\alpha^\dagger, P_\alpha^\dagger\} = \{P_\alpha, P_\alpha\} = 0 \, , \ \ \{P_\alpha^\dagger, P_\beta\} = \delta_{\alpha\beta} \, , \tag{163}$$

where the anticommutator $\{A, B\}$ is

$$\{A, B\} = AB + BA \, . \tag{164}$$

Of course, similar relations hold true for the neutrons, since they are fermions as well. Just replace all operators P_α with N_α. Additionally, we find, that

$$\{P_\alpha^\dagger(\mathrm{or} P_\alpha), N_\beta^\dagger(\mathrm{or} N_\beta)\} = 0 \, , \tag{165}$$

since protons and neutron are clearly different particles.
Let us call a state with no protons and no neutrons $|0\rangle$. Clearly, it is annihilated by the annihilation operators,

$$P_\alpha|0\rangle = N_\alpha|0\rangle = 0 \, . \tag{166}$$

Multi–proton states are then generated via

$$P^\dagger_{\alpha_1} P^\dagger_{\alpha_2} \dots P^\dagger_{\alpha_n} |0\rangle = |p_{\alpha_1} p_{\alpha_2} \dots p_{\alpha_n}\rangle \tag{167}$$

With help of this creation and annihilation operators we can build the Hamiltonian, interaction terms, like for example the mutual repulsion of the protons, would look like

$$H_{\text{inter}} = \sum_{\alpha,\beta} P^\dagger_\alpha P_\alpha V_{\alpha\beta} P^\dagger_\beta P_\beta \, . \tag{168}$$

The operator

$$\mathcal{N}^p_\alpha = \sum_\alpha P^\dagger_\alpha P_\alpha \tag{169}$$

clearly counts the numbers of protons in state α, in contrast the operators

$$P^\dagger_\alpha N_\alpha \text{ and } N^\dagger_\alpha P_\alpha \tag{170}$$

transform neutrons in state α into protons in the same state and vice versa. Writing the Hamiltonian as

$$\mathcal{H} = \mathcal{H}_s + \mathcal{H}_{em} + \mathcal{H}_w \tag{171}$$

and coming back to our formal statement, that the strong interactions contained in $\mathcal{H}_s$ do not feel the difference between protons and neutrons, we have

$$[\mathcal{H}_s, P^\dagger_\alpha N_\alpha] = [\mathcal{H}_s, N^\dagger_\alpha P_\alpha] = 0 \, . \tag{172}$$

Let us define now operators

$$T_+ = \frac{1}{\sqrt{2}} \sum_\alpha P^\dagger_\alpha N_\alpha \text{ and } T_- = \frac{1}{\sqrt{2}} \sum_\alpha N^\dagger_\alpha P_\alpha \, . \tag{173}$$

Obviously, they replace all neutrons by protons and vice versa. Coming back to our algebra we find immediately, that

$$\begin{aligned} [T_+, T_-] &= T_3 \\ &= \frac{1}{2} \sum_{\alpha,\beta} \Big[\Big(P^\dagger_\alpha N_\alpha N^\dagger_\beta P_\beta - N^\dagger_\beta P_\beta P^\dagger_\alpha N_\alpha \Big) \end{aligned}$$

$$
\begin{aligned}
&+ \quad \left(P^\dagger_\alpha N^\dagger_\beta N_\alpha P_\beta - N^\dagger_\beta P^\dagger_\alpha P_\beta N_\alpha\right)\Big] \\
&= \quad \frac{1}{2}\sum_{\alpha,\beta}\left(P^\dagger_\alpha P_\beta \delta_{\alpha\beta} - N^\dagger_\beta N_\alpha \delta_{\alpha\beta}\right) \\
&= \quad \frac{1}{2}\sum_{\alpha}\left(\mathcal{N}^p_\alpha - \mathcal{N}^n_\alpha\right) . \qquad (174)
\end{aligned}
$$

The second term in the second line of the equation above cancels out, since neutron and proton operators commute, and the two terms can be transformed into each other. In other words, the operator T_3 counts the number of protons, subtracts the number of neutrons and divides by two, i.e. it counts the total isospin of our state (note, that in our derivation, we have inserted one convenient 0). Use of the Jacobi identity yields

$$
[\mathcal{H}_s, T_3] = [\mathcal{H}_s, [T_+, T_-]] = 0 . \qquad (175)
$$

We note in passing, that we could exponentiate our generators here and produce a continuous $SU(2)$ symmetry group. But the group elements would in fact be useless, since they relate to unitary transformations on the two dimension space spanned by $|p\rangle$ and $|n\rangle$. But who might be interested in linear combinations of a proton and a neutron, which are clearly no physical states at all ?

Instead, we will go back to the eigenstates of our Hamiltonian, Eq.(171), and analyze them in perturbation theory. This is clearly justified, since the coupling constants of $\mathcal{H}_{em}$ and $\mathcal{H}_w$ are $\alpha \approx 1/137$ and $G_F \approx 10^{-5}/m_p^2$, respectively. Thus in zeroth order of perturbation theory, suitable eigenstates of $\mathcal{H}$ can be labeled by their isospin properties,

$$
|I, I_3, h_s, \alpha\rangle , \qquad (176)
$$

where I is the highest T_3 of the irreducible representation, I_3 is the eigenvalue of T_3, h_s is the eigenvalue of $\mathcal{H}_s$ and α labels all the other quantum numbers necessary. The states with identical I and h_s are said to form an isospin multiplet with isospin I. For example, the proton and the neutron lie in the same isospin multiplet with $I = 1/2$, similarly, H^3 and He^3 lie in the same multiplet, again with $I = 1/2$. Note, that since the other two parts of the Hamiltonian provide us with only small perturbations, the states in some multiplet are nearly degenerate, and the splitting between different multiplets are much larger than the splitting within one multiplet. Furthermore, we find, that isospin is conserved by the strong interactions. This becomes clear immediately, when we consider the time–evolution of some state. It is governed by $\exp(-i\mathcal{H}_s t)$, which commutes with all the T_a. Two more operators are related to conserved quantities, namely the charge

$$
Q = \sum_\alpha P^\dagger_\alpha P_\alpha \qquad (177)
$$

	mass [MeV=mc^2]	spin	I	I_3	lifetime
$\pi^\pm$	139.570	0	1	± 1	$2.6 \cdot 10^{-8}$s
π^0	134.976	0	1	0	$8.4 \cdot 10^{-17}$s

Table 2: Properties of the pion.

and the baryon number

$$B = \sum_\alpha \left(P_\alpha^\dagger P_\alpha + N_\alpha^\dagger N_\alpha \right) . \tag{178}$$

Obviously, the relation

$$Q = T_3 + \frac{1}{2} B \tag{179}$$

relates these three operators with each other.

3.3.3 πN–scattering

So far, we have not really benefitted from the introduction of isospin. At first glance, it seems, as if our developed formal apparatus just reformulated everything in a somewhat more formal, cumbersome way. However, things become interesting, when we go to higher energies, i.e. from the realm of nuclear physics (bound states and their properties) to the domain of particle physics (scattering of particles). The first thing we observe when colliding nucleons is the appearance of the so called pions. They are characterized by properties displayed in Table2. A few things are worth noting here.

- Particle physicists call pions stable, since they live a long time compared to the time it takes to produce them. This time can be estimated to be of the order of $10^{-24}s$, the time, light needs to cross a nucleon.
- The decays are caused by electromagnetism (π^0) and by the weak force ($\pi^\pm$). Pions decay predominantly into either 2 photons ($\pi^0 \to \gamma\gamma$) or into a muon and its neutrino ($\pi^- \to \mu^- \bar{\nu}_\mu$).
- The pions are not built from nucleons in any sensible way, instead you can thing of pions to bind nucleons together inside a nucleus.
- Experimentalists can make beams out of the charged pions and collide them with nucleons. Here, relativity comes to help.

The last point will keep us busy for the rest of this section. In principle, scattering experiments can be described theoretically in the following way :

The experimentalist prepares a pion and a nucleon, say a proton, far apart from each other, so that the particles can not interact with each other. This state can then be described via $|\pi p, i\rangle$, where i reminds us of the fact, that this is the $i$nitial state. Time evolution then proceeds via

$$\exp\left[-i\mathcal{H}(t-t_i)\right]|\pi p, i\rangle \tag{180}$$

and collisions are described by the "sandwich" with some $f$inal state vector. The so–called S–matrix is

$$S_{fi} = \langle \pi p, f|\exp\left[-i\mathcal{H}(t_f-t_i)\right]|\pi p, i\rangle\,. \tag{181}$$

In the examplatory case above, initial and final state coincide. The main point here, is that the probability to produce the final state f in a scattering element with given initial state is proportional to $|S_{fi}|^2$. Ignoring again weak and electromagnetic forces, the time evolution operator above is $\exp\left[-i\mathcal{H}_s t\right]$, which commutes with the T_a. Consequently, on our level of description, isospin is a conserved quantity. Stated in other words,

$$[S_{fi}, T_a] = 0 \tag{182}$$

for all possible processes on the level of nucleons and pions as initial and final states.
The processes that are observable on the level of $\pi N \to \pi' N'$ are:
$\pi^+ p \to \pi^+ p$, $\pi^+ n \to \pi^+ n$, $\pi^+ n \to \pi^0 p$, $\pi^- p \to \pi^- p$, and $\pi^- n \to \pi^- n$. Initial π^0 are quite impossible to prepare, and completely neutral states are in principle undetectable. The five processes listed above then are described by only two amplitudes, one for isospin 3/2 and one for isospin 1/2. The reason is, that we can understand the pion–nucleon states as direct products of $I = 1/2$ (for the nucleons) and $I = 1$ (for the pions) representations of the isospin group. In some analogy with the addition of two spin–1/2 particles, we can identify the states

$$|\pi^+ p\rangle = \left|\frac{3}{2}, \frac{3}{2}\right\rangle\ ,\quad |\pi^- n\rangle = \left|\frac{3}{2}, -\frac{3}{2}\right\rangle\ . \tag{183}$$

Applying now a the ladder operator T_- on the first of these states we find

$$T_-|\pi^+ p\rangle \quad = \quad \sqrt{\frac{2}{3}}|\pi^0 p\rangle + \sqrt{\frac{1}{3}}|\pi^+ n\rangle = \left|\frac{3}{2}, \frac{1}{2}\right\rangle\ . \tag{184}$$

The orthogonal combination of the two states on the right hand side must be in the spin 1/2 subspace, so we have

$$\sqrt{\frac{2}{3}}|\pi^+ n\rangle - \sqrt{\frac{1}{3}}|\pi^0 p\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle . \tag{185}$$

The procedure to construct the states and normalizations above will be put on more formal ground in a later section dealing with the formal addition of spins and the introduction of the Clebsch–Gordan coefficients, Sec.4.3.1. However, we can invert the relations above and find

$$\begin{aligned}
|\pi^0 p\rangle &= \sqrt{\frac{2}{3}}\left|\frac{3}{2},\frac{1}{2}\right\rangle - \sqrt{\frac{1}{3}}\left|\frac{1}{2},\frac{1}{2}\right\rangle \\
|\pi^+ n\rangle &= \sqrt{\frac{2}{3}}\left|\frac{1}{2},\frac{1}{2}\right\rangle + \sqrt{\frac{1}{3}}\left|\frac{3}{2},\frac{1}{2}\right\rangle .
\end{aligned} \tag{186}$$

Finally, let us come back to our scattering processes. By noting, that the isospin generators commute with the scattering matrix, we can write

$$\langle I, I_3, f|S_{fi}|I', I_3', i\rangle = \delta_{II'}\delta_{I_3 I_3'} a_I \tag{187}$$

and we find, that we are concerned only with two isospin amplitudes describing all five processes listed above. In terms of the states we have written in the (I, I_3) representation we have

$$\begin{aligned}
\langle \pi^+ p, f|S|\pi^+ p, i\rangle &= a_{\frac{3}{2}} , \\
\langle \pi^+ n, f|S|\pi^+ n, i\rangle &= \frac{2}{3}a_{\frac{1}{2}} + \frac{1}{3}a_{\frac{3}{2}} \\
\langle \pi^0 p, f|S|\pi^+ n, i\rangle &= \frac{\sqrt{2}}{3}\left(a_{\frac{3}{2}} - a_{\frac{1}{2}}\right) .
\end{aligned} \tag{188}$$

This has been checked experimentally with sufficient agreement of theory and experiment.

4 Angular momentum

4.1 $SO(3)$: The rotation group

4.1.1 From $O(3)$ to $SO(3)$

Let us start from the group $O(3)$, which is related to all possible orthogonal transformations within a three dimensional vector space with appropriate combination law for subsequent transformations. This group is called the rotation–inversion group and can be defined via 3×3–matrices R, which satisfy

$$RR^T = R^T R = \mathbf{1}\,. \tag{189}$$

Since

$$\det(R) = \det(R^T) \tag{190}$$

we find, that for this group

$$\det(R) = \pm 1\,. \tag{191}$$

Consequently, we can project this group onto two sets, those with determinant -1 related to space–inversions (via $-\mathbf{1}$), and the proper rotations with determinant 1. We are lead to write

$$O(3) = SO(3) \otimes (\mathbf{1},\, -\mathbf{1})\,. \tag{192}$$

Since the group $(\mathbf{1},\, -\mathbf{1})$ has only two one–dimensional irreducible representations, the representations of $O(3)$ can be easily obtained from those for $SO(3)$ via direct products.

4.1.2 Euler angles and angular momentum

As already indicated in the case of $SO(2)$, there are various options to construct a representation of a Lie group. For three–dimensional rotations, the standard method is via the so–called Euler–angles α, β and γ. With them, a general rotation R can be written as $R(\alpha, \beta, \gamma)$ defined in the following way:

1. Rotate first through α around the z–axis,
2. then by β around the new y'–axis,
3. and finally by γ around the z'-axis.

Therefore

$$R(\alpha,\beta,\gamma) =$$
$$\begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & \sin\beta\cos\gamma \\ -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & -\sin\beta\sin\gamma \\ -\cos\alpha\sin\beta & -\sin\alpha\sin\beta & -\cos\beta \end{pmatrix}, \tag{193}$$

which is orthogonal and has determinant $+1$. In analogy to the rotations in a plane, the generators for $SO(3)$ can be constructed with help of infinitesimal rotations with angel ϵ around some axis $\vec{h}$. These rotations are isomorphic to $SO(2)$ above, and we can conclude, that three suitable generators for $SO(3)$ are

$$I_{x,y,z} = -\frac{L_{x,y,z}}{\hbar}, \tag{194}$$

cf. Eq.(118). Of course, any rotation operator describing rotations of ϕ around $\vec{h}$ can then be written as

$$R_{\vec{H}}(\phi) = \exp\left[-i\phi\frac{\vec{L}\vec{h}}{\hbar}\right]. \tag{195}$$

When expanding the exponent, one should remember, that the individual components of $\vec{L}$ do not commute with each other. In fact, they obey the same commutation relations as $SU(2)$, Eq.(119). We will use this later on, but in the meantime we will try to construct irreducible representations of $SO(3)$ in a somewhat different approach.
As we have already seen, a good starting point are the characters of the irreducible representations. The main point here is to find a suitable basis – a set of functions – that transforms into some linear combinations under rotations. But we already know such a set, namely the spherical harmonics $Y_l^m(\theta,\phi)$ with some l out of $l = 0, 1, 2, \ldots$ and $-l \leq m \leq l$. If the transformation properties of these functions are known, we immediately have, of course, the matrices of the irreducible representations generated via the Y_l^m. But instead of this, let us dwell first on the class structure of $SO(3)$. As a first step, consider two rotations $R_{\vec{h}}(\phi)$ and $R_{\vec{g}}(\phi)$ with the same angle, but different axes. It should be clear, that we can find some third rotation, $R_{\vec{w}}(\omega)$ that connects these two rotations via $R_{\vec{h}}(\phi) = R^{-1}_{\vec{w}}(\omega)R_{\vec{g}}(\phi)R_{\vec{w}}(\omega)$, where $\vec{w}$ is orthogonal to both $\vec{h}$ and $\vec{g}$, and ω is the angle between $\vec{h}$ and $\vec{g}$. We can conclude, that in $SO(3)$ rotations with the same angle fall into the same class, thus the characters should not depend on the axes but on the angles, only. Consequently, we do not have to work out all the rotation properties of the

Y_l^m around any possible axis, instead we can focus on one axis only, for instance the z–axis. In this case we have

$$R_z(\zeta)Y_l^m(\theta,\phi) = Y_l^m(\theta,\phi-\zeta) = \exp(-im\zeta)Y_l^m(\theta,\phi)\,, \tag{196}$$

and the matrix–representation of $R_z(\zeta)$ is diagonal,

$$R_z(\zeta) = \begin{pmatrix} \exp(-il\zeta) & & & & \\ & \exp(-i(l-1)\zeta) & & \mathbf{0} & \\ & & \ddots & & \\ & \mathbf{0} & & \exp(i(l-1)\zeta) & \\ & & & & \exp(il\zeta) \end{pmatrix} \tag{197}$$

We can read off the corresponding character

$$\begin{aligned} \chi^{(l)}(\zeta) &= \sum_{m=-l}^{l} \exp(im\zeta) = e^{-il\alpha}\left(1 + e^{i\alpha} + \ldots + e^{2i\alpha}\right) \\ &= \frac{\exp[i(l+1/2)\alpha] - \exp[-i(l+1/2)\alpha]}{\exp[i/2\alpha] - \exp[-i/2\alpha]} = \frac{\sin\left[\left(l+\frac{1}{2}\right)\zeta\right]}{\sin\frac{\zeta}{2}}\,. \end{aligned} \tag{198}$$

Let us note, that

$$\chi^{(l)}(0) = 2l + 1 \tag{199}$$

and that we can now write

$$R(\alpha,\beta,\gamma)Y_l^m(\theta,\phi) = \sum_{m'=-l}^{l} D^{(l)}_{m'm}(\alpha,\beta,\gamma)Y_l^{m'}(\theta,\phi)\,, \tag{200}$$

which defines the $(2l+1)$–dimensional representation $D^{(l)}$. The functions $D^{(l)}$ are sometimes called the Wigner functions. Varying $l = 0, 1, \ldots$, these cover all the continuous and single–valued representations of $SO(3)$.

4.2 Representation of $SO(3)$ from $SU(2)$

It should not come as a very big surprise, that we are now in the position to bridge the final gap to some representation of $SO(3)$ by coming back to $SU(2)$. This is motivated by the fact, that the generators of $SO(3)$ and $SU(2)$ obey the same commutation rule. But to proceed here, we will go back to $SU(2)$ and reanalyze it in a somewhat different way than before.

4.2.1 Irreducible representations of $SU(2)$, once more ...

Let us start by constraining the 2×2 matrices forming $SU(2)$. Consider a vector $\vec{u} = (u_1, u_2)^T$ in the two–dimensional vector space over complex numbers and transform it according to

$$(u_1', u_2') = (u_1, u_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} . \tag{201}$$

The condition, that $|\vec{u}'| = |\vec{u}|$ should be conserved, i.e. a restriction on rotations only, leaves us with a unitary matrix. This can be cast into the form

$$\begin{pmatrix} \cos\theta e^{i\alpha} & \sin\theta e^{i\gamma} \\ -\sin\theta e^{i(\beta-\gamma)} & \cos\theta e^{i(\beta-\alpha)} \end{pmatrix} \tag{202}$$

with determinant $\exp(i\beta)$. We see, that we are left with four real angles. Constraining ourselves on matrices with determinant 1 we are left with the group $SU(2)$ and with matrices of the form

$$R(a,b) = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix} \text{ with } aa^* + bb^* = 1 , \tag{203}$$

which has only three free parameters. It is not very surprising, that we have three parameters only, one for each generator. Obviously, matrices of the form in Eq.(203) and Eq.(202) already form representations of $SU(2)$ and $U(2)$, respectively.

4.2.2 ... and of degree $2j$

But we can expand our analysis even further. Let us consider $2j+1$ symmetric products of the components u_1 and u_2 of our complex vector $\vec{u}$ (j is integer or half–integer)

$$f_j^m\Big|_{m=-j,-j+1,\dots,j} = \frac{u_1^{j+m} u_2^{j-m}}{\sqrt{(j+m)!(j-m)!}} . \tag{204}$$

Applying now the $SU(2)$ transformations of Eq.(203) on these functions f_j^m, it is clear, that they will be mapped on their linear combinations, since

$$\begin{aligned} R(a,b) f_j^m &= \frac{(au_1 + bu_2)^{j+m}(-b^* u_1 + a^* u_2)^{j-m}}{\sqrt{(j+m)!(j-m)!}} \\ &= \frac{1}{\sqrt{(j+m)!(j-m)!}} \end{aligned}$$

$$
\begin{aligned}
& \cdot \sum_{k,l} \left[\frac{(j+m)!\,(au_1)^{j+m-k}(bu_2)^k}{k!(j+m-k)!} \frac{(j-m)!\,(-b^* u_1)^{j-m-l}(a^* u_2)^l}{l!(j-m-l)!} \right] \\
= \; & \sum_{k,l} \left[\frac{\sqrt{(j+m)!(j-m)!}}{k!l!(j+m-k)!(j-m-l)!} \right. \\
& \left. \cdot\, a^{j+m-k}(a^*)^l(-b^*)^{j-m-l} b^k u_1^{2j-k-l} u_2^{k+l} \right] . \qquad (205)
\end{aligned}
$$

Note, that here and in the following the limits of the two summations over k and l result from the condition, that all arguments of the factorials in the denominator remain non–negative. Thus, $0 \leq k \leq j+m$ and $0 \leq l \leq j-m$. However, we are free to relabel,

$$
m' = j - k - l\,, \qquad (206)
$$

resulting in

$$
R(a,b) f_j^m = \sum_{m'=-j}^{j} D^{(j)}_{mm'}(a,b) f_j^{m'} \qquad (207)
$$

with

$$
\begin{aligned}
D^{(j)}_{mm'}(a,b) \;=\; & \sum_k \left[\frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j+m-k)!k!(j-m'-k)!(m'-m+k)!} \right. \\
& \left. \cdot\, a^{j+m-k}(a^*)^{j-m'-k} b^k (-b^*)^{m'-m+k} \right] .
\end{aligned}
\qquad (208)
$$

Here,

$$
\left.\begin{array}{ll} m-m' & \text{if } m'-m<0 \\ 0 & \text{else} \end{array}\right\} \leq k \leq \left\{\begin{array}{ll} j+m & \text{if } j+m<j-m \\ j-m & \text{else} \end{array}\right. . \qquad (209)
$$

Let us note, that

$$
\sum_{m=-j}^{j} \left| f_j^m \right|^2 = \sum_m \frac{\left| u_1^{j+m} u_2^{j-m} \right|^2}{(j+m)!(j-m)!} = \frac{(|u_1|^2 + |u_2|^2)^{2j}}{(2j)!}\,, \qquad (210)
$$

which is clearly invariant under $SU(2)$. This proves, that our functions f_j^m clearly constitute $2j+1$ independent functions in the $2j+1$–dimensional

Hilbert space related to the rotations, and that the corresponding representations $D^{(j)}_{mm'}$ are unitary as demanded. Furthermore, we can prove with help of Schur's first lemma, that they are in fact an irreducible representation. To do so, consider two special cases of the $D^{(j)}$,

1.
$$D^{(j)}_{mm'}\left(e^{i\alpha/2},0\right) = \delta_{mm'}e^{im\alpha}\,, \tag{211}$$
and

2.
$$D^{(j)}_{mj}\left(a,b\right) = \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}}\,a^{j+m}(-b^*)^{j-m}\,. \tag{212}$$

Any matrix P commuting with the special form of Eq.(211), which is diagonal, has to be diagonal itself. Considering the (j,m) element of $PD = DP$ and the special form of D in Eq.(212), yields, that P has to be an multiple of the unity matrix. It immediately follows, that the $D^{(j)}$ form an irreducible representation of dimension $2j+1$. Even more, their in-equivalence for different j gives, that $SU(2)$ has exactly one inequivalent representation for every integral order.

4.2.3 Characters of the $D^{(j)}$

We continue our treatment of the representations of $SU(2)$ by investigating the characters of the representations $D^{(j)}$. First, we will work out the class structure of $SU(2)$. This can be done by finding the eigenvalues of the matrices of Eq.(203). The characteristic equation reads

$$\lambda^2 - (a+a^*)\lambda + 1 = 0 \implies \lambda_{1,2} = \frac{\beta \pm \sqrt{\beta^2-4}}{2} \tag{213}$$

with $\beta = (a+a^*)$, which is a real number. The condition $aa^* + bb^* = |a|^2 + |b|^2 = 1$ constrains the real part of a as $-1 \leq Re(a) \leq 1$, ergo $-2 \leq \beta \leq 2$. Thus $\lambda_1 = \lambda_2^*$ and $\lambda_1\lambda_2 = 1$, and we find, that

$$|\lambda_1| = |\lambda_2| = 1\,. \tag{214}$$

Rewriting

$$\cos\frac{\alpha}{2} = \frac{\beta}{2} \tag{215}$$

we see, that

$$\lambda_{1,2} = e^{\pm i\alpha/2} \text{ with } \alpha \in [0, 2\pi[\,. \tag{216}$$

We can conclude, that all elements of $SU(2)$ having the same α, i.e. having the same real part of a, belong to the same class, since they have the same eigenvalue and are therefore conjugate to each other. So, every α labels one class of $SU(2)$. This leads us immediately to the characters of the classes in the irreducible representation $D^{(j)}$. Obviously, we can choose the simplest element of each class to determine the character, and we choose our friends of Eq.(211) above. Thus

$$\chi^{(j)}(e^{i\alpha/2}, 0) = \sum_{m=-j}^{j} D^{(j)}_{mm}(e^{i\alpha/2}, 0) = \sum_{m=-j}^{j} e^{im\alpha} = \frac{\sin[(2j+1)\alpha/2]}{\sin(\alpha/2)}\,. \tag{217}$$

It is definitely not a big surprise, that the characters of $SU(2)$, Eq.(217), equal the ones of $SO(3)$, Eq.(198), with the only difference, that for $SU(2)$, j can take half–integer values as well, whereas for $SO(3)$ j can take integer values, only.

4.2.4 Mapping $SU(2)$ on $SO(3)$

Let us consider functions f_1^m,

$$f_1^1 = \frac{u_1^2}{\sqrt{2}}\,, \quad f_1^0 = u_1 u_2\,, \quad f_1^{-1} = \frac{u_2^2}{\sqrt{2}}\,, \tag{218}$$

and their transformations under some general element $R(a,b)$ of $SU(2)$. We find, that

$$\begin{aligned}
\left(f_1^1\right)' &= R(a,b) f_1^1 = a^2 f_1^1 + ab f_1^0 + b^2 f_1^{-1} \\
\left(f_1^0\right)' &= R(a,b) f_1^0 = -2ab^* f_1^1 + (aa^* - bb^*) f_1^0 + 2a^* b f_1^{-1} \\
\left(f_1^{-1}\right)' &= R(a,b) f_1^{-1} = (b^*)^2 f_1^1 - a^* b^* f_1^0 + (a^*)^2 f_1^{-1}
\end{aligned} \tag{219}$$

and are lead to define new variables

$$x = f_1^1 - f_1^{-1}\,, \quad y = -i(f_1^1 + f_1^{-1})\,, \quad z = f_1^0\,. \tag{220}$$

In these variables, Eq.(219) can be rewritten in a fashion, that the connection between transformed (primed) and original (unprimed) coordinates is via real parameters only.

$$
\begin{aligned}
x' &= \frac{1}{2}(a^2 + a^{*2} - b^2 - b^{*2})x + \frac{i}{2}(a^2 - a^{*2} + b^2 - b^{*2})y + (ab + a^* b^*)z \\
y' &= -\frac{i}{2}(a^2 - a^{*2} - b^2 + b^{*2})x + \frac{1}{2}(a^2 + a^{*2} + b^2 + b^{*2})y - i(ab - a^* b^*)z \\
z' &= -(a^* b + ab^*)x + i(a^* b - ab^*)y + (aa^* - bb^*)z\,. \qquad (221)
\end{aligned}
$$

Let us now map the representation of $SO(3)$ by the Euler angles on some product of suitable rotations in $SU(2)$. To this end we start with our well known friend, the element $R(\exp(i\alpha/2), 0)$ of $SU(2)$ and plug it into Eq.(221). We find

$$
x' = x \cos\alpha - y \sin\alpha\,, \;\; y' = x \sin\alpha + y \cos\alpha\,, \;\; z' = z\,, \qquad (222)
$$

ergo

$$
R\left(e^{i\alpha/2}, 0\right) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \rightarrow \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (223)
$$

and we have mapped one rotation through Euler angle α on some element of $SU(2)$ labeled with half the angle α. Similarly,

$$
R\left(\cos\frac{\beta}{2}, \sin\frac{\beta}{2}\right) = \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \rightarrow \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}. \qquad (224)
$$

Coming back to our construction procedure related with the Euler angles as outlined above, we find, that Eq.(193) translated into representations of $SU(2)$ yields

$$
R(\alpha, \beta, \gamma) \leftarrow \begin{pmatrix} \cos\frac{\beta}{2} e^{i(\alpha+\gamma)/2} & -\sin\frac{\beta}{2} e^{-i(\alpha-\gamma)/2} \\ \sin\frac{\beta}{2} e^{i(\alpha-\gamma)/2} & \cos\frac{\beta}{2} e^{-i(\alpha+\gamma)/2} \end{pmatrix} \qquad (225)
$$

We note, that we do not have an $1 \rightarrow 1$ mapping of elements of $SO(3)$ on elements of $SU(2)$, but rather an $1 \rightarrow 2$ homomorphism. This can be understood trivially, by applying $R(0, 2\pi, 0)$ and $R(0, 0, 0)$, which clearly denote identical elements of $SO(3)$ in Eq.(225). We find, that they are to be associated with the matrices $\mathbf{1}$ and $-\mathbf{1}$ of $SU(2)$. Thus, under $SU(2)$ transitions, state vectors are mapped onto themselves only under 4π rotations. This means, when dealing with particles having spin 1/2 and some orbital angular momentum, the corresponding wave function is mapped onto itself only after 4π.

4.2.5 Finally, the representations

Now we are in the position to find quite easily representations of $SO(3)$ without to much effort. We just have to keep in mind, that $SU(2)$ is $2 \to 1$ holomorphic on $SO(3)$, start with some representation of $SU(2)$ and pick out the correct matrices. Coming back to Eq.(225), we can read off

$$\begin{aligned} D^{(j)}_{mm'}(\alpha,\beta,\gamma) &= D^{(j)}_{mm'}\left(\cos\frac{\beta}{2}e^{i(\alpha+\gamma)/2}, \sin\frac{\beta}{2}e^{i(\alpha-\gamma)/2}\right) \\ &= \sum_k \left[(-1)^{m'-m+k}\frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j+m-k)!k!(j-m'-k)!(m'-m+k)!}\right. \\ &\qquad \left. \cdot\, e^{im\alpha}e^{im'\gamma}\left(\cos\frac{\beta}{2}\right)^{2j+m-m'-2k}\left(\sin\frac{\beta}{2}\right)^{m'-m+2k}\right] . \end{aligned} \tag{226}$$

Again, the characters of this representation can be obtained via our special form with $\beta = \gamma = 0$,

$$\chi^{(j)}(\alpha) = \frac{\sin[(2j+1)\alpha/2]}{\sin(\alpha/2)} , \tag{227}$$

which nicely coincide with what we found before, Eq.(198).

4.2.6 Connection with spherical harmonics (Y_l^m)

Before closing our discussion of the representation of $SO(3)$, let us briefly dwell on the link of our rotation matrices and the spherical harmonics Y_l^m. As is well known, they appear in problems related to single particles in spherical symmetric potentials. For instance, the wave function for a spinless particle in such a potential is known to be separable in spherical coordinates, leading to energy eigenfunctions

$$\Psi_{nlm}(\vec{x}) = \langle \vec{x}\,|l,n,m\rangle = R_{nl}(r)Y_l^m(\theta,\phi) , \tag{228}$$

where n labels the energy eigenvalue and l and m are the eigenvalues of L^2 and L_z, respectively. Note, that here the eigenvalue x of the position eigenvector $|\vec{x}\rangle$ is specified by r and the polar and azimuth angle ϕ and θ. In other words, we can rewrite

$$\left\langle \frac{\vec{x}}{|\vec{x}|}\right| l,m \left.\vphantom{\frac{\vec{x}}{|\vec{x}|}}\right\rangle = Y_l^m(\theta,\phi) . \tag{229}$$

Let us now apply some rotation matrix $D(\alpha,\beta,\gamma)$ on a vector of unit length in z–direction, $|\vec{n}_z\rangle$. We find

$$|\vec{n}\,\rangle = D(\alpha, \beta, \gamma)|\vec{n}_z\,\rangle\,. \tag{230}$$

Specifying $D = D(\alpha = \phi, \beta = \theta, \gamma = 0)$, inserting a suitable **1** and multiplying from the left with $\langle l'm'|$ we can reformulate this as

$$\langle l'm'|\vec{n}\,\rangle = \sum_{l,m}\langle l'm'|D(\phi, \theta, 0)|lm\rangle\langle lm|\vec{n}_z\,\rangle\,. \tag{231}$$

Now,

$$\langle lm|\vec{n}_z\,\rangle = Y_l^{m^*}(\theta = 0, \phi) = \sqrt{\frac{2l+1}{4\pi}}\,\delta_{m0}\,, \tag{232}$$

which can be seen by noting that $|\vec{n}_z\,\rangle$ is an eigenvector of L_z with eigenvalue 0. Returning to Eq.(231), we remember, that

$$\langle l'm'|D(\theta, \phi, 0)|lm\rangle = D^{(l')}_{mm'}(\phi, \theta, 0)\delta_{ll'} \tag{233}$$

and we are left with

$$Y_l^{m'^*}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}}\,D^{(l)}_{0m'}(\phi, \theta, 0) \tag{234}$$

yielding

$$D^{(l)}_{m0}(\phi, \theta, 0) = \sqrt{\frac{4\pi}{2l+1}}Y_l^{m^*}(\theta, \phi)\,. \tag{235}$$

4.3 Addition of angular momentum

Having constructed the representations for $SO(3)$, which are generated by the angular momentum, we will continue now in the spirit of our treatment of $SU(2)$ in Sec.3.2.2 and couple angular momenta, i.e. add them.

4.3.1 Introducing Clebsch–Gordan coefficients

Let us consider two arbitrary angular momentum operators $\vec{J}_1$ and $\vec{J}_2$. Clearly, their components satisfy the usual commutation relations, Eq.(119),

$$[J_{1i}, J_{1j}] = i\epsilon_{ijk}J_{1k}\,,\ \ [J_{2i}, J_{2j}] = i\epsilon_{ijk}J_{2k}$$

and they mutually commute,

$$[J_{1i}, J_{2j}] = 0\,, \tag{236}$$

since they act on different subspaces. The infinitesimal rotation operator is

$$1 - i(\vec{J}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \vec{J}_2) \cdot \vec{n}\delta\phi = (1 - i\vec{J}_1\vec{n}\delta\phi) \otimes (1 - i\vec{J}_2\vec{n}\delta\phi) \tag{237}$$

leading to the total angular momentum operator

$$\vec{J} = \vec{J}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \vec{J}_2 = \vec{J}_1 + \vec{J}_2\,. \tag{238}$$

Obviously, finite rotations are represented by

$$\exp\left(-i\vec{J}_1\vec{n}\phi\right) \otimes \exp\left(-i\vec{J}_2\vec{n}\phi\right)\,. \tag{239}$$

Note that this holds true only for rotations around the same axes with the same angle. This assignment is quite reasonable, one just has to understand $\vec{J}$ acts as an angular momentum operator of the full system, which is rotated as a whole. Again, we have two options to define a set of base vectors for the full system, namely

1. simultaneous eigenvectors of J_1^2, J_2^2, J_{1z} and J_{2z}, which we denote by $|j_1 j_2 m_1 m_2\rangle$. This is analogous to the $\{m_1, m_2\}$–basis of our $SU(2)$ example. Since the four operators commute mutually, this choice is sensible and we have the following eigenvalue assignments

$$\begin{aligned}
J_1^2|j_1 j_2 m_1 m_2\rangle &= j_1(j_1+1)|j_1 j_2 m_1 m_2\rangle \\
J_2^2|j_1 j_2 m_1 m_2\rangle &= j_2(j_2+1)|j_1 j_2 m_1 m_2\rangle \\
J_{1z}|j_1 j_2 m_1 m_2\rangle &= m_1|j_1 j_2 m_1 m_2\rangle \\
J_{2z}|j_1 j_2 m_1 m_2\rangle &= m_2|j_1 j_2 m_1 m_2\rangle\,.
\end{aligned} \tag{240}$$

   Alternatively, we might prefer

2. simultaneous eigenvectors of J^2, J_1^2, J_2^2 and J_z, which we denote by $|j j_1 j_2 m\rangle$. This is in analogy to the $\{s, m\}$ basis outlined in the previous section. Again, the operators commute mutually, see

$$[J^2, J_1^2] = [J_1^2 + J_2^2 + 2J_{1z}J_{2z} + J_{1+}J_{2-} + J_{1-}J_{2+}, J_1^2] = 0\,. \tag{241}$$

   Then we have

$$\begin{aligned} J^2|jj_1j_2m\rangle &= j(j+1)|jj_1j_2m\rangle \\ J_1^2|jj_1j_2m\rangle &= j_1(j_1+1)|jj_1j_2m\rangle \\ J_2^2|jj_1j_2m\rangle &= j_2(j_2+1)|jj_1j_2m\rangle \\ J_z|jj_1j_2m\rangle &= m|jj_1j_2m\rangle\,. \end{aligned} \tag{242}$$

There's one caveat left here, namely, that J^2 does not commute with individual components of the $\vec{J}_{1,2}$. Therefore, we are neither allowed to add J^2 to the set of operators in the $\{m_1m_2\}$ representation nor are we allowed to include for instance J_{1z} to the set related to the $\{s,m\}$ representation.

Of course, since these two possible representations are equivalent, they can be connected via a unitary transformation. In a suggestive manner we are thus allowed to write

$$|jm\rangle = \sum_{m_1,m_2} |m_1m_2\rangle\langle m_1m_2|jm\rangle\,, \tag{243}$$

where we have suppressed the obsolete indices $j_{1,2}$. The elements of the transformation matrix $\langle m_1m_2|jm\rangle$ are called Clebsch–Gordan coefficients. They have quite a lot of properties, which we will list below.

- They vanish, unless

$$m = m_1 + m_2\,. \tag{244}$$

To prove it, use $J = J_1 + J_2$ and, accordingly, $(J - J_1 - J_2)|\alpha\rangle = 0$ for all vectors $|\alpha\rangle$, including $|jm\rangle$. We multiply from the left with $\langle m_1m_2|$ and find

$$(m - m_1 - m_2)\,\langle m_1m_2|jm\rangle = 0\,. \tag{245}$$

- They vanish, unless

$$|j_1 - j_2| \le j \le j_1 + j_2\,. \tag{246}$$

We check this by counting the number of states in both representations. Clearly, in the $|m_1m_2\rangle$ base we have $N_{m_1m_2} = (2j_1+1)(2j_2+1)$ states. In the $|jm>$ base we find, according to the boundaries given above and assuming $j_1 \ge j_2$

$$N_{jm} = \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1) = N_{m_1 m_2}\,. \qquad (247)$$

This motivates the condition above. A more rigorous proof is left for the interested reader.

- They form a unitary matrix and can be taken real (this is a question of convention, though). Thus

$$\langle m_1 m_2 | jm\rangle = \langle jm | m_1 m_2\rangle\,. \qquad (248)$$

Since real unitary matrices are orthogonal, we have the following orthogonality conditions

$$\begin{aligned}\sum_{j,m} \langle m_1 m_2 | jm\rangle\langle jm | m_1' m_2'\rangle &= \delta_{m_1 m_1'}\delta_{m_2 m_2'}\,,\\ \sum_{m_1,m_2} \langle j' m' | m_1 m_2\rangle\langle m_1 m_2 | jm\rangle &= \delta_{jj'}\delta_{mm'}\,.\end{aligned} \qquad (249)$$

Normalization condition for the states can be obtained out of the last equation by setting $j = j'$, $m = m'$.

4.3.2 Recursion relation for Clebsch–Gordan coefficients

Furthermore, the Clebsch–Gordan coefficients satisfy recursion relations, which will be derived now. Let us first fix j_1, j_2 and j. Now we apply ladder operators $J_\pm$ on vectors $|jm\rangle$,

$$J_\pm |jm\rangle = (J_{1\pm} + J_{2\pm}) \sum_{m_1,m_2} |m_1 m_2\rangle\langle m_1 m_2 | jm\rangle\,. \qquad (250)$$

Inserting Eq.(134), we find ($m_1 \to m_1'$,$m_2 \to m_2'$)

$$\begin{aligned}&\sqrt{(j\mp m)(j\pm m+1)}|jm\pm 1\rangle = \\ &\quad\sum_{m_1',m_2'} \langle m_1' m_2' | jm\rangle \Big[\sqrt{(j_1\mp m_1')(j_1\pm m_1'+1)}|m_1'\pm 1 m_2'\rangle \\ &\qquad + \sqrt{(j_2\mp m_2')(j_2\pm m_2'+1)}|m_1' m_2'\pm 1\rangle\Big]\,.\end{aligned} \qquad (251)$$

Multiplying from the left with $\langle m_1 m_2|$ and using the orthogonality of states we find, that

$$\begin{aligned} m_1 &= m_1' \pm 1\,, \;\; m_2 = m_2' \quad \text{for the 1st term} \\ m_2 &= m_2' \pm 1\,, \;\; m_1 = m_1' \quad \text{for the 2nd term}\,. \end{aligned} \tag{252}$$

This leaves us with the desired recursion relations

$$\begin{aligned} &\sqrt{(j \mp m)(j \pm m + 1)}\langle m_1 m_2 | jm \mp 1\rangle = \\ &\quad \Big[\;\sqrt{(j_1 \mp m_1 + 1)(j_1 \pm m_1)}\langle m_1 \pm 1 m_2 | jm\rangle \\ &\quad + \sqrt{(j_2 \mp m_2 + 1)(j_2 \pm m_2)}\langle m_1 m_2 \mp 1 | jm\rangle \Big]\,. \end{aligned} \tag{253}$$

We note, that the non–vanishing condition for the Clebsch–Gordan coefficients in the m–values is now replaced by

$$m_1 + m_2 = m \pm 1\,. \tag{254}$$

Let us elucidate the impact of these recursion relations in a physically significant example. Couple some integer spin (or angular momentum) j and an half–integer spin $s = 1/2$,

$$\begin{aligned} j_1 &= l \in \mathrm{N}_0 \quad & m_1 &\in [-l, l] \\ s &= \tfrac{1}{2} & m_s &= \pm\tfrac{1}{2}\,. \end{aligned}$$

The values for j, m are given by $j = l \pm 1/2$. Thus, for $j = l = 1$ we get two different allowed values for j, namely $j = 1/2$ and $j = 3/2$, which correspond to $p_{1/2}$ and $p_{3/2}$ states in spectroscopic notation. In fact, we have seen this case at work already in our treatment of the isospin scattering amplitudes. However, in the following we will work out some more of the properties of these $l \pm 1/2$–states. The $m_l m_s$ plane for these states is particularly simple. Basically it consists of two rows with $m_s = \pm 1/2$ of points with integer m_l. Let's concentrate on the case $j = l + 1/2$ first. Since m_s is constrained by $m_s = 1/2$, we can use the J_- part of the recursion relation (253) and obtain ($m_1 = m_l = m - 1/2$, $m_2 = m_s = 1/2$)

$$\begin{aligned} &\sqrt{\left(l + \frac{1}{2} + m + 1\right)\left(l + \frac{1}{2} - m\right)}\left\langle m - \frac{1}{2}\frac{1}{2}\right| \left. l + \frac{1}{2} m \right\rangle = \\ &\quad \sqrt{\left(l + \frac{1}{2} + m\right)\left(l + \frac{1}{2} - m\right)}\left\langle m + \frac{1}{2}\frac{1}{2}\right| \left. l + \frac{1}{2} m + 1 \right\rangle\,. \end{aligned} \tag{255}$$

This shows, how we can move horizontally by one unit,

$$\left\langle m - \frac{1}{2}\frac{1}{2} \middle| \, l + \frac{1}{2} m \right\rangle = \sqrt{\frac{l+m+\frac{1}{2}}{l+m+\frac{3}{2}}} \left\langle m + \frac{1}{2}\frac{1}{2} \middle| \, l + \frac{1}{2} m + 1 \right\rangle , \tag{256}$$

a procedure, we can clearly repeat, until we reach the maximal possible value for m_l. Then we have

$$\left\langle m - \frac{1}{2}\frac{1}{2} \middle| \, l + \frac{1}{2} m \right\rangle = \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} \left\langle l \frac{1}{2} \middle| \, l + \frac{1}{2} l + \frac{1}{2} \right\rangle . \tag{257}$$

Choosing now

$$\left\langle l \frac{1}{2} \middle| \, l + \frac{1}{2} l + \frac{1}{2} \right\rangle = 1 \tag{258}$$

for the Clebsch–Gordan coefficient of the maximal total angular momentum (this is clearly a question of convention !), we have

$$\left\langle m - \frac{1}{2}\frac{1}{2} \middle| \, l + \frac{1}{2} m \right\rangle = \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} . \tag{259}$$

But this is not the full story yet. We still have to find the phases related to the non–maximal angular momentum states,

$$\begin{aligned}
\left| j = l + \frac{1}{2}, m \right\rangle &= \sqrt{\frac{l+\frac{1}{2}+m}{2l+1}} \left| m_l = m - \frac{1}{2}, m_s = \frac{1}{2} \right\rangle \\
&\quad + N_1 \left| m_l = m + \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle \\
\left| j = l - \frac{1}{2}, m \right\rangle &= N_2 \left| m_l = m - \frac{1}{2}, m_s = \frac{1}{2} \right\rangle \\
&\quad + N_3 \left| m_l = m + \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle .
\end{aligned} \tag{260}$$

At a first glance this does not look particularly encouraging. However, due to orthogonality of the base vectors, we expect, that fixing m the transformation matrix from the $|m_l m_s\rangle$ to the $|jm\rangle$ representation is orthogonal, i.e. of the form

$$\begin{pmatrix} \cos\eta & \sin\eta \\ -\sin\eta & \cos\eta \end{pmatrix} \tag{261}$$

and we can just read of $\cos\eta$ from Eq.(259). So, up to some sign ambiguity, we find

$$\sin\eta = \pm\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}\,. \tag{262}$$

With the claim, that the states $|l+1/2m\rangle$ are reachable by successive application of J_- we end up with

$$\left\langle m_l = m+\frac{1}{2} m_s = -\frac{1}{2}\right| \left. j = l+\frac{1}{2} m\right\rangle > 0 \tag{263}$$

and therefore the transformation matrix, Eq.(261), becomes

$$\begin{pmatrix} \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} \\ -\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} \end{pmatrix}\,. \tag{264}$$

Finally, we can construct spin–angular functions in two component form (one for each spin–sign) as

$$\begin{aligned} \mathcal{Y}_l^{j=l\pm\frac{1}{2},m} &= \pm\sqrt{\frac{l+\frac{1}{2}\pm m}{2l+1}} Y_l^{m-\frac{1}{2}}(\theta,\phi)\chi_+ \\ &\quad +\sqrt{\frac{l+\frac{1}{2}\mp m}{2l+1}} Y_l^{m+\frac{1}{2}}(\theta,\phi)\chi_- \\ &= \frac{1}{\sqrt{2l+1}} \begin{pmatrix} \pm\sqrt{l+\frac{1}{2}\pm m}\,Y_l^{m-\frac{1}{2}}(\theta,\phi) \\ +\sqrt{l+\frac{1}{2}\mp m}\,Y_l^{m+\frac{1}{2}}(\theta,\phi) \end{pmatrix}\,. \end{aligned} \tag{265}$$

They are simultaneous eigenfunctions of L^2, S^2 and J^2 and also of the operator $L\cdot S$, which depends on the other three.

4.4 The Wigner–Eckart theorem

In the last part of this section, we will introduce the concept of tensor operators and dwell on the issue of how to determine their matrix elements, leading us to the celebrated Wigner–Eckart theorem.

4.4.1 Tensor operators

In quantum mechanics, it is customary to employ notations as $\vec{p}$ or $\vec{L}$ for the momentum or angular momentum operators. Intuitively, the meaning of such

a notation is pretty transparent. However, in the following we will put this intuitive understanding on a somewhat more formal ground.
In general, tensor operators (including vector operators as the examples above) can be characterized with their rotational properties. More specifically, consider a set $\mathcal{O}_{x=1\ldots 2s+1}$ of $2s+1$ linear operators, which obey simple commutation relations with the angular momentum operators J,

$$[J_a, \mathcal{O}_x] = \mathcal{O}_y \left(J_a^s\right)_{yx} , \tag{266}$$

where a sum over y is understood, and $(J)_{yx}$ denotes one element of the matrix generators J. Here, the J_a^s are the generators of the angular momentum operator in the irreducible spin–s representation, and consequently, the $\mathcal{O}_x$ are said to transform according to this representation.
We can rephrase the statement above by employing our understanding of rotations. Let $R(\vec{\phi})$ denote a rotation of coordinates, where $\vec{\phi}$ is parallel to the axis of rotation and its length denotes the magnitude of the rotation angle. We introduce $U_R(\vec{\phi})$, which is the operator corresponding to this rotation and acts on other operators, functions, and the like. If $\mathcal{O}$ and $\mathcal{O}'$ denote operators related by U_R we have

$$\mathcal{O}' = U_R^\dagger(\vec{\phi})\,\mathcal{O}\,U_R(\vec{\phi})\,. \tag{267}$$

Simple inspection (remember, $\vec{J}$ is the generator of rotations) tells us, that

$$U_R(\vec{\phi}) = \exp\left(-\frac{i\vec{\phi}\cdot\vec{J}}{\hbar}\right) . \tag{268}$$

Then, to first order in ϕ, suitable for infinitesimal rotations, Eq.267 becomes

$$\begin{aligned} \mathcal{O}' &\approx \left(1+\frac{i\vec{\phi}\cdot\vec{J}}{\hbar}\right)\mathcal{O}\left(1-\frac{i\vec{\phi}\cdot\vec{J}}{\hbar}\right) \\ &\approx \mathcal{O}+\frac{i}{\hbar}\,[\vec{\phi}\cdot\vec{J},\mathcal{O}]\,. \end{aligned} \tag{269}$$

As an example, consider a particle in a spherically symmetric potential. Its position operator $\vec{x}$ is a tensor operator – in fact a vector operator – because

$$[J_a, x_b] = \epsilon_{acd}\,[x_c p_d, x_b] = -\epsilon_{acb} x_c = x_c \left(J_a^1\right)_{cb} . \tag{270}$$

Here we have identified the matrix element cb of the angular momentum operator directly.
This basic example can be generalized easily to more complicated cases, and we will discuss scalar, vector and tensor operators in the following.

1. Scalar operators :
   We define, that a quantity is *scalar*, if its is invariant under all rotations. Obvious examples are mass, energy and the like. Similarly, *scalar operators* are invariant under rotations. For example, the Hamiltonian of an isotropic system like a particle in a spherically symmetric potential, is a scalar operator. Another example is the Hamiltonian of an atom, and it is easy to prove, that also the operator $\vec{L} \cdot \vec{S}$ of the spin–orbit interaction is scalar. In group theoretical terms, we might say, that a scalar operator generates the irreducible representation $D^{(0)}$ of the rotation group. It is then easy to understand, that all matrix elements of such a scalar operator between states belonging to different irreducible representations (different values of j) or to different columns of the same irreducible representation (different values of m) of the rotation group $SO(3)$ must vanish. Moreover, it can be shown, that the matrix element of a scalar operator and two states belonging to the same $\{j, m\}$–pair is independent of m [1]. Taking everything together, we find, that any scalar operator $\mathcal{S}$ commutes with the angular momentum operator in order to guarantee invariance under rotations, i.e.

$$[J, \mathcal{S}] = 0\,. \tag{271}$$

   We are then in the position to define the *reduced matrix element*

$$\mathcal{S}^{(j)}_{NN'} = \langle Njm|\mathcal{S}|N'jm\rangle\,. \tag{272}$$

2. Vector operators :
   We proceed with vector operators. Again it is paramount to examine what we understand by a vector. Apart from having a magnitude – like the scalars discussed above – a vector has a direction, and it is defined by its transformation properties under rotations. We define $\vec{r}$ to be a vector transforming into $\vec{r}'$ under infinitesimal rotations characterized by $\vec{\phi}$,

$$\vec{r}' = \vec{r} + \vec{\phi} \times \vec{r}\,. \tag{273}$$

   Similarly, we define a *vector operator* by its properties under infinitesimal rotations. An operator $\vec{\mathcal{V}}$ with three (Cartesian) components is a vector operator, if it transforms like

$$U_R(\vec{\phi})^\dagger\, \vec{\mathcal{V}}\, U_R(\vec{\phi}) = \vec{\mathcal{V}} + \vec{\phi} \times \vec{\mathcal{V}} \tag{274}$$

[1] A proof of this property can be found in the appendix

under rotations realized by $U_R(\vec{\phi})$. Comparison with Eq.269 tells us, that vector operators satisfy

$$\frac{i}{\hbar}[\vec{\phi}\cdot\vec{J},\vec{\mathcal{V}}] = \vec{\phi}\times\vec{\mathcal{V}}\,. \tag{275}$$

We are now in the position to check immediately, that position, linear momentum, orbital angular momentum, spin angular momentum, and so on and so forth are in fact vector operators. Thus, there is no need to give our intuition a bad name.

In light of our introduction of tensor operators in this section, it is worthwhile, to consider two more formulations of the commutator algebra above. Let us start with what is known as the spherical components of vector operators,

$$\mathcal{V}_{\pm 1} = -\frac{1}{\sqrt{2}}(\mathcal{V}_x \pm i\mathcal{V}_y)\ ,\quad \mathcal{V}_0 = \mathcal{V}_z\,. \tag{276}$$

In this base, the commutation relations read

$$\begin{aligned}[][J_\pm,\mathcal{V}_q] &= \sqrt{2-q(q+1)}\,\hbar\,\mathcal{V}_{q\pm 1}\ ,\\ [J_z,\mathcal{V}_q] &= q\hbar\,\mathcal{V}_q\,.\end{aligned} \tag{277}$$

At least at this stage it becomes clear, that vector operators generate a representation $D^{(1)}$ of $SO(3)$.

Alternatively to our approach outlined above, we can use our knowledge about the representations of $SO(3)$. As we already learned, a vector $\vec{r}$ transforms as

$$R(\alpha,\beta,\gamma)\,r_m \equiv r'_m = \sum_{m'=-1}^{1} D^{(1)}_{m'm}(\alpha,\beta,\gamma) r_{m'}\,. \tag{278}$$

A similar relation then holds true for vector operators, with $R(\alpha,\beta,\gamma)$ replace by the operators U_R and $U_R^\dagger$. To be specific, we give the form of $D^{(1)}$ explicitly,

$$D^{(1)}(\alpha,\beta,\gamma) = \begin{pmatrix} \cos^2\frac{\beta}{2}e^{i(\alpha+\gamma)} & -\frac{1}{\sqrt{2}}\sin\beta e^{i\gamma} & \sin^2\frac{\beta}{2}e^{i(\gamma-\alpha)} \\ \frac{1}{\sqrt{2}}\sin\beta e^{i\alpha} & \cos\beta & -\frac{1}{\sqrt{2}}\sin\beta e^{-i\alpha} \\ \sin^2\frac{\beta}{2}e^{i(\alpha-\gamma)} & \frac{1}{\sqrt{2}}\sin\beta e^{-i\gamma} & \cos^2\frac{\beta}{2}e^{-i(\alpha+\gamma)} \end{pmatrix}. \tag{279}$$

However, a last comment on vector operators is in order here. Under pure rotations as the ones above, both polar and axial vector operators transform identically. But they change sign under space inversion. In fact polar and axial vector operators are defined via their transformation properties under space inversion U_J, namely

$$U_J^\dagger \vec{\mathcal{V}} U_J = \begin{cases} -\vec{\mathcal{V}} & \text{for polar vectors} \\ +\vec{\mathcal{V}} & \text{for axial vectors} \end{cases} . \tag{280}$$

Examples of polar vector operators are the position, linear momentum and the electric dipole moment operators, axial vectors are for instance the angular momentum and the magnetic dipole operators.

3. Tensor operators :
Tensors of rank two can be constructed from two vectors, $\vec{u}$ and $\vec{v}$, by

$$\mathcal{T}_{mn} = u_m v_n \text{ with } m, n = -1, 0, 1 . \tag{281}$$

These nine quantities transform according to

$$R(\alpha,\beta,\gamma)\mathcal{T}_{mn} = u'_m v'_n = \sum_{p,q=-1}^{1} D^{(1)}_{mp}(\alpha,\beta,\gamma) u_p D^{(1)}_{nq}(\alpha,\beta,\gamma) v_q \tag{282}$$

under rotations $R(\alpha,\beta,\gamma)$. In full analogy, the nine operators $\mathcal{U}_m \mathcal{V}_n$ constructed from the components of vector operators $\vec{\mathcal{U}}$ and $\vec{\mathcal{V}}$ constitute the components of a tensor operator $\mathcal{U}\mathcal{V}$ of rank two, if the transformed components

$$\mathcal{U}'_m \mathcal{V}'_n = U_R^\dagger \mathcal{U}_m \mathcal{V}_n U_R \tag{283}$$

can be expressed by an equation similar to Eq.282. The product $\mathcal{U}\mathcal{V}$ is dubbed the Kronecker or direct product of the two vector operators. It does not come as a surprise, that such tensor operators generate a nine–dimensional representation $D^{(1)} \otimes D^{(1)}$ of $SO(3)$. In fact, this representation can be reduced to $D^{(0)} \oplus D^{(1)} \oplus D^{(2)}$. It follows, that we are capable to construct suitable linear combinations of the components of $\mathcal{U}\mathcal{V}$, such that

- one of these is invariant under all rotations,
- three of them generate $D^{(1)}$,

- and the remaining five combinations generate $D^{(2)}$.

Let's start with the invariant linear combination. This is obviously the scalar product of the two vector operators, i.e. the trace of the tensor

$$\mathcal{T}_0^0 = \vec{\mathcal{U}} \cdot \vec{\mathcal{V}} = \mathcal{U}_x \mathcal{V}_x + \mathcal{U}_y \mathcal{V}_y + \mathcal{U}_z \mathcal{V}_z \equiv \sum_{m=-1}^{m} (-1)^m \mathcal{U}_m \mathcal{V}_{-m} \, . \tag{284}$$

We continue with the three combinations generating $D^{(1)}$. It can be shown, that they are given by the vector product

$$\mathcal{T}^{(1)} = \vec{\mathcal{U}} \times \vec{\mathcal{V}}. \tag{285}$$

The Cartesian components are

$$\begin{pmatrix} \mathcal{T}_x^{(1)} \\ \mathcal{T}_y^{(1)} \\ \mathcal{T}_z^{(1)} \end{pmatrix} = \begin{pmatrix} \mathcal{U}_y \mathcal{V}_z - \mathcal{U}_z \mathcal{V}_y \\ \mathcal{U}_z \mathcal{V}_x - \mathcal{U}_x \mathcal{V}_z \\ \mathcal{U}_x \mathcal{V}_y - \mathcal{U}_y \mathcal{V}_x \end{pmatrix} \tag{286}$$

and it is quite easy to check, that they transform like x, y, and z under rotations. In terms of spherical components we have

$$\mathcal{T}_{\pm 1}^{(1)} = \mathcal{U}_{\pm 1} \mathcal{V}_0 - \mathcal{U}_0 \mathcal{V}_{\pm 1} \, , \;\; \mathcal{T}_{\pm 0}^{(1)} = \mathcal{U}_1 \mathcal{V}_{-1} - \mathcal{U}_{-1} \mathcal{V}_1 \, , \tag{287}$$

which transforms like $x \pm iy$ and z or like the spherical harmonics of degree one, respectively.

Up to constant factors, the Cartesian components generating $D^{(2)}$ are

$$\begin{aligned} \mathcal{T}_{xy}^{(2)} &\sim \mathcal{U}_x \mathcal{V}_y + \mathcal{U}_y \mathcal{V}_x \\ \mathcal{T}_{yz}^{(2)} &\sim \mathcal{U}_y \mathcal{V}_z + \mathcal{U}_z \mathcal{V}_y \\ \mathcal{T}_{zx}^{(2)} &\sim \mathcal{U}_z \mathcal{V}_x + \mathcal{U}_x \mathcal{V}_z \\ \mathcal{T}_{x^2-y^2}^{(2)} &\sim \mathcal{U}_x \mathcal{V}_x - \mathcal{U}_y \mathcal{V}_y \\ \mathcal{T}_{2z^2-x^2-y^2}^{(2)} &\sim 2\mathcal{U}_z \mathcal{V}_z - \mathcal{U}_x \mathcal{V}_x - \mathcal{U}_y \mathcal{V}_y \end{aligned} \tag{288}$$

and in spherical components we find

$$\begin{aligned}
\mathcal{T}^{(2)}_{\pm 2} &\sim \mathcal{U}_{\pm 1}\mathcal{V}_{\pm 1} \propto Y_2^{\pm 2}(\theta,\phi) \\
\mathcal{T}^{(2)}_{\pm 1} &\sim \mathcal{U}_{\pm 1}\mathcal{V}_{0} + \mathcal{U}_{0}\mathcal{V}_{\pm 1} \propto Y_2^{\pm 1}(\theta,\phi) \\
\mathcal{T}^{(2)}_{0} &\sim 2\mathcal{U}_{0}\mathcal{V}_{0} + \mathcal{U}_{1}\mathcal{V}_{-1} + \mathcal{U}_{-1}\mathcal{V}_{1} \propto Y_2^{0}(\theta,\phi)\,.
\end{aligned} \tag{289}$$

Again we denoted the transformation behavior, which goes along the lines of the spherical harmonics.

Taken together, we can write

$$\begin{aligned}
\mathcal{U}_m\mathcal{V}_n &= \frac{1}{3}\vec{\mathcal{U}}\cdot\vec{\mathcal{V}}\delta_{mn} + \frac{1}{2}\epsilon_{ijk}\left(\vec{\mathcal{U}}\times\vec{\mathcal{V}}\right)_k \\
&\quad + \left[\frac{1}{2}(\mathcal{U}_m\mathcal{V}_n + \mathcal{U}_n\mathcal{V}_m - \frac{1}{3}\vec{\mathcal{U}}\cdot\vec{\mathcal{V}}\delta_{mn}\right]
\end{aligned} \tag{290}$$

and we recover clearly the combinations above.

4.4.2 Irreducible tensor operators

The tensor operators $\mathcal{T}^{(0)}$, $\mathcal{T}^{(1)}$, and $\mathcal{T}^{(2)}$ constructed above are dubbed *irreducible tensor operators*. In general, we define an irreducible tensor operator $T(k)$ of degree $k = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ as an operator having $2k+1$ components $T_q^{(k)}$, that transform like

$$U_R^\dagger \mathcal{T}_q^{(k)} U_R = \sum_{m=-k}^{k} D^{(k)}_{mq}(R) T^{(k)}_m \tag{291}$$

under rotations. Rephrased in other words : $\mathcal{T}^{(k)}$ is an irreducible tensor operator, if its $2k+1$ components generate the irreducible representation $D^{(k)}$ of $SO(3)$. In this sense we can understand the $2l+1$ spherical harmonics $Y_l^m(\theta,\phi)$ as standard (or spherical) components of an operator $\mathcal{Y}^{(l)}$ of degree l. Similarly, the 2^l electric or magnetic multipole moment operator forms an irreducible tensor operator of degree l. For $l = 0, 1, 2, 3 \ldots$ we have the monopole, dipole, quadrupole, octopole ... moment operators.
The observation, that the components of an irreducible tensor operator $\mathcal{T}^{(k)}$ for integral k have the same transformation properties as the spherical harmonics of degree k, gives a clear hindsight, that they have the same commutation relations with the components of the total angular momentum operator $\vec{J}$, namely

$$\begin{aligned}
\left[J_z, \mathcal{T}_q^{(k)}\right] &= q\hbar\, \mathcal{T}_q^{(k)} \\
\left[J_\pm, \mathcal{T}_q^{(k)}\right] &= \sqrt{k(k+1) - q(q\pm 1)}\hbar\, \mathcal{T}^{(k)}_{q\pm 1}\,.
\end{aligned} \tag{292}$$

These relations provide another suitable definition of irreducible tensor operators due to Racah : If the $2k+1$ operators $\mathcal{T}_q^{(k)}$ fulfill the commutation relations above, they are dubbed the standard components of an irreducible tensor operator of degree k.

4.4.3 Direct products of irreducible tensor operators

Remember once more our construction of tensor operators from two vector operators and let us extend some of the considerations to the product of two irreducible tensor operators.
Let $\mathcal{X}_{q_1}^{(k_1)}$ and $\mathcal{Y}_{q_2}^{(k_2)}$ denote two irreducible tensors of rank k_1 and k_2, respectively. Then the tensor

$$\mathcal{T}_q^{(k)} = \sum_{q_1,q_2} \langle k_1 k_2; q_1 q_2 | k_1 k_2; kq \rangle \mathcal{X}_{q_1}^{(k_1)} \mathcal{Y}_{q_2}^{(k_2)} \tag{293}$$

forms an irreducible tensor of rank k.
To proof this, we have to investigate the transformation properties of $T_q^{(k)}$ under rotations. In fact, it has to transform like

$$\begin{aligned} U_R^\dagger \mathcal{T}_q^{(k)} U_R &= \sum_{q_1,q_2} \langle k_1 k_2; q_1 q_2 | k_1 k_2; kq \rangle U_R^\dagger \mathcal{X}_{q_1}^{(k_1)} U_R U_R^\dagger \mathcal{Y}_{q_2}^{(k_2)} U_R \\ &= \sum_{q_1,q_2} \langle k_1 k_2; q_1 q_2 | k_1 k_2; kq \rangle \sum_{q_1' q_2'} \mathcal{X}_{q_1'}^{(k_1)} D_{q_1 q_1'}^{(k_1)}(R^{-1}) \mathcal{Y}_{q_2'}^{(k_2)} D_{q_2 q_2'}^{(k_2)}(R^{-1}) \\ &= \sum_{q_1,q_2} \langle k_1 k_2; q_1 q_2 | k_1 k_2; kq \rangle \sum_{q_1' q_2'} \sum_{k'' q'' q'} \langle k_1 k_2; q_1' q_2' | k_1 k_2; k'' q' \rangle \\ &\quad \times \ \langle k_1 k_2; q_1 q_2 | k_1 k_2; k'' q'' \rangle \mathcal{X}_{q_1'}^{(k_1)} \mathcal{Y}_{q_2'}^{(k_2)} D_{q' q''}^{(k'')}(R^{-1}) \,, \end{aligned} \tag{294}$$

where we have used the Clebsch–Gordan expansion

$$\begin{aligned} & D_{m_1 m_1'}^{(j_1)}(R) D_{m_2 m_2'}^{(j_2)}(R) = \\ & \qquad \sum_{j,m,m'} \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle \langle j_1 j_2; m_1' m_2' | j_1 j_2; jm' \rangle D_{mm'}^{(j)}(R) \end{aligned} \tag{295}$$

with $|j_1 - j_2| \leq j \leq j_1 + j_2$. The proof of this relation can be found in the appendix.
Using now the orthogonality of the Clebsch–Gordan coefficients in Eq.294 we find

$$U_R^\dagger \mathcal{T}_q^{(k)} U_R = \sum_{k'',q_1',q_2',q'',q'} \delta_{kk''} \delta_{qq''} \langle k_1 k_2; q_1' q_2' | k_1 k_2; k'' q' \rangle \mathcal{X}_{q_1'}^{(k_1)} \mathcal{Y}_{q_2'}^{(k_2)} D_{q' q''}^{(k'')}(R^{-1})$$

$$
\begin{aligned}
&= \sum_{q',q'_1,q'_2} \langle k_1 k_2; q'_1 q'_2 | k_1 k_2; kq' \rangle \mathcal{X}^{(k_1)}_{q'_1} \mathcal{Y}^{(k_2)}_{q'_2} D^{(k)}_{q'q}(R^{-1}) \\
&= \sum_{q'} T^{(k)}_{q'} D^{(k)}_{q'q}(R^{-1}) = \sum_{q'} D^{(k)*}_{q'q}(R)\, T^{(k)}_{q'} \,. \qquad (296)
\end{aligned}
$$

From this we learn, that the construction of irreducible tensor out of other irreducible tensors proceeds in exactly the same manner in which we construct angular momentum eigenstates by coupling two angular momenta, exactly the same Clebsch–Gordan coefficients emerge, up to replacements $k_{1m2} \to j_{1,2}$, $q_{1,2} \to m_{1,2}$.

4.4.4 Matrix elements of tensor operators, finally the Wigner–Eckart theorem

Quite often we have to evaluate matrix elements of tensor operators with respect to angular momentum eigenstates, for instance, if we consider the interactions of electromagnetic fields with atoms and nuclei, this task is paramount. In fact, this evaluation is not only of big importance, but also of formidable difficulty. However, there are several properties of those matrix elements which follow directly from geometry or kinematic considerations, and which somewhat diminish the computational pain. First of all, there is a simple m–selection rule, namely

$$
\langle \alpha', j'm' | \mathcal{T}^{(k)}_q | \alpha, jm \rangle = 0 \;\; \text{unless}\; m' = m + q \,. \qquad (297)
$$

Going back to Eq.292 we can write

$$
\begin{aligned}
&\left\langle \alpha', j'm' \,\middle|\, \left[J_z, \mathcal{T}^{(k)}_q \right] - \hbar q \}, \mathcal{T}^{(k)}_q \,\middle|\, \alpha, jm \right\rangle = \\
&\qquad [(m' - m) - q]\, \hbar \langle \alpha', j'm' | \mathcal{T}^{(k)}_q | \alpha, jm \rangle = 0 \qquad (298)
\end{aligned}
$$

proving the statement above. Another way to see this is to employ the transformation properties of $\mathcal{T}^{(k)}_q |\alpha, jm\rangle$ under rotations around the z–axis,

$$
\begin{aligned}
D(\vec{e}_z, \phi) \mathcal{T}^{(k)}_q | \alpha, jm \rangle &= D(\vec{e}_z, \phi) \mathcal{T}^{(k)}_q D^\dagger(\vec{e}_z, \phi) D(\vec{e}_z, \phi) | \alpha, jm \rangle \\
&= e^{-iq\phi} e^{-im\phi} \mathcal{T}^{(k)}_q | \alpha, jm \rangle \,, \qquad (299)
\end{aligned}
$$

which is orthogonal to $|\alpha', j'm'\langle$, unless $m' = q + m$.
But more importantly, beyond this simple selection rule, there is the celebrated *Wigner–Eckart theorem*, one of the most important theorems in quantum mechanics. It states, that the matrix elements of tensor operators with respect to angular momentum eigenstates satisfy

$$
\langle \alpha', j'm' | \mathcal{T}^{(k)}_q | \alpha, jm \rangle = \langle jk; mq | jk; j'm' \rangle \frac{\alpha' j' || T^{(k)} || \alpha j \rangle}{\sqrt{2j+1}} \,, \qquad (300)
$$

where the so–called *double–bar matrix element* is independent of m', m, and q. The proof of this theorem goes as follows. Using Eq.292 we have

$$\left\langle \alpha', j'm' \left| \left[J_\pm, \mathcal{T}_q^{(k)} \right] \right| \alpha, jm \right\rangle = \\ \hbar\sqrt{(k \mp q)(k \pm q + 1)} \left\langle \alpha', j'm' \left| \mathcal{T}_{q\pm 1}^{(k)} \right| \alpha, jm \right\rangle . \tag{301}$$

Alternatively we might employ

$$J_\pm |jm\rangle = \sqrt{(j \mp m)(j \pm m + 1)}\hbar |jm \pm 1\rangle \tag{302}$$

and find

$$\begin{aligned} &\sqrt{(j' \pm m')(j' \mp m' + 1)}\langle \alpha', j'm' \mp 1 | \mathcal{T}_q^{(k)} | \alpha, jm\rangle \\ &= \quad \sqrt{(j \mp m)(j \pm m + 1)}\langle \alpha', j'm' \mp 1 | \mathcal{T}_q^{(k)} | \alpha, jm \pm 1\rangle \\ &\quad + \sqrt{(k \mp q)(k \pm q + 1)}\langle \alpha', j'm' \mp 1 | \mathcal{T}_{q\pm 1}^{(k)} | \alpha, jm\rangle . \end{aligned} \tag{303}$$

Replacing $j' \to j$, $m' \to m$, $j \to j_1$, $m \to m_1$, $k \to j_2$, and $q \to m_2$, there is a striking similarity with the recursion relation for the Clebsch–Gordan coefficients, Eq.253. Both recursion relations are of the form

$$a_{ij} x_j = 0 \text{ or } \; a_{ij} y_j = 0 \tag{304}$$

with the same coefficients a_{ij}. We then cannot solve for the x_j or y_j individually, but for the ratios, yielding

$$\frac{x_j}{x_k} = \frac{y_j}{y_k} \implies x_j = c y_j . \tag{305}$$

Here, c is some universal proportionality factor. Identifying $\langle j_1 j_2; m_1 m_2 \pm 1 | j_1 j_2; jm\rangle$ with $\langle \alpha', j'm' \mp 1 | \mathcal{T}_{q\pm 1}^{(k)} | \alpha, jm\rangle$ we see immediately, that

$$\langle \alpha', j'm' \mp 1 | \mathcal{T}_{q\pm 1}^{(k)} | \alpha, jm\rangle = C\langle jk; mq \pm 1 | jk; jm'\rangle , \tag{306}$$

where C is some universal proportionality constant independent of m, q, and m', the double–bar matrix element. This proves the theorem.
Let us now dwell a little bit on its significance. First of all, we see, that the matrix element in Eq.300 is written as the product of two factors. The first factor is the Clebsch–Gordan coefficient for adding $j + k$ to get j'. This factor depends only on the geometry of the problem, i.e. on the orientation of the system with respect to the z–axis. All the dynamics incorporated in the

particular nature of the tensor operator is concentrated in the second factor. For example, radial quantum numbers and the like, denoted by α, which play a dynamical role in the problem can be found in this factor, but there is no information whatsoever on the magnetic quantum numbers m, m', and q, which specify the orientation of the system. This means, that in order to evaluate $\langle\alpha', j'm'|\mathcal{T}_q^{(k)}|\alpha, jm\rangle$ with combinations of the magnetic quantum numbers we just have to know one of the matrix elements, all others are related by geometry, or by factors proportional to Clebsch–Gordan coefficients. The common proportionality factor is, as indicated above, the double–bar matrix element, which makes no reference to any geometrical feature. Selection rules can then be read of immediately, once we remember the requirement for the Clebsch–Gordan coefficients to be non–zero. These selection rules are the m–selection rule and the triangular relation

$$|j-k| \leq j' \leq j+k\,. \tag{307}$$

We will elucidate this nice theorem by some examples.

1. Scalar operators :
   For tensors of rank 0, i.e. scalar operators we have

   $$\langle\alpha', j'm'|\mathcal{T}_0^{(0)}|\alpha, jm\rangle = \delta_{jj'}\delta_{mm'}\frac{\langle\alpha'||\mathcal{T}_0^{(0)}||\alpha\rangle}{\sqrt{2j+1}}\,, \tag{308}$$

   showing, that they cannot change the angular momentum quantum numbers.

2. Vector operators :
   For vector operators in spherical components we find the selection rules

   $$\Delta m = m'-m = \pm 1,\, 0 \;\; \Delta j = j'-j = \pm 1,\, 0\,. \tag{309}$$

   This selection rule is of fundamental importance in the theory of radiation, it is the dipole selection rule we discuss in Sec.5.2.4.

5 Symmetry at work : The hydrogen atom

5.1 Symmetry of the Hamiltonian

5.1.1 Symmetry and degeneracy

In analogy to Noethers theorem in classical mechanics, we can relate symmetries of the Hamiltonian describing a system with constants of motion. Let us consider for the moment single–particle Hamiltonians, which are of the form

$$\mathcal{H} = -\frac{\hbar^2 \nabla^2}{2m} + V \,, \tag{310}$$

where the first term is the kinetic energy operator for a particle with mass m and V contains all the interactions of this particle. We can easily see, that the ∇^2 and thus the first term are invariant under the group $O(3)$, therefore the symmetry of $\mathcal{H}$ is essentially governed by the symmetry of the potential term V. Applying now the operation P_R related to some coordinate transformation R on the Schrödinger Equation, we have

$$P_R \mathcal{H}|\Psi\rangle = P_R \mathcal{H} P_R^{-1} P_R|\Psi\rangle = \mathcal{H}' P_R|\Psi\rangle = E_\Psi P_R|\Psi\rangle \,. \tag{311}$$

If we wish the new Hamiltonian to be equal to the untransformed one, $\mathcal{H}' = \mathcal{H}$, we find that

$$[\mathcal{H}, P_R] = 0 \,. \tag{312}$$

The set of all transformations R which leaves the system invariant, forms a group, dubbed the *symmetry group* G of the Hamiltonian. The commutation relation above, implies, that not only the $|\Psi\rangle$, but also the $P_R|\Psi\rangle$ are eigenvectors of $\mathcal{H}$ with the same eigenvalue E_Ψ. Thus, unless we are talking about simple multiples of the vectors, the states $|\Psi\rangle$ are *degenerate* with $P_R|\Psi\rangle$.

If we start now by some eigenvector $|\Psi_1\rangle$ of $\mathcal{H}$ with eigenvalue E_1 and apply repeatedly the symmetry transformations of G we will end up with a set of independent eigenvectors $|\Psi_{i,\dots,n}\rangle$, where $n \leq g$, the order of the group G. If our group G includes *all* possible symmetry transformations leaving $\mathcal{H}$ invariant, then, in general, the eigenvectors form a basis for an irreducible representation of G. This statement can be reversed as follows : All eigenvectors of an Hamiltonian belonging to an irreducible representation of G can be transformed into each other via corresponding symmetry transformations, hence they're degenerate. Even more, in general, eigenvectors transforming according to different representations of G have different eigenvalues, i.e. they're not degenerate with each other. This is quite easy to

understand. Just consider two vectors belonging to two different irreducible representations. Since there is no way within G to produce any base transformation they do not mix with each other. Therefore they are non–degenerate with each other and thus have different eigenvalues.
However, these statements have to be taken with a grain of salt, since in practice, we can never be certain, that we found the full symmetry and all the invariances of a system. Even more, it may still happen, that for some regions of the parameter space of a problem there are identical eigenvalues for eigenstates of different irreducible representations. But this type of degeneracy is not enforced by symmetry and we cannot infer it with our nice apparatus developed so far. In principle we should be capable to remove this accidental symmetry, at least in theoretically, by fiddling round with the parameters of the problem. As an example for this accidental symmetry consider a two–level system of fermions. If we include spin effects and switch on a magnetic field, we find, that each of the level splits in two because of the interaction of the spin with the background field. Of course, we will be able to find a field configuration, where two of these levels overlap, i.e. have the same energy. But this symmetry disappears as fast as it appeared by just altering the field strength once more.

5.1.2 Good quantum numbers

After briefly mentioning this interesting but somewhat esoteric scenario (as far as the topic of this lecture is concerned) we would like to continue on our path linking symmetries of the Hamiltonian with observable properties of the system it describes. In this framework the importance of the connection between dimensions of irreducible representations of the symmetry group and degeneracies of the energy levels can not be overestimated. In fact, this link provides us in a natural way with a system to label the energy levels and the corresponding eigenfunctions by the irreducible representations. In this sense it gives us an handle on the degeneracies we expect.
Let us rephrase this statement somewhat different to highlight the meaning. Assume, we have an eigenvector $|\Psi^{\alpha}_{pm}\rangle$, that belongs to the irreducible representation α, in which it is labeled by m, and occurs the pth time in the energy level scheme. Then, $E_{\alpha p}$ is a reasonable name for the energy eigenvalue, and m carries the information needed to account for the degenerate states. Take as a practical example the eigenfunctions of the hydrogen atom. Clearly, they can be written as $R_{nl}Y_l^m(\theta, \phi)$, where the energy levels can be denoted with n $(\to p)$ and l $(\to \alpha)$, and m labels the remaining degenerate states. In this sense, n and l are *good quantum numbers* to describe the energy levels of the hydrogen atom. Add m and the set of good quantum numbers to specify the corresponding eigenstates is complete. Thus *good quantum numbers* remaining invariant under symmetry transformations of the system are the most suitable way of labeling the eigenvectors and eigenvalues. Hence, it seems natural to rephrase the time–independent Schrödinger equation as

$$\mathcal{H}|\Psi^{\alpha}_{pm}\rangle = E_{\alpha p}|\Psi^{\alpha}_{pm}\rangle\,. \tag{313}$$

It should not come as a big surprise, that with the choice of the eigenvectors of the irreducible representations of the symmetry group as basis vectors, the Hamiltonian can be cast into a block–diagonal form.

5.1.3 Perturbation and level splitting

However, we still have to deal with the fact, that there's only a small number of problems in quantum mechanics which are exactly solvable. A good approach to construct solutions for the other problems relies on perturbation theory. Here, we write the Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + V\,, \tag{314}$$

where $\mathcal{H}_0$ is simple enough for us to solve it, e.g. determine eigenvectors and eigenvalues, and V has only a small effect on them.
If now G_0 is the symmetry group of $\mathcal{H}_0$, its transformations will, in general, not leave V invariant. In other words, the symmetry group G of $\mathcal{H}$ will in almost all cases be smaller than G_0. We will assume in the following, that G is a subgroup of G_0. The question is now, what happens to the l_α–fold degenerate eigenvectors of $\mathcal{H}_0$ for eigenvalue $E^{(0)}_\alpha$? Before switching "on" the perturbation V, they form a basis for the corresponding irreducible representation $\Gamma^{(0)}_\alpha$ of $G^{(0)}$ related to the eigenvalue $E^{(0)}_\alpha$. Now, the symmetry group $G^{(0)}$ is reduced to its subgroup G, which in general has lower dimension. Thus, the eigenvectors still generate a representation of G, but this representation will in general be a reducible one. We already know how to reduce reducible representations, and of course, we know the shape of the result of such a process. The eigenvectors of the previously irreducible representation $\Gamma^{(0)}_\alpha$ of $G^{(0)}$ will form subsets of lower dimension related to irreducible representations of the subgroup G. The vectors of different subsets labeled by β won't mix with each other. Each subset, i.e. each irreducible representation of G will be related to some eigenvalue $E_{\alpha\beta}$ of G. Thus, the previous energy level $E^{(0)}_\alpha$ splits into a number of levels, the degeneracy is lifted to some extend. In the following we will highlight this effect of perturbations with the example of the hydrogen atom.

5.2 The hydrogen atom without spin

5.2.1 Symmetry group

Let us start with the simplest model of an hydrogen atom, consisting of a spinless "electron" in a Coulomb potential

$$V(\vec{r}) = -\frac{Ze^2}{|\vec{r}|} \tag{315}$$

related to some nuclear charge $Z \cdot e$. Clearly, this potential is rotational invariant and invariant under inversions. Thus, the symmetry group of the full Hamiltonian

$$\mathcal{H} = \frac{\vec{p}^2}{2\mu} - \frac{Ze^2}{|\vec{r}|} \tag{316}$$

seems to have the symmetry group $O(3)$. Here, μ is the reduced mass of the problem, roughly the electron mass. The analogon of classical mechanics is the well–studied Kepler problem, and we know, that the classical orbit of our electron is an ellipse with the nucleus in one of the foci. In fact, this introduces another symmetry into the problem, since we have another invariant, the vector linking the nuclear charge in the focus with either the aphel or the perihel of our electronic orbit. In quantum mechanics, we can show, that the so–called Runge–Lenz vector

$$\vec{M}' = \frac{1}{2\mu}\left(\vec{p} \times \vec{L} - \vec{L} \times \vec{p}\right) - \frac{Ze^2\vec{r}}{|\vec{r}|} \tag{317}$$

is another constant of motion in addition to the angular momentum, which would be our first guess. In fact, we have

$$\vec{M}'\vec{L} = 0\,, \tag{318}$$

since $\vec{r}$, $\vec{p} \times \vec{L}$ and $\vec{L} \times \vec{p}$ lie within the plane of the orbit (and are perpendicular to the trajectory) and $\vec{L}$ is perpendicular to it. As we know,

$$\left[\vec{L}, \mathcal{H}\right] = 0 \;\text{ and }\; \left[\vec{M}', \mathcal{H}\right] = 0\,. \tag{319}$$

This leaves us with six operators (three components of $\vec{L}$ and $\vec{M}'$), which correspond to invariants of the problem. Their commutators are, up to cyclic permutations

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k\,, \tag{320}$$

as we already know. Furthermore, we have

$$\begin{aligned}
[M'_i, L_j] &= i\hbar\epsilon_{ijk}\, M'_k \\
\left[M'_i, M'_j\right] &= -\frac{2i\hbar}{\mu}\,\mathcal{H}\,\epsilon_{ijk}L_k
\end{aligned} \tag{321}$$

and

$$\vec{M}'^2 = \frac{2\mathcal{H}}{\mu} \left(\vec{L}^2 + \hbar^2\right) + Z^2 e^4 \,. \tag{322}$$

Proofs of these relations can be found in the appendix.
The operators L_i form a closed algebra related to $O(3)$, but expanding this algebra with the M_i' leaves us with an algebra, which is not closed anymore ($\mathcal{H}$ appears). However, if we restrict ourselves on some bound state with $E < 0$ and the corresponding subspace of the full space of all energy levels, then we can replace $\mathcal{H}$ by E and close the algebra. Defining a new operator

$$\vec{M} = \sqrt{-\frac{\mu}{2E}}\,\vec{M}' \tag{323}$$

we see, that we can replace M' by M in all commutation relations above, with the only exception of

$$[M_i, M_j] \;=\; i\hbar\epsilon_{ijk} L_k \,. \tag{324}$$

But wait ! Now we have a closed algebra with six generators, i.e. a six parameter Lie group. We will see, that this group is $O(4)$. Let's define a new set of operators, namely

$$J_{ij} = \epsilon_{ijk} L_k \;\text{ and }\; J_{iw} = -J_{wi} = M_i \,, \tag{325}$$

where i, j and k label the space components x, y and z. The commutation relations of the new operators are

$$\begin{aligned} [J_{xy}, J_{yz}] &= i\hbar J_{zx} \,, & [J_{zw}, J_{xw}] &= i\hbar J_{zx} \,, \\ [J_{xw}, J_{yz}] &= 0, & [J_{xw}, J_{zx}] &= i\hbar J_{zw} \,. \end{aligned}$$

These operators define the infinitesimal operations of $O(4)$, which leaves $x^2 + y^2 + z^2 + w^2$ invariant under rotations.

5.2.2 Energy levels

But we can do even better. Let us consider two more helpers,

$$\vec{A} = \frac{1}{2}(\vec{L} + \vec{M}) \,, \;\; \vec{B} = \frac{1}{2}(\vec{L} - \vec{M}) \,. \tag{326}$$

They have the nice property, that their commutators read

$$[A_x, A_y] \quad = \quad \hbar A_z \text{ and cycl. }, \ \ [B_x, B_y] = \hbar B_z \text{ and cycl. }, \ \ [A_i, B_j] = 0 \tag{327}$$

or, in a nice shorthand,

$$\vec{A} \times \vec{A} \quad = 7i\hbar\vec{A} \ , \ \ \vec{B} \times \vec{B} = i\hbar\vec{B} \ , \ \ [A_i, B_j] = 0 \,. \tag{328}$$

Of course, since $\vec{L}$ and $\vec{M}$ commute with $\mathcal{H}$ these two entities do as well, but they form two closed subalgebras. The nicest feature of these subalgebras is, that we already know them, in fact, they're our good old friends, the $SU(2)$! Thus $O(4)$ can be decomposed as $SU(2) \otimes SU(2)$. In fact, we have two Casimir operators, one for each $SU(2)$ or $\vec{L}^2$ and $\vec{M}^2$. However, we prefer to start with

$$\vec{A}^2 = a(a+1)\hbar^2 \ \text{ and } \ \vec{B}^2 = b(b+1)\hbar^2 \tag{329}$$

with a and b all nonnegative integers or half–integers. Now we combine these two operators as sum and difference to find

$$C = \vec{A}^2 + \vec{B}^2 = \frac{1}{2}(\vec{L}^2 + \vec{M}^2) \ , \ \ C' = \vec{A}^2 - \vec{B}^2 = \vec{L} \cdot \vec{M} = 0 \tag{330}$$

with help of Eq.(318). We find immediately, that $\vec{A}^2 = \vec{B}^2$ and, discarding negative values for a and b, $a = b$. Therefore, the physically sensible spectrum of $O(4)$ in the case of the hydrogen atom is constrained to representations of the form (a, a). Consequently,

$$C = 2a(a+1)\hbar^2 \tag{331}$$

and inserting the definitions of Eqs.(329,323 and 322) we find

$$\begin{aligned} C &= \frac{1}{2}\left\{\vec{L}^2 - \frac{\mu}{2E}\left[\frac{2E}{\mu}(\vec{L}^2 + \hbar^2) + Z^2 e^4\right]\right\} \\ &= -\frac{1}{2}\left[\hbar^2 + \frac{\mu Z^2 e^4}{2E}\right] \end{aligned} \tag{332}$$

and, finally,

$$E = -\frac{\mu Z^2 e^4}{2\hbar^2(2a+1)^2} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} \,. \tag{333}$$

This leaves us with energy levels labeled by $n = 1, 2, \ldots$, where each level has an n^2–fold degeneracy. We note in passing, that this is quite good news for the anxious bookkeeper. In fact, we found the extra symmetry, that would explain, why the energy levels have no $2l+1$–fold degeneracy (with angular momentum $l = 0, 1, \ldots$), but an $(2l+1)^2$–fold degeneracy (with $l = 2a$). To highlight this point, just remember, that the second energy level (related with $l = 1$), has obviously 4 orbitals instead of 3.

5.2.3 Parity of the states

Before we continue our journey through the spectrum of the hydrogen atom, let us dwell briefly on the issue of parity, that is, the behavior of a state etc. under space–inversions. Actually, this behavior will tell us something about the allowed transitions between different states of an atom. Recall, that our Coulomb potential, Eq.315, is invariant not only under rotations in space , but also under inversions, that is, the symmetry group of it is rather $O(3)$ than $SO(3)$. Both groups are related via

$$O(3) = SO(3) \otimes \{E, I\} \tag{334}$$

with I the inversion operator (its effect is to switch the sign of any space coordinate). The two irreducible representations of the group $\{E, I\}$ are

$$\begin{aligned} \Gamma^{+}(E, I) &= +1 \\ \Gamma^{-}(E, I) &= \pm 1 \,, \end{aligned} \tag{335}$$

and we recover the identity representation in Γ^{+}. Consequently, irreducible representations of $O(3)$ are direct products of irreducible representations of $SO(3)$ and $\{E, I\}$,

$$D^{(l,\sigma)} = D^{(l)} \otimes \Gamma^{\sigma} \,. \tag{336}$$

Obviously, since the Γ^{σ} are of dimension 1, the dimension of the representation $D^{(l,\sigma)}$ is the same as that of $D^{(l)}$, namely $2l+1$. We conclude, that the group $O(3)$ has two different irreducible representations for any value of l. We are now going to construct these beasts.
To obtain the actual matrices representing $O(3)$, we note, that we can divide the elements of the group into two sets, $\{X\}$ and $\{IX\}$, where X runs over all elements of $SO(3)$. Thus,

$$\begin{aligned} D^{(l,\sigma)}(EX) &= D^{(l)}(X) \otimes \Gamma^{\sigma}(E) = D^{(l)}(X) \\ D^{(l,\sigma)}(IX) &= D^{(l)}(X) \otimes \Gamma^{\sigma}(I) = \sigma D^{(l)}(X) \,. \end{aligned} \tag{337}$$

This behavior can be tracked into the class structure of $O(3)$, and we find

$$\begin{aligned} \chi^{(l,\sigma)}(E\alpha) &= \chi^{(l)}(\alpha) \\ \chi^{(l,\sigma)}(I\alpha) &= \sigma \chi^{(l)}(\alpha) \end{aligned} \tag{338}$$

for the characters of class α. The eigenfunctions of an electron in atom may belong to any of the irreducible representations above. If it belongs to $D^{(l,+)}$ it is dubbed as to be of *even parity*, and it remains invariant under space

l	name	parity	representation	degeneracy
0	$1s$	even	$D^{(0,+)}$	1
0	$2s$	even	$D^{(0,+)}$	1
1	$2p$	odd	$D^{(1,-)}$	3
0	$3s$	even	$D^{(0,+)}$	1
1	$3p$	odd	$D^{(1,-)}$	3
2	$3d$	even	$D^{(2,+)}$	5
0	$4s$	even	$D^{(0,+)}$	1
1	$4p$	odd	$D^{(1,-)}$	3
2	$4d$	even	$D^{(2,+)}$	5
3	$4f$	odd	$D^{(3,-)}$	7

Table 3: Naming scheme, parity, representation and degeneracy of electronic states in one–electron systems with different orbital angular momentum. We did not include the effect of electron spin here, which first doubles and then lifts the degeneracy.

inversions. If it belongs to $D^{(l,-)}$ it is of *odd parity.* Obviously, the parity of a state is determined by its dependence on the space coordinates, but it does not depend on the spin angular momentum (which is invariant under inversions). However, for an one–electron system, not all combinations of l and σ can occur. This can be easily explained by noting, that the one–electron wave functions are homogenous polynomials of degree l in the space coordinates, where l is the orbital quantum number. Coordinate inversion effectively multiplies these polynomials with $(-1)^l$, when switching $x \to -x$, $y \to -y$ and $z \to -z$. Thus, permissible irreducible representations are the ones with even parity for even angular momentum and odd parity for odd angular momentum, i.e. $D^{(0,+)}$, $D^{(1,-)}$, $D^{(2,+)}$, $D^{(3,-)}$, ... For the spectroscopic nomenclature and the parity of electronic levels in one–electron systems, see Table3.
The statements above for the parity of one–electron systems do not hold true in many–electron systems, since here one–electron states of different parity can get mixed, prohibiting the clear identification outlined above.
In the following we will list the symmetry behavior of the spherical harmonics, see Table4, which are related to the representations $D^{(l)}$ of $SO(3)$ via linear combinations (in Table5).

5.2.4 Basic selection rules for atomic transitions

We are now equipped with the notion of parity for atomic systems and we have learned, that we have only a limited set of irreducible representations of $O(3)$ at our disposal for constructing the eigenfunctions of one–electron systems. But what does this tell us about physics ? Well, for a beginning, let us start with transitions of the electron within the atom. Usually, they are triggered by the emission or absorption of some photon, i.e. by interaction

l	Symmetry	Angular part
0	$s : 1$	constant
1	$p : x$	$\sin\theta\cos\phi$
	$p : y$	$\sin\theta\sin\phi$
	$p : z$	$\cos\theta$
2	$d : xy$	$\sin^2\theta\sin 2\phi$
	$d : yz$	$\sin\theta\cos\theta\sin\phi$
	$d : zx$	$\sin\theta\cos\theta\cos\phi$
	$d : x^2 - y^2$	$\sin^2\theta\cos 2\phi$
	$d : 2z^2 - x^2 - y^2$	$3\cos^2\theta - 1$

Table 4: Angular symmetries of atomic eigenfunctions.

$D^{(l)}$	The linear combination	is proportional to
$D^{(0)}$	1	$Y_0^0 = \sqrt{\frac{1}{4\pi}}$
$D^{(1)}$	$x \pm iy$	$Y_1^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\exp(\pm i\phi)$
	z	$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$
$D^{(2)}$	$2z^2 - x^2 - y^2$	$Y_2^0 = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1)$
	$zx \pm iyz$	$Y_2^{\pm 1} = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\exp(\pm i\phi)$
	$(x^2 - y^2) \pm 2ixy$	$Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta\exp(\pm 2i\phi)$

Table 5: The standard symmetrized linear combinations, that generate the representation $D^{(l)}$ of $SO(3)$.

with an electromagnetic field, which acts in the form of dipoles. For instance, the electric dipole momentum operator has the form

$$\vec{\mu} = e\vec{e}_r = e(x, y, z)^T , \tag{339}$$

which is a polar vector and changes sign under inversion. In fact, it can be represented by $D^{(1,-)}$. The transition of a state related to $D^{(l,\sigma)}$ to a state $D^{(l',\sigma')}$ under the influence of the electric dipole operator is only allowed, if the direct product representation $D^{(1,-)} \otimes D^{(l,\sigma)}$ contains $D^{(l',\sigma')}$, because otherwise, no overlap of the wave functions would be possible. We can decompose the direct product as a direct sum by means of the techniques employed in Eq.147. The result is

$$D^{(1,-)} \otimes D^{(l,\sigma)} = D^{(l-1,-\sigma)} \oplus D^{(l,-\sigma)} \oplus D^{(l+1,-\sigma)} , \tag{340}$$

with $l = 0$ forbidden. Noting, that, if $D^{(l,\sigma)}$ exists, $D^{(l,-\sigma)}$ does not in one–electron systems, we find the rule, that *electric dipole transitions flip parity and induce a change in orbital angular momentum of one unit,* $\Delta l = \pm 1$. On the other hand, the magnetic dipole operator is related to an axial vector (no change under inversion) and is therefore represented by $D^{(1,+)}$. Again, we decompose the direct product and find

$$D^{(1,+)} \otimes D^{(l,\sigma)} = D^{(l-1,\sigma)} \oplus D^{(l,\sigma)} \oplus D^{(l+1,\sigma)} , \tag{341}$$

and find as a selection rule, that *parity and orbital angular momentum remain invariant under magnetic dipole transitions.*

5.2.5 Lifting degeneracies I : The linear Stark effect

As a first example, of how degeneracies can be lifted by diminishing the underlying symmetry of the Hamiltonian, let us consider the so–called linear Stark effect. This effect appears when applying a uniform electric field on an one–electron atom, more precisely its excited states. Remember, that so far, we have not included any spin effects yet, and consequently, the energy of the bound states depends only on their principal quantum number n. Thus, states with different orbital angular momentum l but identical n are degenerate. To be more specific, let us concentrate now on the $n = 2$ level. Here, we find a $l = 0$ state $2s$ and three $l = 1$ states, which differ in $m = \pm 1, 0$. Recalling Eq.333,

$$E = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} = -\frac{e^2 a_0}{2n^2} \tag{342}$$

we therefore find identical energies $E_2 = e^2 a_0/8$ for these four states. Applying now a weak electrical field in z–direction, the perturbation operator reads

$$V = -ez|\vec{E}|\,. \tag{343}$$

To calculate the resulting level–splitting, we first have to diagonalize this operator, when applied to the $n = 2$–states. (Note, that we inherently assumed, that the perturbation is small, i.e. its effect on the levels is smaller than the energy difference between levels of different n.) Recalling the selection rules above, we find, that the electrical dipole operator related to the perturbation potential has non–vanishing matrix elements only, if the parity is flipped and if l differs by one unit. Therefore, ordering the rows and columns as $2s$, $2p, m = 0$, $2p, m = \pm 1$ we have

$$V = \begin{pmatrix} 0 & \langle 2s|V|2p, m=0\rangle & 0 & 0 \\ \langle 2p, m=0|V|2s\rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{344}$$

with

$$\langle 2s|V|2p, m=0\rangle = \langle 2p, m=0|V|2s\rangle = 3ea_0|\vec{E}|\,. \tag{345}$$

Noting, that the upper left part of our perturbation matrix is proportional to σ_x we can immediately write down the answer for the energy eigenvectors and the energy shifts related to them. We find, that combinations

$$|\pm\rangle = \frac{1}{\sqrt{2}}\left(|2s, m=0\rangle \pm |2p, m=0\rangle\right) \tag{346}$$

diagonalize V and lead to energy shifts

$$\Delta_\pm = \pm 3ea_0|\vec{E}|\,. \tag{347}$$

We find, that applying an electric field $\vec{E}$ pointing into z–direction splits the four degenerate states with principal quantum number $n = 2$ into three levels, namely the two states $|2p, m = \pm 1\rangle$ with the original energy, one state with energy lowered by Δ_- related to the $-$–vector and one with energy raised by the same amount linked to the $+$–vector.

5.3 Including spin

So far, we have been concerned with spinless hydrogen atoms only, which isn't quite found in nature. To close in on nature, we are going to include spin now, but we will stay in our non–relativistic framework. In the following we are going to discuss two effects related to the inclusion of the electrons spin, namely the spin–orbit coupling giving rise to an extra potential and the fine–structure, and the Zeeman–effect (sometimes called the *anomalous Zeeman effect*), which occurs, when applying a magnetic field to the atom.

5.3.1 Spin–orbit coupling and fine structure

Let us start with the fine structure introduced by the spin of the electron and its interaction with the orbital angular momentum. We can understand this spin–orbit coupling in a qualitative fashion as follows. The Coulomb potential V_c of Eq.315 is clearly a central force, and the electron thus experiences the electrical field

$$\vec{E} = -\frac{1}{e}\nabla V_c \,. \tag{348}$$

But, as we know from classical physics, whenever a moving particle experiences an electrical field, this can be re-expressed as an effective magnetic field,

$$\vec{B}_{\text{eff}} = -\vec{v} \times \vec{E} \,. \tag{349}$$

The spin $\vec{S}$ provides the electron with a magnetic moment

$$\vec{\mu} = \frac{g_s e\vec{S}}{2m_e} \,, \tag{350}$$

Bohr's magneton. The factor g_s is called the gyromagnetic factor, which is $g_s = 2$ for an electron. This magnetic moment might interact with the magnetic field as

$$\begin{aligned} V_{LS} &= -\vec{\mu} \cdot \vec{B}_{\text{eff}} = \vec{\mu} \cdot (\vec{v} \times \vec{E}) \\ &= -\left(\frac{e\vec{S}}{m_e}\right) \cdot \left[\frac{\vec{p}}{m_e} \times \frac{\vec{r}}{r}\frac{dV_c}{edr}\right] \\ &= \frac{1}{m_e^2 r}\frac{dV_c}{dr}\vec{L} \cdot \vec{S} \,. \end{aligned} \tag{351}$$

Comparing this ansatz we find, that it seems to be working very well up to a factor of 2, which needs the full relativistic theory of QED.
We return now to our problem of solving the new Hamiltonian including the spin–orbit interaction V_{LS},

$$H = H_0 + V_{LS} = \frac{\vec{p}^2}{2m} + V_c + V_{LS} \,. \tag{352}$$

Without the LS–part, we are free to choose either the $\{m_1, m_2\}$–type base of vectors given by eigenvectors of L^2, S^2, L_z, S_z or the $\{j, m\}$–like basis defined by the common eigenvectors of $L^2, S^2, J^2 = (L+S)^2, J_z$. Including V_{LS} the first set of base vectors is not as good as the second set. In fact it leads to a painful calculation to diagonalize V_{LS} in the L_z, S_z representation just to find out , that the J^2, J_z eigenvectors do the job perfectly well. Actually, this is very good news. Remember, that to obtain the first–order energy shift due to a perturbation in a diagonal base, we only have to take the expectation value of it. This is not a tremendous task. Just take a look into our spin–angular two–component functions of Eq.265

$$\mathcal{Y}_l^{j=l\pm\frac{1}{2},m} = \frac{1}{\sqrt{2l+1}} \begin{pmatrix} \pm\sqrt{l+\frac{1}{2}\pm m}\,Y_l^{m-\frac{1}{2}}(\theta,\phi) \\ +\sqrt{l+\frac{1}{2}\mp m}\,Y_l^{m+\frac{1}{2}}(\theta,\phi) \end{pmatrix} \tag{353}$$

and construct wave functions as

$$\Psi_{nlm} = \mathcal{R}_{nl}(r)\mathcal{Y}_l^{j=l\pm 1/2,m} . \tag{354}$$

Thus, we obtain

$$\Delta_{nlj} = \frac{1}{4m_e^2}\left\langle \frac{1}{r}\frac{dV_c}{dr}\right\rangle \cdot \begin{cases} l & for \quad j = l+\frac{1}{2} \\ -(l+1) & for \quad j = l-\frac{1}{2} \end{cases} . \tag{355}$$

In the derivation of the equation for the level splitting above above, we have used that

1. $$\left\langle \frac{1}{r}\frac{dV_c}{dr}\right\rangle = \int dr r^2 \mathcal{R}_{nl}^*(r)\frac{dV_c}{rdr}\mathcal{R}_{nl}(r) \tag{356}$$

   and,

2. that with $\vec{J} = \vec{L} + \vec{S}$ we have

   $$\vec{S}\cdot\vec{L} = \frac{1}{2}\left(\vec{J}^2 - \vec{L}^2 - \vec{S}^2\right) \tag{357}$$

   and, consequently, that we can rewrite the angular integral as

   $$\int d\Omega \mathcal{Y}^\dagger \vec{S}\cdot\vec{L}\mathcal{Y} = \frac{1}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right] \tag{358}$$

   employing, that $j = l \pm 1/2$.

We are now going to estimate the size of the resulting level–splitting for hydrogen atoms with $Z = 1$.
We estimate the expectation value as

$$\left\langle \frac{1}{r}\frac{dV_c}{dr} \right\rangle \sim \frac{e^2}{a_0^3} \tag{359}$$

on dimensional grounds (recall, that $V_c \sim 1/r$, and that a_0 is the characteristic length scale of the problem). Thus, the fine–structure splitting is typically of order $e^2/(m_e^2 a_0^3)(\hbar/c)^2$, which is to be compared with the Balmer splittings of order e^2/a_0. To cast this into a form useful for quick numerical estimates, remember that the ratio of classical electron radius, Compton wavelength of the electron and Bohr radius is

$$\frac{e^2}{m_E c^2} : \frac{\hbar}{m_e c} : a_0 = 1 : 137 : 137^2 \, , \tag{360}$$

where we have used, that

$$\frac{e^2}{\hbar c} = \frac{1}{137} \, . \tag{361}$$

Thus the ratio of fine structure splittings and Balmer splittings are of the order of

$$\left(\frac{e^2}{a_0^3} \frac{\hbar^2}{m_e^2 c^2} \right) : \frac{e^2}{a_0} = \left(\frac{1}{137} \right)^2 : 1 \tag{362}$$

explaining immediately the term *fine structure*.
Note in passing, that relativistic mass effects are roughly of the same order numerically, just consider the expansion

$$\sqrt{m_e^2 c^4 + p^2 c^2} - m_e c^2 \sim \frac{p^2}{2m_e} - \frac{p^4}{8 m_e^3 c^2} \, . \tag{363}$$

However, to give another, more concise example of the effect of the fine structure splitting, consider a sodium atom with ground state configuration

$$(1s)^2 (2s)^2 (2p)^6 (3s)^1 \, , \tag{364}$$

i.e. hydrogen–like. We will assume, that the inner ten electrons form a kind of a spherically symmetric cloud in a manner, that the original pure Coulomb form of the potential is somewhat modified. Let's think now about the transition of the outermost $3s$–electron to $3p$. This $3p$ state already is split off the $3s$–state due to the modifications in the potential. However, the spin–orbit coupling induces an even finer level splitting between the $3p_{1/2}$ and $3p_{3/2}$–states, where the subscript labels j. This splitting has been confirmed experimentally. In fact, it refers to the two sodium D–lines, one at 5896 A and one at 5890 A. Here, $3p_{3/2}$ lies higher, since the corresponding radial integral in Eq.356 is positive.

5.3.2 Lifting degeneracies II : The Zeeman effect

Let us now repeat the procedure we already encountered when discussing the linear Stark effect, namely the application of an external field. But in contrast to the treatment of the Stark effect we would like to investigate the impact of uniform magnetic fields leading to the (anomalous) Zeeman effect. Using the fact, that an uniform magnetic field can be derived from a vector potential

$$\vec{A} = \frac{1}{2}\vec{r} \times \vec{B} = -\frac{1}{2}(By\vec{e}_x - Bx\vec{e}_y) \tag{365}$$

for a field $\vec{B} = B\vec{e}_z$ in z–direction we employ, that the substitution (this is in fact the so–called minimal coupling of the vector potential)

$$\vec{p} \to \vec{p} - e\vec{A} \tag{366}$$

leads to the gauge invariant construction of a suitable Hamiltonian, we find

$$H = \frac{\vec{p}^2}{2m_e} + V_c(r) - \frac{e}{2m_e}(\vec{p}\cdot\vec{A} + \vec{A}\cdot\vec{p}) + \frac{e^2\vec{A}^2}{2m_e}\,. \tag{367}$$

Whenever $\vec{\nabla}\vec{A} = 0$, which is the case for our vector potential, we are free to replace $\vec{p}\vec{A}$ with $\vec{A}\vec{p}$. Noting, that with

$$\vec{A}\vec{p} = B\frac{xp_y - yp_x}{2} = \frac{BL_z}{2} \tag{368}$$

and

$$\vec{A}^2 = \frac{B^2(x^2+y^2)}{4} \tag{369}$$

we obtain, adding the spin–orbit coupling,

$$H = \frac{\vec{p}^2}{2m_e} + V_c(r) - \frac{e}{2m_e}BL_z + \frac{e^2}{8m_e}B^2(x^2+y^2) - \frac{e}{m_e}BS_z\,. \tag{370}$$

Neglecting the quadratic term $\sim B^2$, we can rewrite the Hamiltonian as three distinct terms, namely

$$
\begin{aligned}
H_0 &= \frac{p^2}{2m_e} + V_c(r) \\
H_{LS} &= \frac{1}{2m_e^2 r}\frac{dV_c(r)}{dr}\vec{L}\cdot\vec{S} \\
H_B &= -\frac{eB}{2m_e}(L_z + 2S_z)\,. \qquad (371)
\end{aligned}
$$

The factor two in front of the spin operator S_z merely reflects the fact, that the g–factor of the electron is 2. Treating B as small perturbation – consistently we omitted the B^2–term – we should better study the effect of the H_B–part using the suitable eigenvectors for the "unperturbed" problem, namely the eigenvectors of J^2 and J_z. Identifying

$$L_z + 2S_z = J_z + S_z \qquad (372)$$

the energy shift due to the H_B term to first order is given by

$$\Delta_B = -\frac{eB}{2m_e}\langle J_z + S_z\rangle_{j=l\pm 1/2,m}\,. \qquad (373)$$

According to the definition of its eigenvectors and –values, the expectation value for J_z is m, and we just have to care for the expectation value $\langle S_z\rangle$. For this, recall, that

$$
\begin{aligned}
\left|j = l\pm\frac{1}{2}, m\right\rangle &= \pm\sqrt{\frac{l\pm m+\frac{1}{2}}{2l+1}}\left|m_l = m-\frac{1}{2}, m_s = \frac{1}{2}\right\rangle \\
&\quad +\sqrt{\frac{l\mp m+\frac{1}{2}}{2l+1}}\left|m_l = m+\frac{1}{2}, m_s = -\frac{1}{2}\right\rangle\,. \qquad (374)
\end{aligned}
$$

Then, the expectation value of S_z is easily computed as

$$
\begin{aligned}
\langle S_z\rangle_{j=l\pm 1/2,m} &= \frac{1}{2}\frac{1}{2l+1}\left[\left(l\pm m+\frac{1}{2}\right) - \left(l\mp m+\frac{1}{2}\right)\right] \\
&= \pm\frac{m}{2l+1}\,. \qquad (375)
\end{aligned}
$$

Finally, sticking together the single parts, we obtain the energy shift

$$\Delta_B = -\frac{eB}{2m_e}m\left[1\pm\frac{1}{2l+1}\right]\,. \qquad (376)$$

Obviously, the energy shift is proportional to the magnetic quantum number m, as we already might have guessed, and it dies out, as the orbital angular momentum goes to infinity. This completes our treatment of the various effects destroying the nice symmetry of the hydrogen atom and splits is levels.

6 More group theory

In this section we will introduce some more concepts of group theory, namely the roots and weights and their importance for constructing the representations of arbitrary simple Lie–algebras. For better understanding, it might be worthwhile to read Sec.7 parallel, especially Sec.7.1.

6.1 Roots and weights

We leave again physics for a while with the satisfying feeling, that group theory can be quite a nice tool, and return to mathematics. In fact, we will try now to generalize the highest weight construction outlined above for the case of $SU(2)$ to an arbitrary simple Lie–algebra, i.e. a Lie algebra without non–Abelian subalgebras.
The basic idea for this is quite simple : We divide the generators of the algebra into two sets. One set with hermitian operators to be diagonalized, like J_3 in the case of $SU(2)$, the other set consisting of non–hermitian ladder operators, like the $J_\pm$. So, suppose we have a simple Lie algebra with generators X_a in the irreducible representation D. We now construct a maximal set of M hermitian operators $H_{1,\dots,M}$, so that

- the $H_i = c_{ia}X_a$ are linear combinations of the X_a with real coefficients c_{ia} (corresponding to a suitable choice of basis),
- they are Hermitian operators,
- the H_i commute,

$$[H_i, H_j] = 0\,, \tag{377}$$

- and they are normalized according to

$$\mathrm{Tr}(H_i H_j) = k_D \delta_{ij}\,. \tag{378}$$

We see, that the H_i are in analogy with J_3. Obviously, their eigenvalues can then be used to label any state of the representation D in an unambiguous way. This again is in analogy with $SU(2)$, where the spin j is used to denote states. Consequently, M parameters are needed to label the states completely, and M is therefore dubbed the *rank* of the algebra.
Since the H_i are Hermitian, they can be diagonalized, and we have eigenvectors $|\mu, D\rangle$ with

$$H_i|\mu, D\rangle = \mu_i|\mu, D\rangle\,. \tag{379}$$

These eigenvalues are called *weights*, the vectors with components μ_i are called *weight vectors* and the set of the H_i form the *Cartan subalgebra*. In other words, a Cartan subalgebra is the maximal abelian subalgebra of a Lie–algebra.

6.1.1 Properties of the roots

Now, for every representation of our group we define weights using the same generators H_i. To be more specific, let us turn now to the adjoint representation defined by Eq.101. Beyond this formal definition, we can think of the adjoint representation as a realization of the group in a basis, where the generators themselves become states, i.e. $X_a \to |X_a\rangle$. Scalar products of such states are then given by

$$\langle X_a|X_b\rangle = -\frac{1}{\lambda}\mathrm{Tr}(X_a^\dagger X_b) \tag{380}$$

with normalization λ and we have

$$X_a|X_b\rangle = |[X_a, X_b]\rangle \tag{381}$$

for the action of our generators on the states. Let us now find a convenient basis for this space. But wait. We already know a very nice set of operators (in fact, states, in the adjoint representation), the H_i, with weight zero, yielding

$$H_i|H_j\rangle = |[H_i, H_j]\rangle = 0 \tag{382}$$

(see Eq.377). Now we diagonalize the rest of the space. We find states $|E_\alpha\rangle$ with

$$H_i|E_\alpha\rangle = \alpha_i|E_\alpha\rangle\,, \tag{383}$$

which correspond to generators E_α via

$$[H_i, E_\alpha] = \alpha_i E_\alpha\,. \tag{384}$$

At this stage we would like to stress again, that the H_i and the E_α are linear combinations of the generators X_a of the fundamental representation, and thus states in the adjoint representation. Therefore, the $\alpha_{i=1,\dots,M}$ are components of the weight vectors α in the adjoint representation.
We proceed by noting, that the E_α are not hermitian,

$$[H_i, E_\alpha^\dagger] = -\alpha_i E_\alpha^\dagger \quad\Longrightarrow\quad E_\alpha^\dagger = E_{-\alpha}\,. \tag{385}$$

We can see this property by noting, that the H_i are hermitian and by taking the hermitian conjugate of Eq.384. The weights of the adjoint representation, the α_i are called roots, forming root vectors. Finally, we will normalize

$$
\begin{aligned}
\langle E_\alpha | E_\beta \rangle &= -\frac{1}{\lambda}\mathrm{Tr}(E_\alpha^\dagger E_\beta) = \delta_{\alpha\beta} \\
\langle H_i | H_j \rangle &= -\frac{1}{\lambda}\mathrm{Tr}(H_i H_j) = \delta_{ij} \,.
\end{aligned} \tag{386}
$$

Comparing the second of these equations with Eq.378 we can now identify $k_D = \lambda$, if D is the adjoint representation. In fact, once the normalization λ is fixed, all other k_D are as well. The weights of the adjoint representation, i.e. the α_i are called *roots*.

6.1.2 The E_α as ladder operators

From

$$
H_i E_\alpha |\mu, D\rangle = [H_i, E_\alpha]|\mu, D\rangle + E_\alpha H_i |\mu, D\rangle = (\alpha_i + \mu_i)\, E_\alpha |\mu, D\rangle \tag{387}
$$

we see immediately, that the E_α are raising and lowering operators for the weights. This is in correspondence with the $J_\pm$ of $SU(2)$, where their particlar root vector α in fact is just a number, namely $\alpha = 1$.
We will dwell on this correspondence and derive properties of the roots and weights which are in close analogy to the statement, that in $SU(2)$ the eigenvalues of J_3, i.e. the weights, are half–integers. Let us start with any state $|\mu, D\rangle$. If there is more than one state with weight μ, we will choose our starting state $|\mu, D\rangle$ to be an eigenstate of $E_{-\alpha}E_\alpha$ for some fixed α. We pose this condition on the state, to ensure, that the state is not a linear combination of states with the same μ but from different representations. Let us define a normalization

$$
E_{\pm\alpha}|\mu, D\rangle = N_{\pm\alpha,\mu}|\mu \pm \alpha, D\rangle \,. \tag{388}
$$

Now, turning back to the adjoint representation, we find, that the state $E_\alpha|E_{-\alpha}\rangle$ has weight zero, since

$$
\begin{aligned}
H_i \left[E_\alpha | E_{-\alpha}\rangle\right] &= H_i |[E_\alpha, E_{-\alpha}]\rangle \\
&= |\,[H_i, [E_\alpha, E_{-\alpha}]]\rangle = 0
\end{aligned} \tag{389}
$$

by employing the Jacobi–identity and Eq.384. We can now make the following ansatz for this state

$$
E_\alpha |E_{-\alpha}\rangle = \beta_i |H_i\rangle \,, \tag{390}
$$

which is motivated by the fact, that $[H_i, H_j] = 0$ and identifying $[E_\alpha, E_{-\alpha}] \to H_j$. Multiplying from the left with $\langle H_j|$ and with the normalization condition for the products $\langle H_j | H_i\rangle$ we find for the β_i

$$\beta_i = \langle H_i | E_\alpha | E_{-\alpha} \rangle = \frac{\mathrm{Tr}(H_i [E_\alpha, E_{-\alpha}])}{\lambda} = \frac{\mathrm{Tr}(E_{-\alpha}[H_i, E_\alpha])}{\lambda} = \alpha_i \,. \quad (391)$$

Thus,

$$[E_\alpha, E_{-\alpha}] = \alpha_i H_i \,. \quad (392)$$

To analyze the commutator $[E_\alpha, E_\beta]$ we start from an equation similar to Eq.387,

$$\begin{aligned} H_i[E_\alpha, E_\beta]|\mu, D\rangle &= \{[H_i, E_\alpha] + E_\alpha H_i\} E_\beta |\mu, D\rangle - \{[H_i, E_\beta] + E_\beta H_i\} E_\alpha |\mu, D\rangle \\ &= (\alpha_i + \beta_i + \mu_i) [E_\alpha, E_\beta]|\mu, D\rangle \,. \quad (393) \end{aligned}$$

From this we can conclude, that

$$[E_\alpha, E_\beta] \sim E_{\alpha+\beta} \,. \quad (394)$$

We return now to our representation D. Consider first

$$\begin{aligned} &\langle \mu, D | [E_\alpha, E_{-\alpha}] | \mu, D\rangle \\ &= \langle \mu, D | E_\alpha E_{-\alpha} | \mu, D\rangle - \langle \mu, D | E_{-\alpha} E_\alpha | \mu, D\rangle = |N_{-\alpha,\mu}|^2 - |N_{\alpha,\mu}|^2 \\ &= \alpha \cdot \mu = \langle \mu, D | \alpha_i H_i | \mu, D\rangle \,. \quad (395) \end{aligned}$$

But, we also have

$$\begin{aligned} N_{-\alpha,\mu} &= \langle \mu - \alpha, D | E_{-\alpha} | \mu, D\rangle = \langle \mu - \alpha, D | E_\alpha^\dagger | \mu, D\rangle \\ &= \langle \mu, D | E_\alpha | \mu - \alpha, D\rangle^* = N^*_{\alpha,\mu-\alpha} \quad (396) \end{aligned}$$

resulting in

$$|N_{\alpha,\mu-\alpha}|^2 - |N_{\alpha,\mu}|^2 = \alpha \cdot \mu \,. \quad (397)$$

Since our Lie–algebra is finite–dimensional, we must eventually get zero when applying $E_{\pm\alpha}$ repeatedly on some state $|\mu, D\rangle$. Suppose, this is the case for

$$E_\alpha |\mu + p\alpha, D\rangle = 0 \text{ and } E_{-\alpha} |\mu - q\alpha, D\rangle = 0 \quad (398)$$

with positive integers p and q. Then we might climb down the full ladder,

$$
\begin{array}{lcllcl}
|N_{\alpha,\mu+(p-1)\alpha}|^2 & - & 0 & = & \alpha\cdot[\mu+p\alpha] \\
|N_{\alpha,\mu+(p-2)\alpha}|^2 & - & |N_{\alpha,\mu+(p-1)\alpha}|^2 & = & \alpha\cdot[\mu+(p-1)\alpha] \\
 & & \cdots & & \\
|N_{\alpha,\mu}|^2 & - & |N_{\alpha,\mu+\alpha}|^2 & = & \alpha\cdot[\mu+\alpha] \\
|N_{\alpha,\mu-\alpha}|^2 & - & |N_{\alpha,\mu}|^2 & = & \alpha\cdot\mu \\
 & & \cdots & & \\
|N_{\alpha,\mu-q\alpha}|^2 & - & |N_{\alpha,\mu-(q-1)\alpha}|^2 & = & \alpha\cdot[\mu-(q-1)\alpha] \\
0 & - & |N_{\alpha,\mu-q\alpha}|^2 & = & \alpha\cdot[\mu-q\alpha]
\end{array}
\tag{399}
$$

Summing the differences on the left and the scalar products on the right we find

$$
\begin{aligned}
0 &= (p+q+1)\alpha\cdot\mu + \left(\frac{p(p+1)}{2} - \frac{q(q+1)}{2}\right)\alpha^2 \\
&= (p+q+1)\left(\alpha\cdot\mu + \alpha^2\frac{p-q}{2}\right) .
\end{aligned}
\tag{400}
$$

But this tells us immediately, that

$$
\frac{\alpha\cdot\mu}{\alpha^2} = -\frac{p-q}{2} \tag{401}
$$

is a half–integer. Correspondingly, we are now equipped to determine all $|N|^2$ in terms of μ, α, p, and q. In fact, this holds true for any representation, but we will see, that it has particularly simple and strong consequences for the adjoint representation.
In the adjoint representation, μ is a root, say β with

$$
\frac{\alpha\cdot\beta}{\alpha^2} = -\frac{p-q}{2} = \frac{m}{2} \tag{402}
$$

and integer m. Let us now apply $E_{\pm\beta}$ on $|E_\alpha\rangle$ leaving us with

$$
\frac{\beta\cdot\alpha}{\beta^2} = -\frac{p'-q'}{2} = \frac{m'}{2} \tag{403}
$$

and integer m'. Multiplying Eqs.402 and 403 we find, that

$$
\frac{mm'}{4} = -\frac{(\alpha\cdot\beta)^2}{\alpha^2\beta^2} \equiv \cos^2\theta\,, \tag{404}
$$

where θ is the angle between the root vectors. But this tells us, that only a limited set of angles between root vectors is allowed at all, namely

$$
\begin{array}{ll}
mm' & \theta \\
0 & \pi/2 \\
1 & \pi/3\,,\ 2\pi/3 \\
2 & \pi/4\,,\ 3\pi/4 \\
3 & \pi/6\,,\ 5\pi/6 \\
4 & 0\,,\ \pi\,.
\end{array}
\tag{405}
$$

So far we were assuming implicitly, that for every non–zero root vector α, there is some unique generator E_α. Now we are in the position to prove this. Suppose there are two generators E_α and E'_α associated with the same root. We can choose them to be orthogonal,

$$
\langle E_\alpha | E'_\alpha \rangle \sim \mathrm{Tr}(E_\alpha E'_\alpha) = 0\,. \tag{406}
$$

Applying now $E_{\pm\alpha}$ on the state $|E'_\alpha\rangle$ we find

$$
\frac{\alpha\cdot\alpha}{\alpha^2} = -\frac{p-q}{2}\,. \tag{407}
$$

But since

$$
\begin{aligned}
E_{-\alpha}|E'_\alpha\rangle &= \beta_i |H_i\rangle \\
\text{with } \beta_i &= \langle H_i|E_{-\alpha}|E'_\alpha\rangle \sim \mathrm{Tr}(H_i[E_{-\alpha},E'_\alpha]) \\
&= \mathrm{Tr}(E'_\alpha[H_i,E_{-\alpha}]) = -\alpha_i \mathrm{Tr}(E_{-\alpha}E'_\alpha) = 0
\end{aligned}
\tag{408}
$$

we can conclude, that

$$
E_{-\alpha}|E'_\alpha\rangle \;=\; 0 \tag{409}
$$

and therefore $q = 0$. Ergo,

$$
\frac{\alpha\cdot\alpha}{\alpha^2} = -\frac{p}{2} < 0\,. \tag{410}
$$

But this is impossible, and we see, that E_α is unique.

6.1.3 Connection to $SU(2)$

The steps leading to Eq.401 are essentially the same we did in Sec.3.2.2 to determine the representation of $SU(2)$. This connection might become even clearer in the following.
First, let us define rescaled operators

$$
\begin{aligned}
E_{\pm} &= \frac{1}{|\alpha|} E_{\pm\alpha} \\
E_3 &= \frac{1}{|\alpha|^2} \alpha_i H_i
\end{aligned} \tag{411}
$$

with $|\alpha|^2 = \alpha \cdot \alpha$. Using Eqs.384 385, and 392, it is easy to show that these operators satisfy the $SU(2)$ algebra,

$$
[E_3, E_{\pm}] = \pm E_{\pm} \, , \;\; [E_+, E_-] = E_3 \, . \tag{412}
$$

Thus, each E_α is associated with an $SU(2)$ subalgebra, comprising E_α and its adjoint $E_{-\alpha}$ and a linear combination of Cartan operators H_i. If $|\mu, D\rangle$ is an eigenstate of $E_{-\alpha} E_\alpha$, then the states $|\mu + k\alpha, D\rangle$ obtained by multiple application of the ladder operators $E_{\pm\alpha}$ form an irreducible representation of the $SU(2)$ subalgebra of Eq.412. The condition, that $|\mu, D\rangle$ is an eigenstate of $E_{-\alpha} E_\alpha$ insures, that this state is not merely a linear combination of states with the same μ from different representations D. However, the newly obtained states satisfy

$$
E_3 |\mu + k\alpha, D\rangle = \frac{\alpha \cdot \mu + k\alpha^2}{|\alpha|^2} |\mu + k\alpha, D\rangle \, . \tag{413}
$$

The highest and lowest eigenvalues of E_3 in the representation corresponding to $|\mu + p\alpha, D\rangle$ and $|\mu - q\alpha, D\rangle$, respectively, must then be

$$
\begin{aligned}
j &= \frac{\alpha \cdot \mu}{\alpha^2} + p \\
-j &= \frac{\alpha \cdot \mu}{\alpha^2} - q \\
0 &= 2 \left[\frac{\alpha \cdot \mu}{\alpha^2} + \frac{p - q}{2} \right] \, .
\end{aligned} \tag{414}
$$

6.2 Simple roots and fundamental representations

6.2.1 Generalization of the highest weight

We are now equipped to generalize our concept of highest weight already encountered in our treatment of $SU(2)$ for arbitrary Lie groups. The basic idea is to define "positivity" for weight vectors. After fixing the Cartan basis, i.e. the H_1, H_2, $\ldots$, the components of our weight vectors μ_1, μ_2, $\ldots$ are fixed, too. Then we call a weight vector positive (negative), it its first non–zero component is positive (negative). Consequently it is dubbed zero, if all components are zero. This is an arbitrary and somewhat sloppy division of the space into two halves, and intuitively it is not clear at all, that such a

definition is any good. In fact, this definition will turn out to be surprisingly fruitful, allowing us to define raising and lowering operators. The alert reader might now worry, if this arbitrary ordering does not induce any prejudices into our notation and spoil the full story. Well, in the proceeding of our treatment, we will show, that this is not the case. Even more, we will see, that any definition of positivity will do the job.
Now, after introducing positivity, we can talk about some ordering. In fact, we call $x > y$ if $x - y > 0$, i.e. positive. Now, obviously, the highest weight of any representation is the weight, which is larger than any other weight. In our framework of positivity its the weight with the largest non–zero first component. Again, the highest weight is of special importance, as was the case for $SU(2)$. It will turn out, that it is the same reason, that makes it unique : Starting from the highest weight we can construct the whole representation.
We can apply this concept to the roots, the weights of the adjoint representation. They are either positive or negative and they correspond to raising or lowering operators, respectively. Obviously, if we act on the state of highest weight in any representation with some of the raising operators $E_{\alpha>0}$, we must get zero.
A *simple root* is a positive root, which cannot be decomposed as a sum of two positive roots. The simple roots play a special role, since they determine the whole structure of the group. We will see this in the following.

- If α and β are simple roots, then $\beta - \alpha$ is not a root.

  Proof : Assume $\beta - \alpha > 0$ is a root. Then $\beta = \alpha + (\beta - \alpha)$, the sum of two positive roots, ergo β is not simple. The same reasoning holds true for $\alpha - \beta > 0$ and a root.

  We conclude, that, since $\beta - \alpha$ is not a root, $E_{-\alpha}|E_\beta\rangle$ is not a weight in the adjoint representation, i.e. no root. Therefore, $E_{-\alpha}|E_\beta\rangle = 0$, and we have

$$\frac{\alpha \cdot \beta}{\alpha^2} = -\frac{p-q}{2} = -\frac{p}{2}, \tag{415}$$

  see Eq.401. Remember, q equals zero here, because there are exactly zero times $E_{-\alpha}$ has to be applied to annihilate $|E_\beta\rangle$, if both are simple roots.

- Thus, knowing all the integers p is equivalent to knowing all the relative angles between simple roots and their relative length :

$$\frac{2\alpha \cdot \beta}{\alpha^2} = -p\,, \quad \frac{2\alpha \cdot \beta}{\beta^2} = -p' \tag{416}$$

  implies

$$\cos\theta_{\alpha\beta} = -\frac{1}{2}\sqrt{pp'}\,, \quad \frac{\beta^2}{\alpha^2} = \frac{p}{p'}\,. \tag{417}$$

Since p,$p' > 0$, the angle between any two simple roots satisfies

$$\frac{\pi}{2} \leq \theta < \pi\,. \tag{418}$$

Simple geometry tells us, that any set of vectors with this constraint is linearly independent.

Proof : If they were not, we could find constants x_α with

$$\sum_\alpha x_\alpha \alpha = 0 \tag{419}$$

for the simple roots α. We could divide them now into two sets, $\Gamma_\pm$ with $\alpha \in \Gamma_+$ if $x_\alpha > 0$ and vice versa. Then,

$$y \equiv \sum_{\alpha\in\Gamma_+} x_\alpha \alpha = \sum_{\alpha\in\Gamma_-} (-x_\alpha)\alpha \equiv z\,, \tag{420}$$

where now all x_α are non–negative. Both y and z are then positive (or at least non–zero) vectors by construction. But this is impossible, since

$$y^2 = y \cdot z \leq 0\,, \tag{421}$$

which follows from $\alpha \cdot \beta \leq 0$ for any two simple roots α and β.

- Consequently, any positive root ϕ can be written as a sum of simple roots with non–negative integer coefficients k_α.

  Proof : By induction. If ϕ is simple, we are done. If not, we can split ϕ into two positive roots ϕ_1 and ϕ_2. If they are not simple we can continue and split them again.

- The number of simple roots is equal to the rank of the group M.

  Proof : Because the roots are vectors with M components, there can not be more than M linearly independent of their kind. Now assume, we have less than M simple roots, so that they do not span the full space of roots. Then, we can choose a basis, in which the first component of all simple roots α vanish. But this means, that the first component of every root vanishes. This in turn has the consequence, that

$$[H_1, E_\phi] = 0 \text{ for all roots } \phi\,. \tag{422}$$

But wait. Since $[H_1, H_i] = 0$ by definition of the Cartan subalgebra, H_1 commutes with everything – it is an invariant subalgebra in its own rights ! This is in contradiction to our assumption, that the group we are dealing with, is a simple Lie group.

- If we are capable to figure out, which of the linear combinations

$$\sum_\alpha k_\alpha \alpha\ , \quad k_\alpha \text{ nonnegative integers} \tag{423}$$

are roots, we can determine all the roots of a group. We can do this by induction. For $k = \sum k_\alpha = 1$, the sums are just the simple roots. The case $k = 2$ can be constructed by acting on the $k = 1$ configurations with the generators, which correspond to the simple roots. We then use Eq.401 to check, whether we found a new root or annihilated the state. For instance, $\alpha + \beta$ is a root, if $\alpha \cdot \beta < 0$, because then

$$\frac{2\alpha \cdot \beta}{\alpha^2} = -p \quad \Longrightarrow \quad p > 0\,, \tag{424}$$

where p is the largest integer, so that $\alpha + p\beta$ is a root.

In this specific case, we knew, that $q = 0$, because for $k = 2$, α and β are simple roots. But for more complicated cases, $k = n$ plus the action of a simple root α on it leading to $k = n + 1$, more care has to be taken. In principle, q is not necessarily zero, but we can determine it by examining the roots with $k < n$. Knowing q, we can fix p with help of our beloved Eq.401,

$$\frac{2\alpha \cdot \gamma}{\alpha^2} = -(p - q)\,. \tag{425}$$

If $p \neq 0$, $\gamma + \alpha$ is a root. In this manner we can find all roots associated with $k = n + 1$, which can be written as sums $\gamma + \alpha$, where γ stems from the step $k = n$ and α is simple. But that's it already for $k = n + 1$. There are no roots different from $\gamma + \alpha$.

Proof : Suppose, there's some root ρ, which does not have the desired form $\gamma + \alpha$. Then, since $\rho - \alpha$ is not a root, $q = 0$, and we have

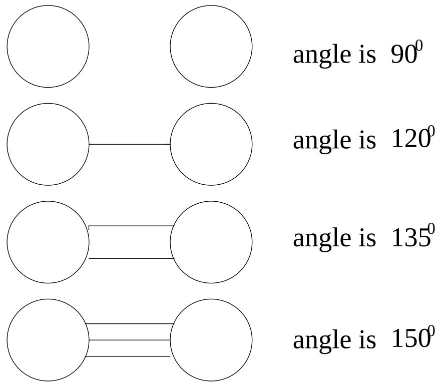


Figure 1: The rules for the angles between simple roots for Dynkin diagrams.

$$\frac{2\alpha \cdot \rho}{\alpha^2} = -p < 0 \tag{426}$$

and ρ is linearly independent of all simple roots α leading to a contradiction.

This last point is quite instructive. We have determined all roots inductively starting from the simple roots by using only lengths and angles between them and by employing our master formula Eq.401. Obviously, the full treatment is independent of the basis in root space and of the specific notion of positivity.

6.2.2 Dynkin diagrams

Let us remember, that the simple roots are all, what we really need to know in order to construct a representation of the group. The knowledge of the simple roots can be used to construct all roots, which we can use in turn to fix the normalization factors $N_{\alpha,\beta}$ of Eq.388 and thus describe the full algebra.
The treatment of the preceeding section allows us to construct a diagrammatic shorthand for pinning down the simple roots. This shorthand consists of open circles, representing the simple roots, which are connected by lines. The number of lines linking two circles tells us about the angle between the simple roots as in Fig.1.
As examples consider the Dynkin diagrams for $SU(2)$ and $SU(3)$ as displayed in Fig.2.

6.2.3 Fundamental weights

Let us label the simple roots as α^κ with $\kappa = 1, \ldots, M$. Some weight μ of an arbitrary irreducible representation, D, is the highest weight, if and only if

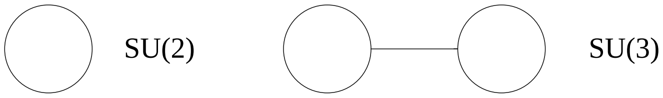


Figure 2: Dynkin diagrams for $SU(2)$ and $SU(3)$.

$\mu + \phi$ is not a weight in this representation for all positive roots ϕ. Recalling the decomposition of positive roots in terms of simple roots, it is then sufficient to claim, that $\mu + \alpha^\kappa$ for any κ is not a weight. This means, that acting with E_{α^κ} on $|\mu\rangle$ we must have $p = 0$ and ergo

$$\frac{2\alpha^\kappa \cdot \mu}{\alpha^{\kappa\,2}} = q^\kappa \tag{427}$$

with non–negative integers q^κ. But this helps us to determine μ, since the α^κ are linearly independent. So every set of q^κ gives us the highest weight of some irreducible representation of our group, and starting from the highest weight we can apply the lowering operators $E_{-\alpha^\kappa}$ to construct the entire representation.

What does this mean for practical purposes ? To gain some insight let us consider weight vectors μ^λ with

$$\frac{2\alpha^\kappa \mu^\lambda}{\alpha^{\kappa\,2}} = \delta^{\kappa\lambda} \,. \tag{428}$$

Thus, any μ^λ is the highest weight of an irreducible representation with $q^\lambda = 1$ and $q^{\kappa\neq\lambda} = 0$. Therefore, any highest weight μ can be written as linear combination

$$\mu = \sum_\kappa q^\kappa \mu^\kappa \tag{429}$$

and we see, that we can construct representations with arbitrary highest weight μ by gluing together representations with highest weight μ^1, μ^2 and so on. This is in full analogy to our way to construct spin–$n/2$ representations of $SU(2)$ by sticking together n spin–$1/2$ representations. The vectors μ^κ are dubbed *fundamental weights* and the m irreducible representations with highest weights given by the fundamental weights are called the *fundamental representations.*

In the treatment above, one big caveat is in order : Please, do not confuse the lower indices which label the components of all the weight and root vectors with the upper indices labeling the simple roots and fundamental representations. Both indices run from 1 to M, but the first ones are vector components and the latter ones just labels for different vectors.

7 More fun with $SU(3)$

7.0.4 Structure of the Lie–algebra

$SU(3)$ is the group of 3×3 unitary matrices with determinant 1. The generators of this group are the 3×3 traceless hermitian matrices. The tracelessness guarantees hermiticity and, that the determinant in fact equals one. The standard physics basis of $SU(3)$ consists of the so–called *Gell–Mann* λ–matrices

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\
\lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \tfrac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{aligned} \tag{430}$$

The generators are given by $T_a = \lambda_a/2$. They are normalized according to

$$\mathrm{Tr}(T_a T_b) = \frac{1}{2}\delta_{ab}\,. \tag{431}$$

Notice that T_1, T_2, and T_3 form an $SU(2)$ subgroup of $SU(3)$ which is dubbed the isospin subgroup for obvious reasons.

7.0.5 Introducing $SU(3)$: The isotropic harmonic oscillator

To become acquainted with $SU(3)$ we will consider first the isotropic harmonic oscillator in three dimensions. Its Hamiltonian is given by

$$\mathcal{H} = \frac{\vec{p}^2}{2m} + \frac{k\vec{r}^2}{2} = \frac{1}{2m}\sum_{j=1}^{3}(p_j^2 + m^2\omega^2 r_j^2) \tag{432}$$

A first naive guess is, that its symmetry group is $O(3)$. But as we already had to learn in the case of the hydrogen atom, first guesses often enough miss the truth. In fact, as we will learn, the symmetry group of the three–dimensional isotropic harmonic oscillator is $SU(3)$. To see this, we switch to the common raising and lowering operators

$$
\begin{aligned}
a_j &= \frac{p_j - i\omega m r_j}{\sqrt{2m\omega\hbar}} \\
a_j^\dagger &= \frac{p_j + i\omega m r_j}{\sqrt{2m\omega\hbar}} .
\end{aligned}
\tag{433}
$$

Their commutation relations are

$$
[a_i, a_j^\dagger] = \delta_{ij} \ , \ \ [a_i, a_j = [a_i^\dagger, a_j^\dagger] = 0 \, , \tag{434}
$$

and they follow directly from the commutation relations for the position and momentum operators. In fact, they are merely generalizations of the commutation relations for one dimension to the ones for three dimensions. With help of these operators, the Hamiltonian can be rewritten as

$$
\mathcal{H} = \hbar\omega \sum_j \left(a_j^\dagger a_j + \frac{1}{2} \right) = \frac{\hbar\omega}{2} \sum_j \left\{ a_j^\dagger, a_j \right\} \ , \tag{435}
$$

where the anticommutator $\{a, b\} = ab + ba$. The commutators of the raising and lowering operators with the Hamiltonian are

$$
\left[\mathcal{H}, a_i^\dagger \right] = \hbar\omega a_i^\dagger \ , \ \ [\mathcal{H}, a_i] = -\hbar\omega a_i \tag{436}
$$

explaining their names. The occupation number operator

$$
\mathcal{N}_i = a_i^\dagger a_i \tag{437}
$$

has integer eigenvalues $n_i \geq 0$. The eigenvalues of the Hamiltonian are thus

$$
E_n = \left(n_1 + n_2 + n_3 + \frac{3}{2} \right) \hbar\omega \, , \tag{438}
$$

see Eq.435. After some bookkeeping we count $(n+1)(n+2)/2$ ways, in which any positive integer number n can be divided into three non–negative integers n_1, n_2,and n_3. Therefore, for every energy level labeled by the main quantum number n, we find $(n+1)(n+2)/2$ states. We construct now the angular momentum operator. In its components it is given by

$$
L_j = (\vec{r} \times \vec{p})_j = \frac{i\hbar}{2} \sum_{k,l=1}^{3} \epsilon_{jkl} (a_k a_l^\dagger - a_k^\dagger a_l) \, . \tag{439}
$$

It can be shown, that in general, operators of the form $a_i^\dagger a_j$ commute with $\mathcal{H}$, and so does the angular momentum operator. The physical effect of such operators $a_i^\dagger a_j$ is to annihilate a quantum in the j–direction and transfer it to the i–direction. It therefore leaves the number of quanta fixed. Therefore, the total quantum number operator

$$\mathcal{N} = \sum_{j=1}^{3} a_j^\dagger a_j = \frac{\mathcal{H}}{\hbar\omega} - \frac{3}{2} \tag{440}$$

commutes with all nine operators of the form $a_i^\dagger a_j$. But from these we might switch to some basis, which is more convenient for investigations. Starting with $\mathcal{N}$, we can construct eight other independent linear combinations of the $a_i^\dagger a_j$, namely

$$\begin{array}{lcllcl}
\lambda_1 & = & a_1^\dagger a_2 + a_2^\dagger a_1\,, & \lambda_2 & = & -i(a_1^\dagger a_2 - a_2^\dagger a_1)\,, \\
\lambda_3 & = & a_1^\dagger a_1 - a_2^\dagger a_2\,, & \lambda_4 & = & a_1^\dagger a_3 + a_3^\dagger a_1\,, \\
\lambda_5 & = & -i(a_1^\dagger a_3 - a_3^\dagger a_1)\,, & \lambda_6 & = & a_2^\dagger a_3 + a_3^\dagger a_2\,, \\
\lambda_7 & = & -i(a_2^\dagger a_3 - a_3^\dagger a_2)\,, & \lambda_8 & = & \frac{1}{\sqrt{3}}(a_1^\dagger a_1 + a_2^\dagger a_2 - 2a_3^\dagger a_3)\,.
\end{array} \tag{441}$$

From this we see immediately, that the dynamical symmetry group of the three–dimensional isotropic harmonic oscillator is in fact $SU(3)$ as already advertised.

7.1 Representations of $SU(3)$

7.1.1 Roots and weights of $SU(3)$

We start our analysis of the representations of $SU(3)$ by choosing the Cartan subalgebra. It will turn out, that a convenient choice is the generator T_3. The only other generator commuting with T_3 is T_8. So we have

$$H_1 = T_3 \text{ and } H_2 = T_8 \tag{442}$$

in the language of Sec.6.1. Accordingly, $SU(3)$ is rank 2. A good reason for this choice is, that Mr. Gell–Mann has already diagonalized these two operators for us. The eigenvectors and the corresponding states labeled by their weights $|\mu_1, \mu_2\rangle$ are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rightarrow \left|\frac{1}{2}, \frac{1}{2\sqrt{3}}\right\rangle$$

$$
\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rightarrow \left| -\frac{1}{2}, \frac{1}{2\sqrt{3}} \right\rangle
$$

$$
\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rightarrow \left| 0, -\frac{1}{\sqrt{3}} \right\rangle . \tag{443}
$$

In the μ_1–μ_2 plane the weight vectors form an equilateral triangle ! This can be seen in Fig.3
We know, that the root vectors E_α act as ladder operators on the weights, taking us from one weight to the other. So we can be certain, that the roots in the μ_1–μ_2 plane are differences of the weights. In fact they are $(1,0)$, $(1/2, \sqrt{3}/2$, $(-1/2, \sqrt{3}/2)$, and minus these vectors, see Fig.4.
To find the corresponding generators we look for matrices that take us from one state to another. It is not too hard to guess, that these matrices are the ones with only one single off–diagonal element. We can look them up easily from the matrices of Eq.430,

$$
\begin{aligned}
E_{\pm 1,0} &= \frac{1}{\sqrt{2}}(T_1 \pm iT_2) \\
E_{\pm 1/2, \pm\sqrt{3}/2} &= \frac{1}{\sqrt{2}}(T_4 \pm iT_5) \\
E_{\mp 1/2, \pm\sqrt{3}/2} &= \frac{1}{\sqrt{2}}(T_6 \pm iT_7) .
\end{aligned} \tag{444}
$$

These roots form a regular hexagon in the μ_1–μ_2 plane. We find, that all angles in $SU(3)$ are multiples of $\pi/3$.

7.1.2 Simple roots and fundamental representations

Taking a closer look on Eq.444, we see immediately, that
$(1,0) = (1/2, \sqrt{3}/2) + (1/2, -\sqrt{3}/2)$ is the sum of the two other positive roots. Thus, the simple roots are

$$
\alpha^1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right) \ , \ \alpha^2 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2} \right) \tag{445}
$$

with

$$
{\alpha^1}^2 = {\alpha^2}^2 = 1 \, , \ \ \alpha^1 \cdot \alpha^2 = -\frac{1}{2} \tag{446}
$$

leaving us with

$$
\frac{2\alpha^1 \cdot \alpha^2}{{\alpha^1}^2} = \frac{2\alpha^1 \cdot \alpha^2}{{\alpha^2}^2} = -1 = -(p-q) = -p \, . \tag{447}
$$

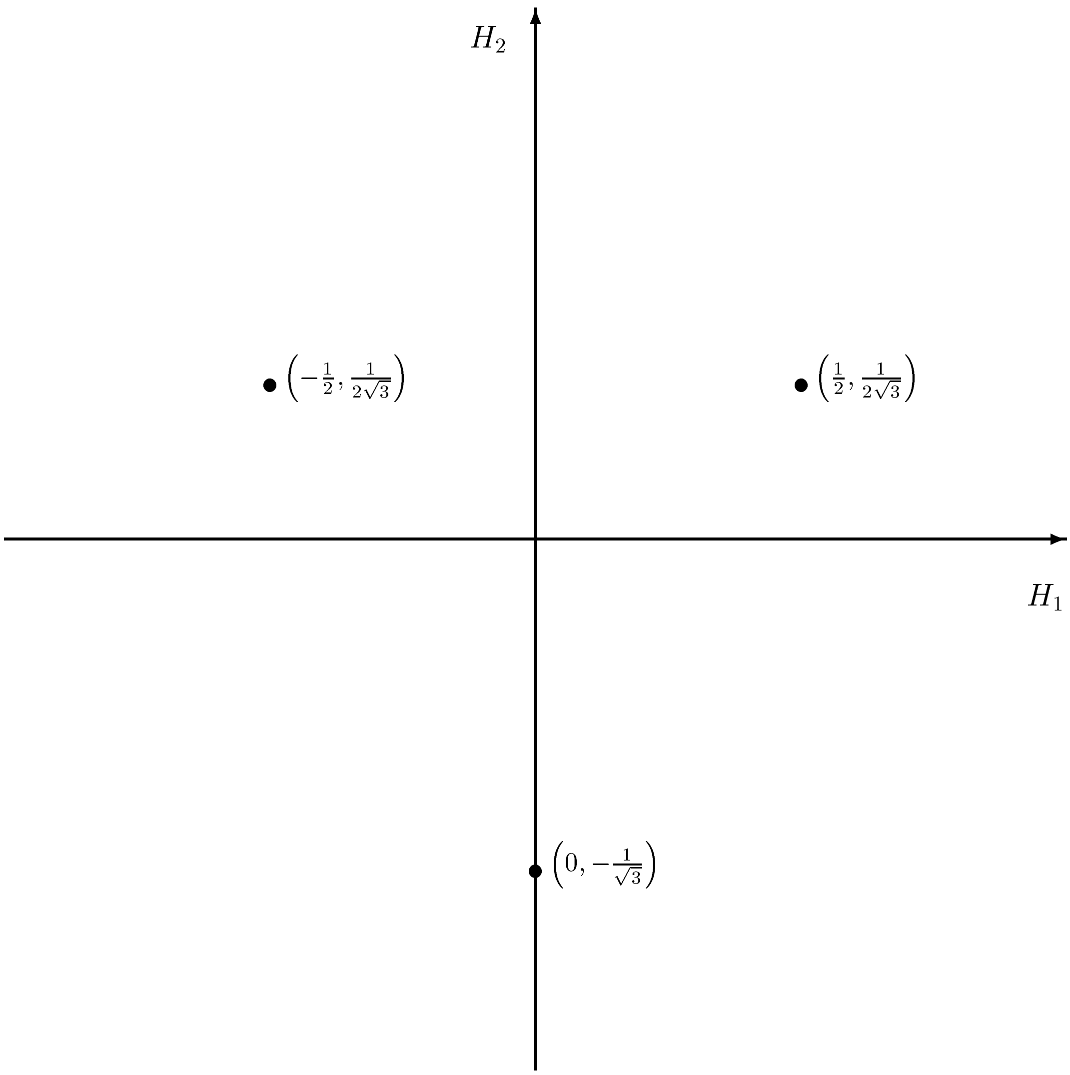


Figure 3: Weights of $SU(3)$

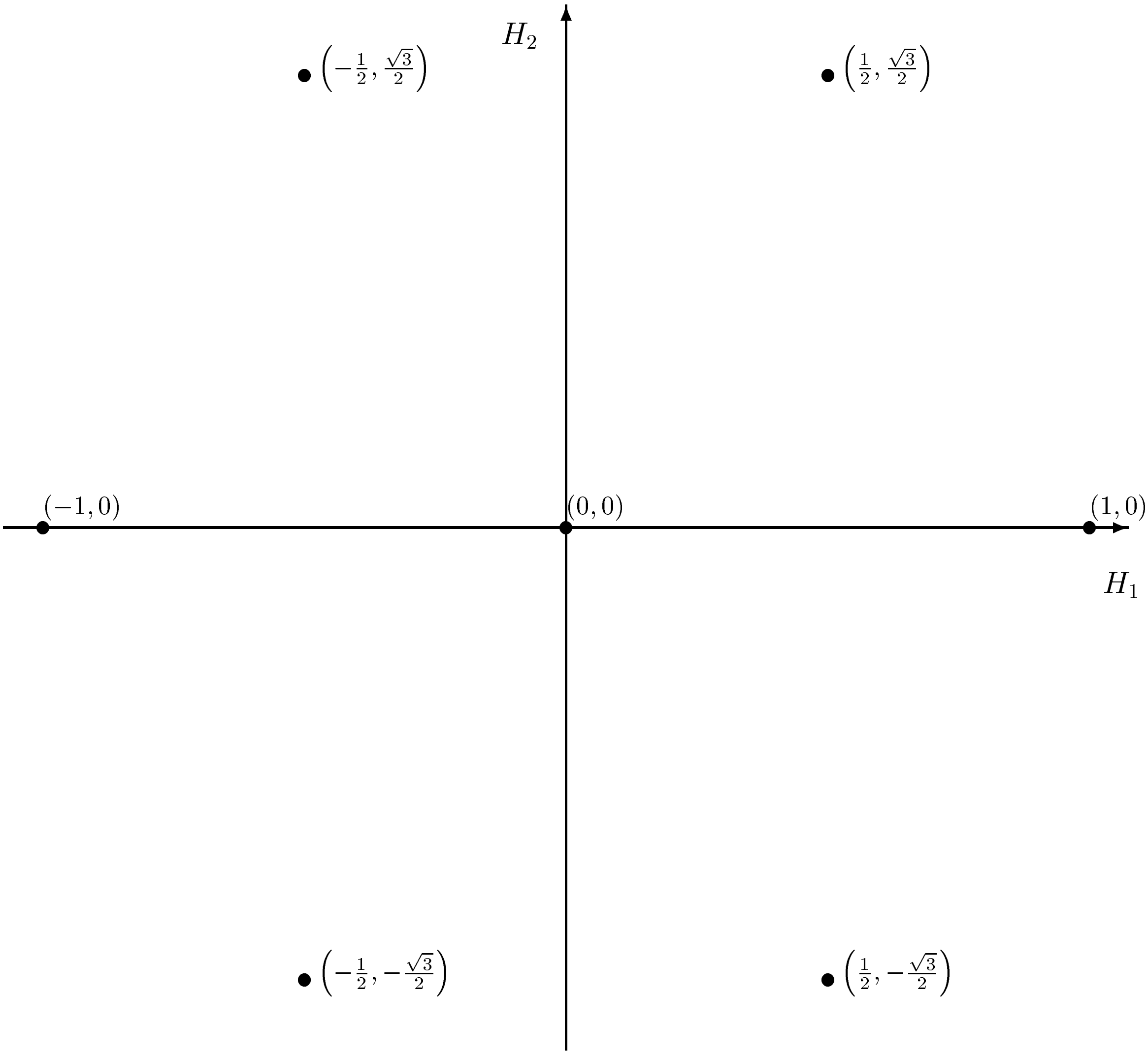


Figure 4: Roots of $SU(3)$

To illustrate the meaning of the previous chapter, consider $E^1_\alpha|E^2_\alpha\rangle$. Because $p=1$, $\alpha^1+\alpha^2$ is a root, but for example $2\alpha^1+\alpha^2$ is not.
From the α^i we obtain the fundamental weights

$$\mu^1 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \ , \ \ \mu^2 = \left(\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right) \ . \tag{448}$$

μ^1 is the highest weight of the defining representation of $SU(3)$ presented before, generated by the T_a matrices of Eq.430. In turn, μ^2 is the highest weight of a different representation, which we will construct in the following. In principle, we could find this representation with a trick, but before we talk about tricks, let us do something more concise. We will build first all states obtained by applying lowering operators on the state with the highest weight, $|\mu^2\rangle$. Since for this fundamental representation $\mu^1=0$ and $\mu^2=1$, it follows, that $\mu^2-\alpha^2$ is a weight, but $\mu^2-2\alpha^2$ and $\mu^2-\alpha^1$ are not. Therefore, there is a state

$$|\mu^2-\alpha^2\rangle \sim E_{-\alpha^2}|\mu^2\rangle \,, \tag{449}$$

which is annihilated when acting on it again with $E_{-\alpha^2}$, since $\mu^2-2\alpha^2$ is not a weight. Acting on it with $E_{-\alpha^1}$ we compute

$$\frac{2\alpha^1\cdot(\mu^2-\alpha^2)}{\alpha^1\cdot\alpha^1} = \frac{-2\alpha^1\cdot\alpha^2}{\alpha^1\cdot\alpha^1} = 1 = q-p\,, \tag{450}$$

where we used the explicit form of the roots and the fact, that $\alpha^1\cdot\mu^2 \sim \delta^{12}$, cf. Eq.428. To continue, we fix p first. Remember, that p labels the number of times, a raising operator can be applied on a state without annihilating it. So let us consider the effect of E_{α^1} on the state $|\mu^2-\alpha^2\rangle$. We note, that $\mu^2-\alpha^2+\alpha^1$ is not a weight, since

$$E_{\alpha^1}E_{-\alpha^2}|\mu^2\rangle = E_{-\alpha^2}E_{\alpha^1}|\mu^2\rangle = 0\,, \tag{451}$$

because $\alpha^1-\alpha^2$ is not a root. Therefore, $p=0$ in Eq.450, and thus $q=1$ and $\mu^2-\alpha^2-\alpha^1$ is a weight, corresponding to the state

$$|\mu^2-\alpha^2-\alpha^1\rangle \sim E_{-\alpha^1}E_{-\alpha^2}|\mu^2\rangle \,. \tag{452}$$

This state gets annihilated when acting on it with either $E_{-\alpha^1}$ or $E_{-\alpha^2}$. To see this, we just have to follow the same path outlined above and use similar arguments. Consider first the effect of E_{α^1} on the state. We find in analogy to Eq.450

$$\frac{2\alpha^1 \cdot (\mu^2 - \alpha^2 - \alpha^1)}{\alpha^1 \cdot \alpha^1} = \frac{-2\alpha^1 \cdot (\alpha^1 + \alpha^2)}{\alpha^1 \cdot \alpha^1} = -1 = q - p\,. \tag{453}$$

Thus, $p = 1$ (remember, $\mu^2 - \alpha^2$ is a weight), and $q = 0$. Hence, $E_{-\alpha^1}$ indeed annihilates $|\mu^2 - \alpha^2 - \alpha^1\rangle$. Similarly,

$$\frac{2\alpha^2 \cdot (\mu^2 - \alpha^2 - \alpha^1)}{\alpha^2 \cdot \alpha^2} = 0 = q - p\,. \tag{454}$$

Here, $p = 0$, since $\mu^2 - \alpha^1$ is not a weight, and again $q = 0$ with the same consequences as above.
If we plot the three weights, μ^2, $\mu^2 - \alpha^2$, and $\mu^2 - \alpha^2 - \alpha^1$, of this representation we find the inverted equilateral triangle of Fig.3, see Fig.5.
The question, that naturally arises here, is, whether all three of the weights correspond to unique states in the representation or whether there are some degeneracies, which remain invisible in our neat diagrams. We are going to answer this question in general terms.

7.1.3 Weyl–reflections

To answer the question outlined above, let us consider some arbitrary irreducible representation D with highest weight μ. Let $|\mu\rangle$ be any state with weight μ and form

$$E_{\xi^1} E_{\xi^2} \dots E_{\xi^n} |\mu\rangle\,, \tag{455}$$

where the ξ are any roots. If we take into account all n, these states span the full representation. But, any state with positive ξ can be dropped, because with help of the commutation relations we can move all raising operators to the right, until they annihilate $|\mu\rangle$. Therefore, we can take all ξ to be negative. We can further decompose these negative roots into sums of simple roots with non–positive coefficients, so we only need to consider

$$E_{-\chi^1} E_{-\chi^2} \dots E_{-\chi^n} |\mu\rangle\,, \tag{456}$$

where the χ^i are simple roots. This shows, that the state $|\mu\rangle$ is unique, as are states $E_{-\chi^i}|\mu\rangle$ and $\left[E_{-\chi^i}\right]^m |\mu\rangle$. this shows, that in our example above, $|\mu^2 - \alpha^2\rangle$ in fact is unique. Let us now check, whether the other state is unique. To see this, we have to examine first, what we mean by "obvious symmetry of the representation". This symmetry results from the trivial fact, that in each root direction, there is an $SU(2)$ symmetry group, and that representation of $SU(2)$ are symmetric. How can we see this ?
Let us reconsider Sec.6.1.3, where we made contact between the $E_{\pm\alpha}$ and $SU(2)$. There we found, that each E_α has an associated $SU(2)$ algebra, where

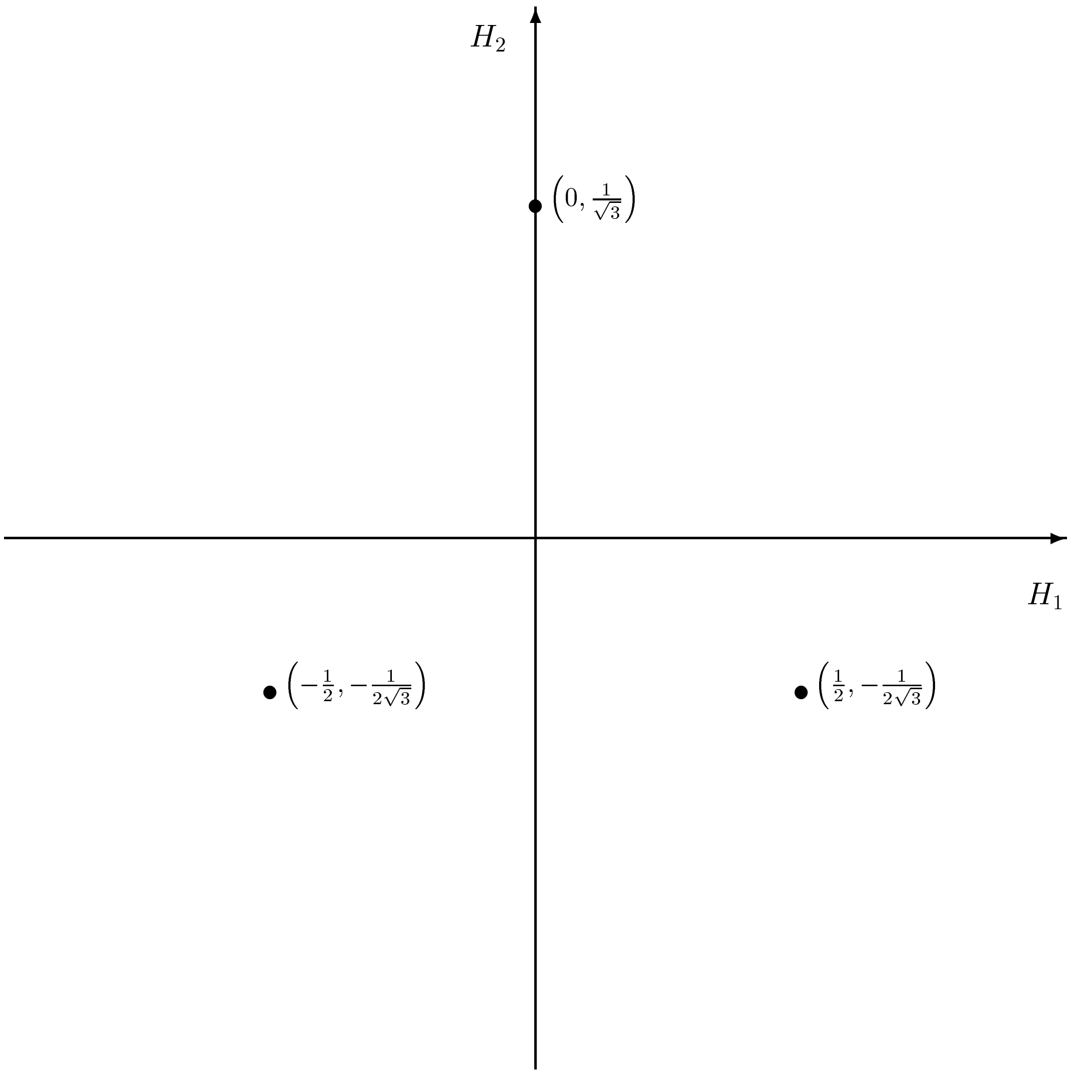


Figure 5: Weights of $SU(3)$ in the $\mu^2 \equiv (0,\,1)$–representation.

the $E_{\pm\alpha}$ form the analogon to the ladder operators $J_\pm$ and some set of operators from the Cartan subalgebra form the analogon to J_3. We also found there, that if $|\mu\rangle$ is a state corresponding to the weight μ and an eigenstate of $E_{-\alpha}E_\alpha$, then the states obtained by multiple application of $E_{\pm\alpha}$ from $|\mu\rangle$, $|\mu+k\alpha\rangle$, form an irreducible representation of $SU(2)$. How can we use this ? Let us take some weight μ and some root α. Then

$$\mu - \frac{2\alpha\cdot\mu}{\alpha^2}\alpha \tag{457}$$

is also a weight. This weight is the reflection of μ in the hyperplane perpendicular to α. Any representation is invariant under all such reflections, which are dubbed *Weyl reflections.* The set of all such reflections for all roots form a group, which is called the *Weyl group* of the algebra.
Let us now return to our example of $SU(3)$. We already learned, that μ^2 and $\mu^2-\alpha^2$ correspond to unique states. We can construct the weight $\mu^2-\alpha^2-\alpha^1$ by a Weyl reflection of μ^2 in the hyperplane perpendicular to $\alpha^2+\alpha^1$ (note, that the root defining the Weyl reflection does not need to be simple):

$$\mu^2 - \frac{2(\alpha^1+\alpha^2)\cdot\mu^2}{(\alpha^1+\alpha^2)^2}(\alpha^1+\alpha^2) = \mu^2-\alpha^2-\alpha^1 \tag{458}$$

by the explicit form of the roots and weight. Since there is only one way to construct the root defining the Weyl refection, the ladder is unique, and so is the state.
This shows, that the second fundamental representation is three–dimensional, as is the first. Its weights are just the negatives of the weights already encountered in the first fundamental representation, as you can read of Figs.5 and 3. This connection is so simple, that there should be a simple trick to go from the first to the second fundamental representation.

7.1.4 Complex representations

And here is the trick. In fact, it is pretty simple. We just have to persuade ourselves, that if T_a are the generators of some representation D of some Lie algebra, then the matrices $-T_a^*$ follow exactly the same algebra. Hence, they also generate a representation, which is obviously of the same dimension as D. It should be not a big surprise, that this representation is called the *complex conjugate representation.* It is denoted by $\bar{D}$. If a representation is equivalent to its complex conjugate it is called *real*, otherwise it is called *complex.*
Since the Cartan generators of $\bar{D}$ are $-H_i^*$, the weight of $\bar{D}$ equal $-\mu$ if μ is a weight in D. Consequently the highest weight μ in D is the negative of the lowest weight in $\bar{D}$ and vice versa. Ergo, if the highest weight of D equals the negative of its lowest weight, D is real.
What does this tell us for our aim to construct the μ^2 representation from the s μ^1 representation ? Quite simple : In the defining (μ^1) representation of $SU(3)$, μ^1 is the highest weight and the lowest is $-\mu^2$. Therefore

$$D_{\mu^2} = \bar{D}_{\mu^1} \,, \tag{459}$$

where the subscripts label the highest weight of the representation. We now know, that the generators of D_{μ^2} are $-T_a^*$, where the T_a are defined via Eq.430.
In physics, there are several notations around to denote representations of $SU(3)$:

1. The most explicit one is to label representations by ordered pairs of integers, (q^1, q^2).

2. Another choice, more common to particle physicists, and quite convenient for small dimensions, is to give just the dimension of the representation and to distinguish between a representation and its complex conjugate with a bar.

Ergo

$$(1,\, 0) = \mathbf{3} \text{ and } (0,\, 1) = \bar{\mathbf{3}}\,. \tag{460}$$

7.1.5 Higher representations

Let us now construct some higher representations of $SU(3)$.

1. $\underline{(2,\, 0) :}$
   This representation is with $q^1 = 2$ and $q^2 = 0$. The highest weight is

   $$2\mu^1 = \left(1,\, \frac{1}{\sqrt{3}}\right) \,, \tag{461}$$

   other weights corresponding to unique states are $2\mu^1 - \alpha^1$ and $2\mu^1 - 2\alpha^1$, as we can conclude from the considerations of Sec.7.1.2. We also already know, that $2\mu^1 - \alpha^2$ is not a weight. The Weyl reflections now allow us to construct other weights, in fact, as we will see, all other weights. They are depicted in Fig.6. We see, that this representation has dimension six, therefore it is sometimes dubbed the **6**-representation of $SU(3)$. Its complex conjugate is $\bar{\mathbf{6}}$, and we will leave it to the reader's fantasy to figure out its weights.

2. $\underline{(1,\, 1) :}$
   Next, we will consider the $(1, 1)$–representation with highest weight $\alpha^1 + \alpha^2$. But wait. We already know this representation ! This is the adjoint representation, whose weights are the roots shown in Fig.4. Note, that the zero weight is doubly degenerate, since there are two ways to

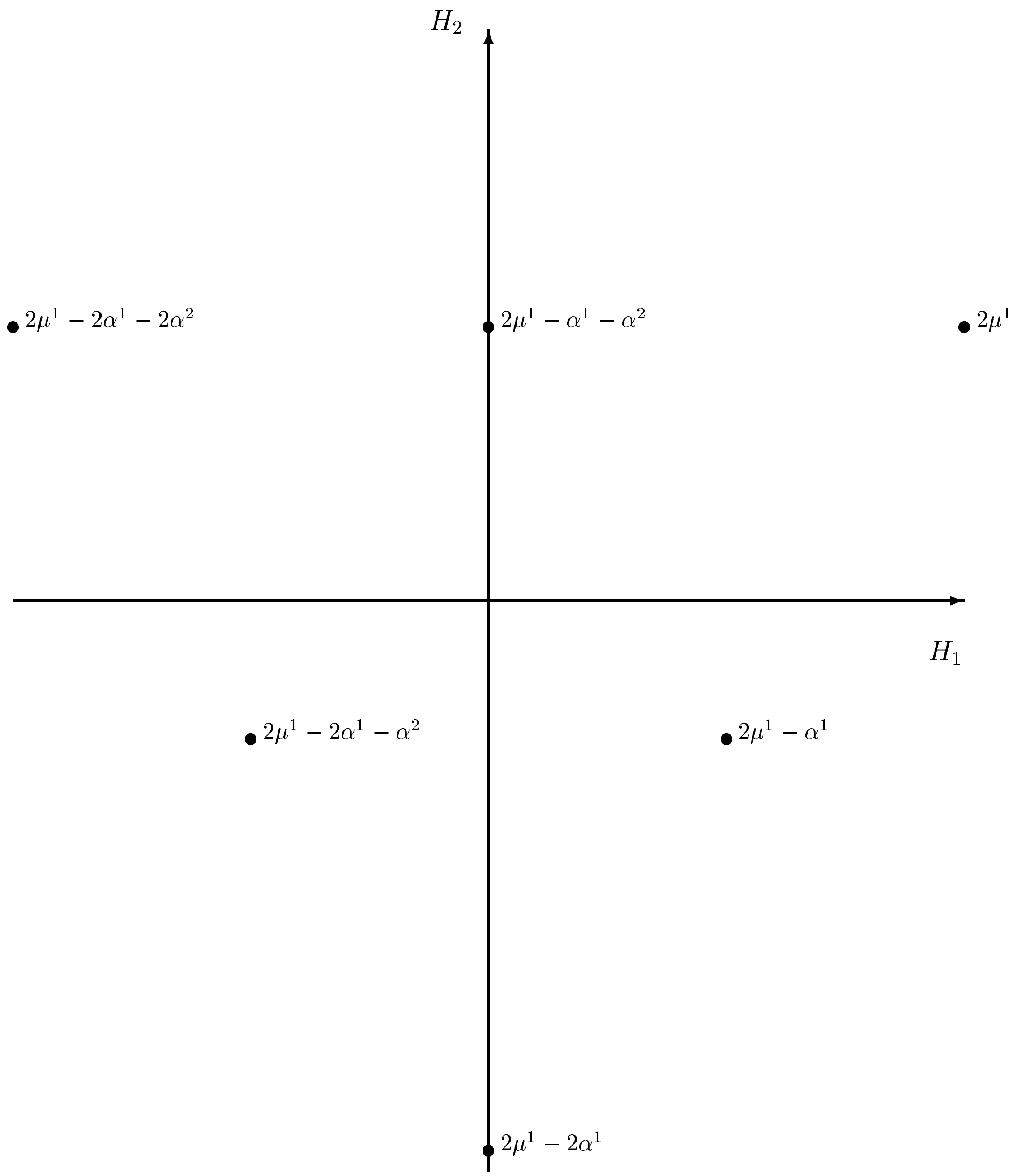


Figure 6: Weights of $SU(3)$ in the $\mathbf{6} \equiv (2,\, 0)$–representation.

reach it starting from the highest weight,and because the two states $A = E_{-\alpha^1} E_{-\alpha^2} |\alpha^1 + \alpha^2\rangle$ and $B = E_{-\alpha^2} E_{-\alpha^1} |\alpha^1 + \alpha^2\rangle$ are linearly independent. This can be seen by calculating the matrix elements $\langle A|A\rangle$, $\langle A|B\rangle$, $\langle B|A\rangle$, and $\langle B|B\rangle$ of the two states and by noting that they are linearly dependent only, if $\langle A|B\rangle\langle B|A\rangle = \langle A|A\rangle\langle B|B\rangle$. Therefore, the dimension of this representation is 8, and it is also known as the **8**–representation of $SU(3)$.

3. (3, 0) :
   We continue our rallye with the (3, 0)–representation. The highest weight is

$$3\mu^1 = \left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right) . \tag{462}$$

   From this we can construct other weights, namely $3\mu^1 - m\alpha^1$ with $m = 1, \ldots, 3$ and their Weyl reflections. The former populate the right side of the triangle in the figure. However, all weights correspond to unique states, see Fig.7.

   Obviously, there is another weight in the center, located at (0, 0). But does it correspond to an unique state ? Basically, there are two distinct ways to reach it from the highest weight, namely

$$E_{-\alpha^1} E_{-\alpha^2} E_{-\alpha^1} |3\mu^1\rangle \text{ and } \quad E_{-\alpha^2} E_{-\alpha^1} E_{-\alpha^1} |3\mu^1\rangle . \tag{463}$$

   The question now is, whether they are linearly independent. To see this, rewrite the second state as

$$\begin{aligned} & E_{-\alpha^2} E_{-\alpha^1} E_{-\alpha^1} |3\mu^1\rangle \\ &= [E_{-\alpha^2}, E_{-\alpha^1}] E_{-\alpha^1} |3\mu^1\rangle + E_{-\alpha^1} E_{-\alpha^2} E_{-\alpha^1} |3\mu^1\rangle \\ &= E_{-\alpha^1} [E_{-\alpha^2}, E_{-\alpha^1}] |3\mu^1\rangle + E_{-\alpha^1} E_{-\alpha^2} E_{-\alpha^1} |3\mu^1\rangle \\ &= E_{-\alpha^1} E_{-\alpha^2} E_{-\alpha^1} |3\mu^1\rangle . \end{aligned} \tag{464}$$

   We have employed the fact, that the commutator $[E_{-\alpha^2}, E_{-\alpha^1}]$ is proportional to $E_{-\alpha^1-\alpha^2}$ and thus commutes with $E_{-\alpha^1}$, because $-2\alpha^1 - \alpha^2$ is not a root. Furthermore, we noted that $E_{-\alpha^2}$ annihilates $|3\mu^1\rangle$. Taken together, we see, that the two states are not linearly independent, therefore the state corresponding to (0, 0) is unique and we have a representation of dimension ten called the **10**–representation of $SU(3)$.

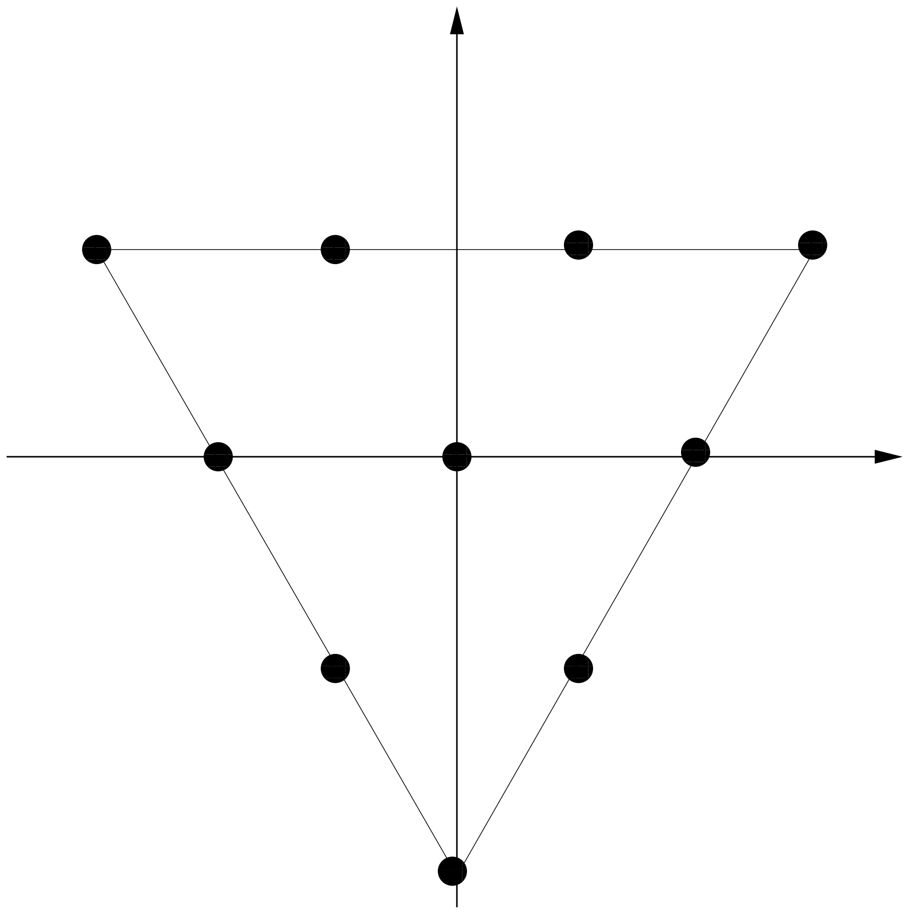


Figure 7: Weights of $SU(3)$ in the $\mathbf{10} \equiv (3,\, 0)$–representation.

7.1.6 Tensor methods

Before we return to physics, we introduce a useful tool for doing calculations explicitly. In some sense it is a generalization of the techniques to do direct products of representations of groups. We will introduce these techniques in close contact with $SU(3)$.
Let us start by relabeling the states of the **3**–representation of $SU(3)$.

$$\begin{aligned} \left|\frac{1}{2}, \frac{1}{2\sqrt{3}}\right\rangle &= |_1\rangle \\ \left|-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right\rangle &= |_2\rangle \\ \left|0, -\frac{1}{\sqrt{3}}\right\rangle &= |_3\rangle \ , \end{aligned} \tag{465}$$

where we chose the convention of naming because of the explicit form of the eigenvectors in Eq.443. Note, that the indices are somewhat lower for reasons to become clear in a minute.
Let us define a set of representation matrices with one upper and one lower index as

$$[T_a]^i_j = \frac{1}{2}\,[\lambda_a]_{ij} \tag{466}$$

with the Gell–Mann matrices λ (or whatever generators belong to the group under consideration). Then the triplet $|_i\rangle$ transforms under the algebra as

$$T_a|_i\rangle = [T_a]^j_i\,|_j\rangle\,, \tag{467}$$

where the Einstein convention is used. Accordingly, we have for the states of the $\bar{\mathbf{3}}$–representation

$$\begin{aligned} \left|-\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right\rangle &= |^1\rangle \\ \left|\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right\rangle &= |^2\rangle \\ \left|0, \frac{1}{\sqrt{3}}\right\rangle &= |^3\rangle \end{aligned} \tag{468}$$

and

$$T_a|^i\rangle = -\,[T_a]^i_j\,|^j\rangle\,. \tag{469}$$

This follows from

$$-\left[T_a^*\right]_i^j = -\left[T_a^T\right]_i^j = -\left[T_a\right]_j^i \ . \tag{470}$$

Now we are in the position to define a state in the product space of n $\mathbf{3}$'s and m $\bar{\mathbf{3}}$'s as

$$\left|{}^{i_1\ldots i_n}_{j_1\ldots j_m}\right\rangle = \left|{}_{j_1}\right\rangle \ldots \left|{}_{j_m}\right\rangle \left|{}^{i_1}\right\rangle \ldots \left|{}^{i_n}\right\rangle \ . \tag{471}$$

These beasts transform as

$$\begin{aligned} T_a \left|{}^{i_1\ldots i_n}_{j_1\ldots j_m}\right\rangle &= \sum_{l=1}^{m} \left[T_a\right]_{j_l}^k \left|{}^{i_1\ldots i_n}_{j_1\ldots j_{l-1} k j_{l+1}\ldots j_m}\right\rangle \\ &\quad - \sum_{l=1}^{n} \left[T_a\right]_k^{i_l} \left|{}^{i_1\ldots i_{l-1} k i_{l+1}\ldots i_n}_{j_1\ldots j_m}\right\rangle \ . \end{aligned} \tag{472}$$

Now, let a tensor $|v\rangle$ be any state in this tensor product space,

$$|v\rangle = v^{j_1\ldots j_n}_{i_1\ldots i_m} \left|{}^{i_1\ldots i_m}_{j_1\ldots j_n}\right\rangle \ . \tag{473}$$

Then following the path of Eq.472, the action of the generators on $|v\rangle$ is given by

$$\left[T_a v\right]^{j_1\ldots j_n}_{i_1\ldots i_m} = \sum_{l=1}^{n} \left[T_a\right]_k^{j_l} v^{j_1\ldots k\ldots j_n}_{i_1\ldots i_m} - \sum_{l=1}^{m} \left[T_a\right]_{i_l}^k v^{j_1\ldots j_n}_{i_1\ldots k\ldots i_m} \ . \tag{474}$$

Now, let us pick out the states in this tensor product, which correspond to the irreducible representation $(n\,, m)$. The state with highest weight is

$$|\mu\rangle = \left|{}^{2\,2\,\ldots}_{1\,1\,\ldots}\right\rangle \ , \tag{475}$$

and starting from this state we can obtain all the others by applying lowering operators. The state $|\mu\rangle$ corresponds to the tensor

$$|\mu\rangle \Longleftrightarrow v^{j_1\ldots j_n}_{i_1\ldots i_m} = \delta^{j_1 1} \ldots \delta^{j_n 1} \delta_{i_1 2} \ldots \delta_{i_m 2} \ . \tag{476}$$

Of course, we can also construct all other states in the representation by acting on this tensor with lowering operators, and this is even nicer. Two important points come to our help here

1. v is symmetric in upper and lower indices and it satisfies

$$\delta^{i_1}_{j_1} v^{j_1\ldots j_n}_{i_1\ldots i_m} = 0 \ . \tag{477}$$

2. These properties are preserved under $v \to T_a v$.

The δ^i_j is called an invariant tensor, because

$$[T_a]^j_i \delta^i_j = [T_a \delta] = 0 \tag{478}$$

and we have

$$[T_a]^j_i \delta^k_j - [T_a]^k_j \delta^j_i = 0\,. \tag{479}$$

Since the lowering operators are just linear combinations of the T_a, we can conclude, that all states in the (n, m)–representation therefore correspond to traceless tensors of the form of v, symmetric in upper and lower indices. This correspondence holds in b both directions, i.e. every traceless tensor with the correct number of entries n and m and symmetry in the indices forms a state in the (n, m)–representation.
These tensors are extremely nice and suitable for practical work, it is pretty easy to multiply them to construct tensors with larger numbers of indices, and they are easily decomposed into δ's and ϵ_{ijk}, the other invariant tensors (they are invariant due to the tracelessness of both the ϵ–tensor and the T_a).
If we take a bra $\langle v|$ instead of the ket $|v\rangle$, we can write this object as

$$\langle v| = \langle^{i_1 \dots i_n}_{j_1 \dots j_m}| \left(v^{j_1 \dots j_m}_{i_1 \dots i_n}\right)^* \,. \tag{480}$$

But here OEM care has to be taken, because the bras transform under the algebra with an extra minus sign, so for instance the triplet transforms like

$$-\langle_i| T_a = -\langle_i| T_a |_j\rangle \langle_j| = -\langle_j| [T_a]^i_j \,. \tag{481}$$

We conclude, that bras with upper index transform like kets with lower index, this just reflects the process of complex conjugation needed to go from kets to bras. Thus, we might identify

$$\left(v^{j_1 \dots j_m}_{i_1 \dots i_n}\right)^* \equiv \bar{v}^{i_1 \dots i_n}_{j_1 \dots j_m} \,, \tag{482}$$

with the property, that, if $\langle v|$ transforms as $-\langle v|T_a$, $|\bar{v}\rangle$ transforms as $T_a|\bar{v}\rangle$. To exemplify this behavior, consider the matrix element $\langle u|v\rangle$. Obviously, both $|u\rangle$ and $|v\rangle$ have to live in the same space and must be tensors of the same kind. Then

$$\langle u|v\rangle = \bar{u}^{k_1 \dots k_m}_{l_1 \dots l_n} v^{j_1 \dots j_n}_{i_1 \dots i_m} = \bar{u}^{i_1 \dots i_m}_{j_1 \dots j_n} v^{j_1 \dots j_n}_{i_1 \dots i_m} \tag{483}$$

to guarantee, that there is no "pending index" left. So, all indices get contracted.

7.1.7 Dimension of (n, m)

The dimensionality of a representation (n, m) is simple to determine using the tensor concepts developed so far. First, let us calculate the number of independent tensor components for a symmetric tensor with n lower (or upper) indices. Only the number of indices with each value $\{1, 2 \text{ or } 3\}$ matters. Therefore the number of independent components equals the number of ways one can separate n identical objects with two identical partitions. Combinatorics tells us, that this number is $(n+2)!/(n!2!)$. Thus, a tensor with n upper and m lower indices has

$$\frac{(n+2)!}{2!n!}\frac{(m+2)!}{2!m!} = B(n, m) \tag{484}$$

independent components. But the trace condition says, that out of this a symmetric object with $n-1$ upper and $m-1$ lower indices vanishes, leaving us with dimension

$$D(n, m) = B(n, m) - B(n-1, m-1) = \frac{(n+1)(m+1)(n+m+2)}{2} \tag{485}$$

for the traceless tensor associated with the representation of $(n\,, m)$.

7.1.8 Clebsch–Gordan decomposition with tensors

Let us now turn this into something with practical implications. We will analyze now the composite representation $\mathbf{3} \otimes \mathbf{3}$ using the methods outlined above. The tensor product is just $v^i u^j$, which can be rewritten as

$$v^i u^j = \frac{1}{2}\left(v^i u^j + v^j u^i\right) + \frac{1}{2}\left(\epsilon^{ijk}\epsilon_{klm} v^l u^m\right) \,. \tag{486}$$

With help of this identity, we have explicitly decomposed the product into a symmetric tensor

$$\frac{1}{2}\left(v^i u^j + v^j u^i\right) \,, \tag{487}$$

which is a $\mathbf{6}$, and a $\bar{\mathbf{3}}$, namely the lower index object

$$\epsilon_{klm} v^l u^m \,. \tag{488}$$

Therefore,

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \mathbf{3}\,. \tag{489}$$

Now let us consider $\mathbf{3} \otimes \bar{\mathbf{3}}$. Here, the corresponding product can be cast into

$$v^i u_j = \left(v^i u_j - \frac{1}{3} \delta^i_j v^k u_k \right) + \left(\frac{1}{3} \delta^i_j v^k u_k \right) . \tag{490}$$

The first object denotes traceless tensors with one lower and one upper index, namely the T_a. Therefore it belongs to the $\mathbf{8}$–representation. The trivial tensor $v^k u_k$ in contrast has no indices left, therefore it is just $(0, 0) = \mathbf{1}$. Hence,

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1} . \tag{491}$$

We learn, that in general every time, a tensor is not completely symmetric, we can trade two upper indices for one lower index plus an ϵ–tensor and vice versa.

7.2 $SU(3)$ of flavor

7.2.1 Hypercharge and strangeness

Let us extend our knowledge about hadrons, i.e. the particles experiencing the strong interactions. So far we talked about protons, neutrons and the pions only, and we used the isospin group $SU(2)$ to analyze their interactions. Beyond these guys, the next lightest particles are the so called *strange* particles, which carry a quantum number dubbed *strangeness*. Indeed, when these particles were discovered, they seemed strange, and during the decades ever since many interesting aspects of physics, in particular symmetries and their violation in particle physics, were discovered in the sector of the strange hadrons.

The lightest particles carrying strangeness are the kaons or K–mesons. They are organized as an isospin doublet K^+ and K^0 (with upper index indicating electric charge in units of proton charge) with strangeness $+1$ and an isospin doublet of their antiparticles K^- and $\bar{K^0}$ with strangeness $S = -1$. Since the strong interactions conserve strangeness, the kaons and other strange particles are produced in pairs like in the reactions

$$p + p \rightarrow p + p + K^+ + K^- \quad \text{or} \quad p + p \rightarrow p + p + K^0 + \bar{K^0} . \tag{492}$$

The kaons produced are relatively short lived, because they can decay via weak interactions, which does not conserve strangeness. There lifetimes range between 10^{-10} and 10^{-8} seconds, but this gives experimentalists enough time to study them in great detail (again, special relativity comes to help). Like the pions, the kaons have spin zero and baryon number B equal to zero.

But of course there are also strange baryons : an isotriplet of strangeness -1 $(\Sigma^+, \Sigma^0, \Sigma^-)$ and an isosinglet (Λ^0), all of them with strangeness $S = -1$. Additionally, there is an isodoublet (Ξ^0, Ξ^-) with $S = -2$. All the particles

meson	mass [MeV]	Mean life [s]	flavors
$\pi^{\pm}$	139.566	$2.6 \cdot 10^{-8}$	$u\bar{d}$, $d\bar{u}$
π^0	134.976	$8.4 \cdot 10^{-17}$	$\frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right)$
η	547.3	1.18 keV	$\frac{2}{\sqrt{6}}\left(u\bar{u} + d\bar{d} - 2s\bar{s}\right)$
$K^{\pm}$	439.67	$1.2 \cdot 10^{-8}$	$u\bar{s}$, $s\bar{u}$
K^0, $\bar{K^0}$	497.672	*	$d\bar{s}$, $s\bar{d}$
η'	957.78	203 keV	$\frac{1}{\sqrt{3}}\left(u\bar{u} + d\bar{d} + s\bar{s}\right)$
$\rho^{\pm}$	770.0	150 MeV	$u\bar{d}$, $d\bar{u}$
ρ^0	770.0	150 MeV	$\frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right)$
ω	781.94	8.4 MeV	$\frac{1}{\sqrt{2}}\left(u\bar{u} + d\bar{d}\right)$
ϕ	1019.41	4.4 MeV	$s\bar{s}$
$K^{*\pm}$	891.66	50.8 MeV	$u\bar{s}$, $s\bar{u}$
K^{*0}	896.1	50.5 MeV	$d\bar{s}$

Table 6: Some of the properties of the lowest lying $SU(3)$–multiplets of pseudoscalar and vector mesons. Masses are in MeV, lifetimes are given either in seconds or as total decay width. Note, that the neutral kaons, K^0 and $\bar{K^0}$ are not the physically observable states, instead, linear combinations are the relevant particles. These linear combinations are dubbed K_S^0 and K_L^0, where the subscripts denote "short" or "long" according to their drastically different lifetimes. For K_S^0 the mean life is $\tau \approx 0.8 \cdot 10^{-10}$ and for K_S^0 $\tau \approx 5.2 \cdot 10^{-8}$. The flavors denote the composition of the mesons in terms of quark degrees of freedom, which will be discussed later. Note, that in addition to the $SU(3)$–octet mesons we have also included an $SU(3)$–singlet η' into the table.

named so far satisfy a relation like Eq.179, relating the third component of isospin, T_3, with the electrical charge Q and hypercharge $Y = B + S$,

$$Q = T_3 + \frac{Y}{2}\,. \tag{493}$$

Some of the properties of the particles discussed so far are displayed in Tables 6 and 7. When considering the baryons, it becomes pretty obvious, that the mass splitting between the different states is not very large, therefore it is quite natural to talk about multiplets. With the mesons, the situation is slightly different. Here the mass splittings become larger than the meson masses themselves.
However something funny happens, if we plot the particles discussed so far in a table with axes T_3 and Y : We find hexagons, see Figs.8 and 9 ! These pictures and the multiplet of the vector mesons displayed in Fig.10 suggest, that we are dealing with something like the **8** of $SU(3)$, with $\mathcal{H}_1 = T_3$ and $\mathcal{H}_2 = \sqrt{3}/2 \cdot Y$.

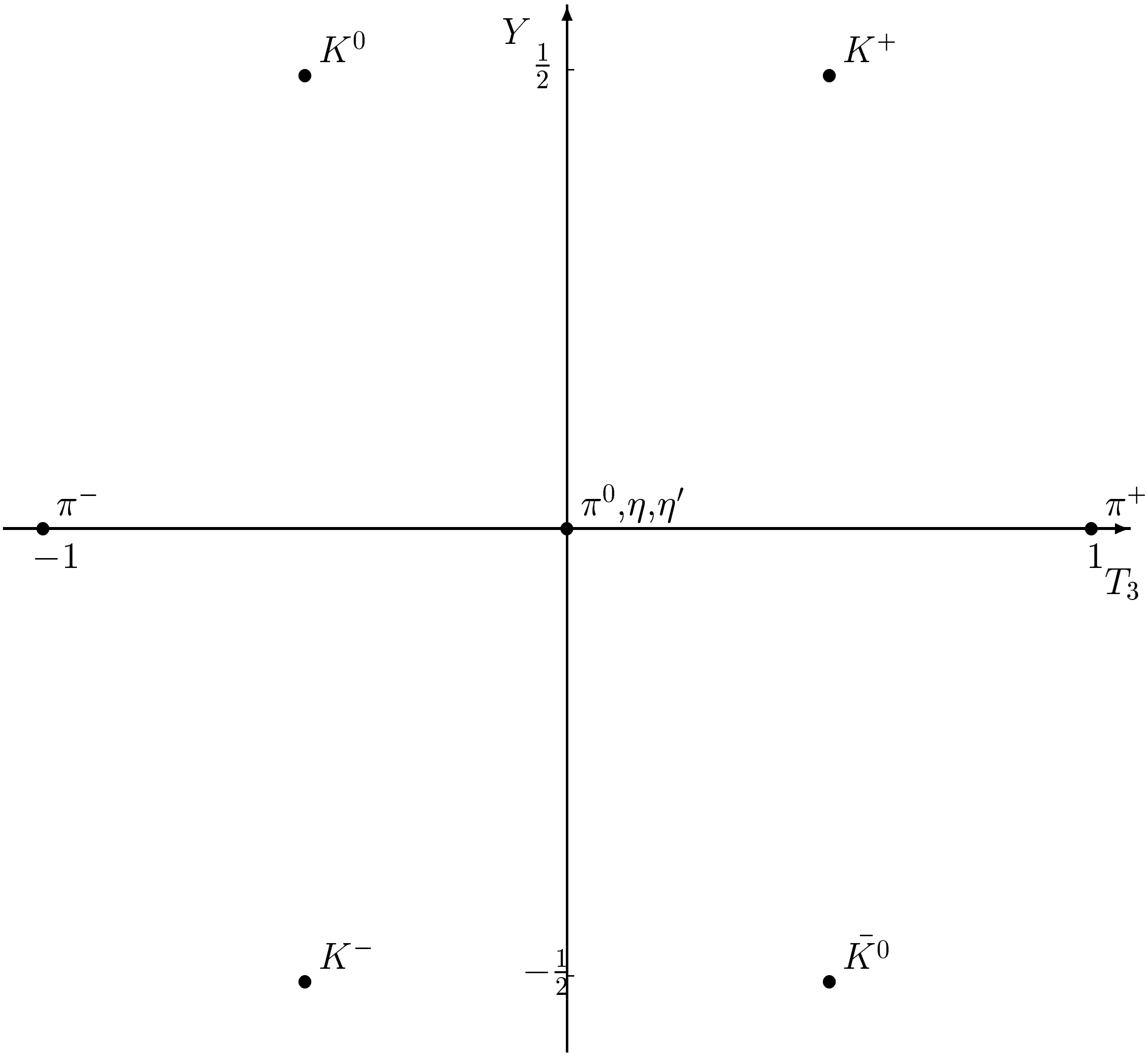


Figure 8: The lowest lying (pseudoscalar) mesons.

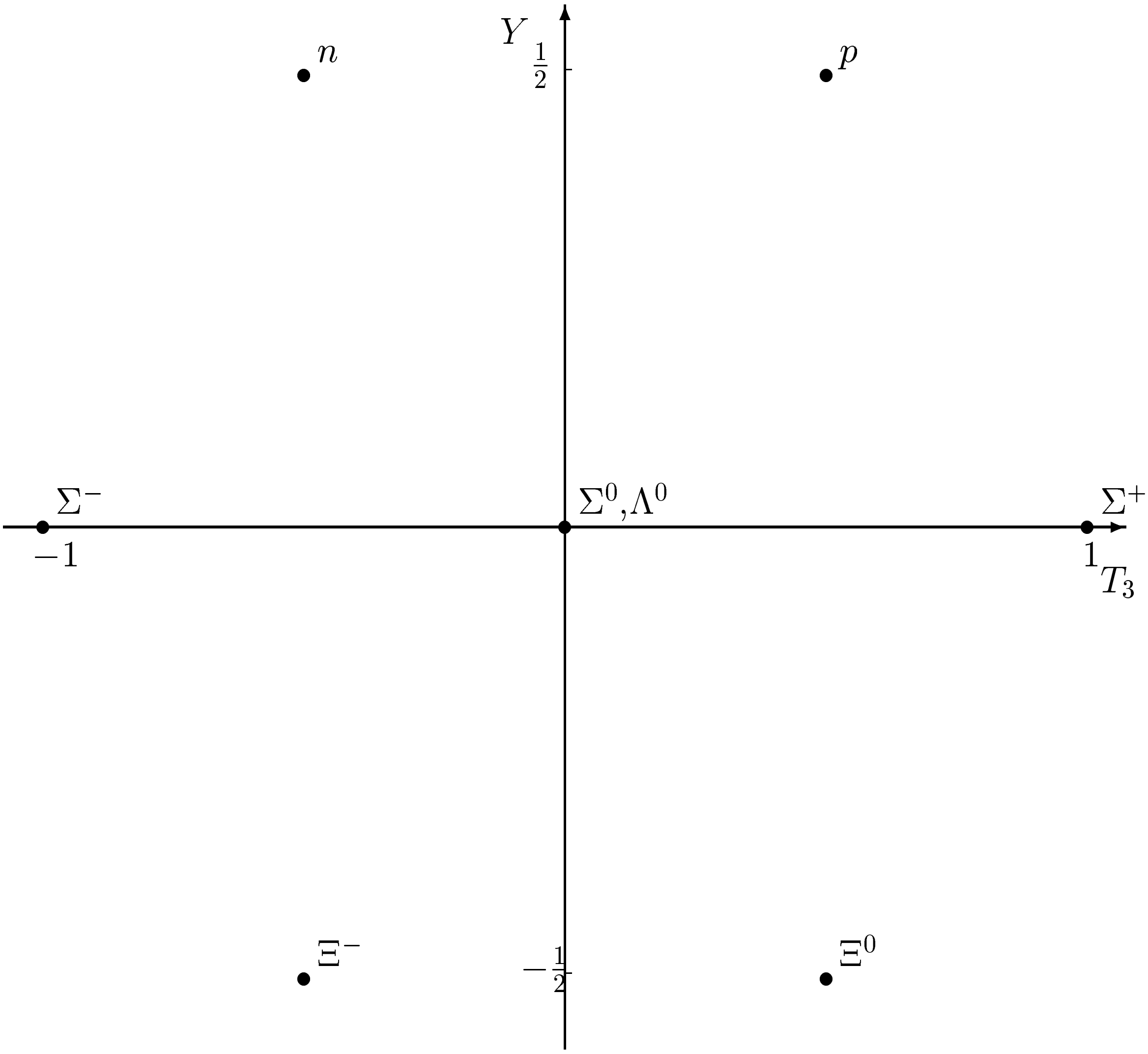


Figure 9: The lowest lying baryons.

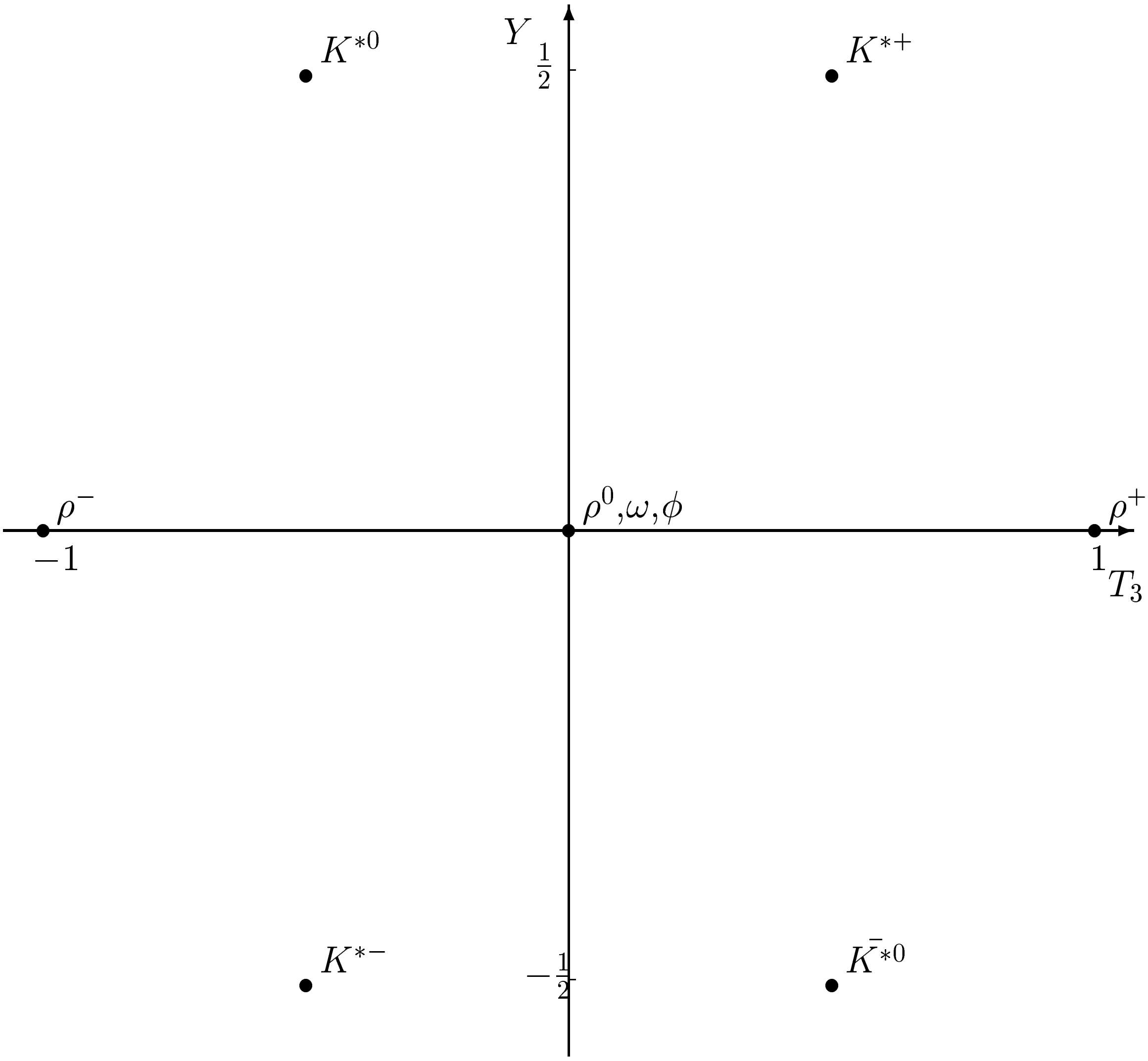


Figure 10: The vector mesons.

baryon	mass [MeV]	Mean life [s]	flavors
p	938.272	$> 1.6 \cdot 10^{25}$	uud
n	939.565	886.7 ± 1.9	udd
Λ^0	1115.683	$2.6 \cdot 10^{-10}$	uds
Σ^0	1192.64	$7.4 \cdot 10^{-20}$	uds
Σ^+	1189.37	$0.8 \cdot 10^{-10}$	uus
Σ^-	1197.45	$1.4 \cdot 10^{-10}$	dds
Ξ^0	1314.9	$2.9 \cdot 10^{-10}$	uss
Ξ^-	1321.32	$1.6 \cdot 10^{-10}$	dss

Table 7: Some of the properties of the lowest lying $SU(3)$–multiplet of baryons. Masses are in MeV, lifetimes are given in seconds. The flavors denote the composition of the baryons in terms of quark degrees of freedom.

7.2.2 The Gell–Mann Okubo relation

The multiplet structure outlined above suggests, that we can think of $SU(3)$ as underlying symmetry determining hadron states and their masses. Neglecting electromagnetic and weak interactions, we are left with the question, how the multiplets split. A simple model is to divide the strong interactions in two parts : One part, which is extremely strong and obeys $SU(3)$, and one medium strong part, which conserves only isospin and hypercharge, but not the other $SU(3)$ generators. The conservation of hypercharge and isospin is motivated by the multiplet structure itself, forming $SU(3)$ multiplets and the clustering within the multiplets in terms of isoplets like p and n. Then, if something like perturbation theory is any good in this medium strong sector, we will find exactly the multiplets above.
Let us use this idea now, to extract more information than a mere classification of states, which alone is a mnemonic device, only. Let us consider the baryon masses first and let us write the baryon states as

$$B^i_j \left| {}^j_i \right\rangle , \tag{494}$$

where B^i_j is a tensor, that labels the baryon displayed in Fig.9. It reads

$$B^i_j = \begin{pmatrix} \frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & -\frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda^0}{\sqrt{6}} & n \\ \Xi^- & \Xi^0 & -\frac{2\Lambda^0}{\sqrt{6}} \end{pmatrix} \tag{495}$$

and accordingly we have

$$B^i_j \left|{}^j_i\right\rangle = p|p\rangle + n|n\rangle + \dots \,. \tag{496}$$

Now, we are in the position to discuss the masses of these states. Ignoring again electromagnetic and weak interactions, the mass of a single particle state $|\Psi\rangle$ can be written as

$$m_\Psi = \langle\Psi|\mathcal{H}|\Psi\rangle = \langle\Psi|\mathcal{H}_{VS}|\Psi\rangle + \langle\Psi|\mathcal{H}_{MS}|\Psi\rangle \,. \tag{497}$$

Here subscripts VS and MS denote "very strong" and "medium strong", respectively. Since the VS part commutes with all the eight $SU(3)$ generators in our toy model, it contributes the same mass to each state. In principle we can not say so much about the contributions of the MS part, which commutes with the isospin generators $T_{1,2,3}$ and with the hypercharge T_8 but not with the other generators. But let us further assume, that $\mathcal{H}_{MS}$ is some component of some tensor operator. Then we can use the Wigner–Eckart theorem to analyze the mass relation above. In particular, we will chose the ansatz

$$\mathcal{H}_{MS} = [T_8]^i_j \; \mathcal{O}^j_i \tag{498}$$

for the medium strong interactions. Later in this chapter, in Sec.7.3 we will give some motivation for this choice. Here, $\mathcal{O}^j_i$ is some octet tensor operator satisfying

$$\left[T_a \, , \, \mathcal{O}^j_i\right] = \mathcal{O}^j_k \, [T_a]^k_i - \mathcal{O}^k_i \, [T_a]^j_k \ . \tag{499}$$

What can we say about the matrix elements

$$\langle B|\mathcal{H}_{MS}|B\rangle \, ? \tag{500}$$

We can – and should – use tensor methods. Obviously, the matrix element of Eq.500 is proportional to the tensors B^i_j, $[T_8]^i_j$ and $\bar{B}^i_j$. But it should be invariant as well, therefore, we have to contract all tensor indices. There are only two ways of doing this, namely by

$$\begin{aligned} \bar{B}^i_j B^j_k \, [T_8]^k_i &= \mathrm{Tr}\left(B^\dagger B T_8\right) \\ \bar{B}^i_j \, [T_8]^k_j \, B^i_k &= \mathrm{Tr}\left(B^\dagger T_8 B\right) \ . \end{aligned} \tag{501}$$

Therefore, Eq.500 can be written in terms of two numbers x and y and the traces as

$$
\begin{aligned}
\langle B|\mathcal{H}_{MS}|B\rangle &= x\mathrm{Tr}\left(B^{\dagger}BT_8\right) + y\mathrm{Tr}\left(B^{\dagger}T_8B\right) \\
&= \quad \frac{x}{12}\left(|\Sigma|^2 + |\Xi|^2 - |\Lambda|^2 - 2|n|^2\right) \\
&\quad + \frac{y}{12}\left(|\Sigma|^2 + |n|^2 - |\Lambda|^2 - 2|\Xi|^2\right) \;, \qquad (502)
\end{aligned}
$$

where a sum over the different particle types within each isomultiplet is understood, i.e.

$$
\begin{aligned}
|\Sigma|^2 &= |\Sigma^+|^2 + |\Sigma^0|^2 + |\Sigma^-|^2 \\
|\Xi|^2 &= |\Xi^0|^2 + |\Xi^-|^2 \\
|n|^2 &= |p|^2 + |n|^2 \,, \qquad (503)
\end{aligned}
$$

and we assume approximately the same masses within each of these isomultiplets. Adding now a common mass m_0 stemming from the very strong interaction part, $\mathcal{H}_{VS}$, we end up with

$$
\begin{aligned}
m_n &= m_0 - \frac{x}{6} + \frac{y}{12} \\
m_\Lambda &= m_0 - \frac{x}{12} - \frac{y}{12} \\
m_\Sigma &= m_0 + \frac{x}{12} + \frac{y}{12} \\
m_\Xi &= m_0 + \frac{x}{12} - \frac{y}{6} \qquad (504)
\end{aligned}
$$

and from here, we can eliminate the unknown x, y and m_0 in favor of one relation among the four masses

$$
2(m_n + m_\Xi) = 3m_\Lambda + m_\Sigma \,. \qquad (505)
$$

This is the Gell–Mann Okubo formula. Putting in rounded values for the masses we find agreement on a level of less than 20 MeV. It works ! Some simple symmetry arguments and a toy model for the mass splitting within the multiplets gave us a surprisingly good answer. This is even more surprising, since the mass splitting within the isospin multiplets is of the order of 5 MeV itself (this difference is at least a hint on the size of the neglected terms). The only assumptions we used, were the obvious hints from the mass structure within the multiplets leading to our guess about the form of the Hamiltonian plus a somewhat extended version of the Wigner–Eckart theorem.

7.2.3 Resonances - the decuplet

We will continue our journey through $SU(3)$ of flavor by considering higher multiplets. Historically, these particles were found by observing an

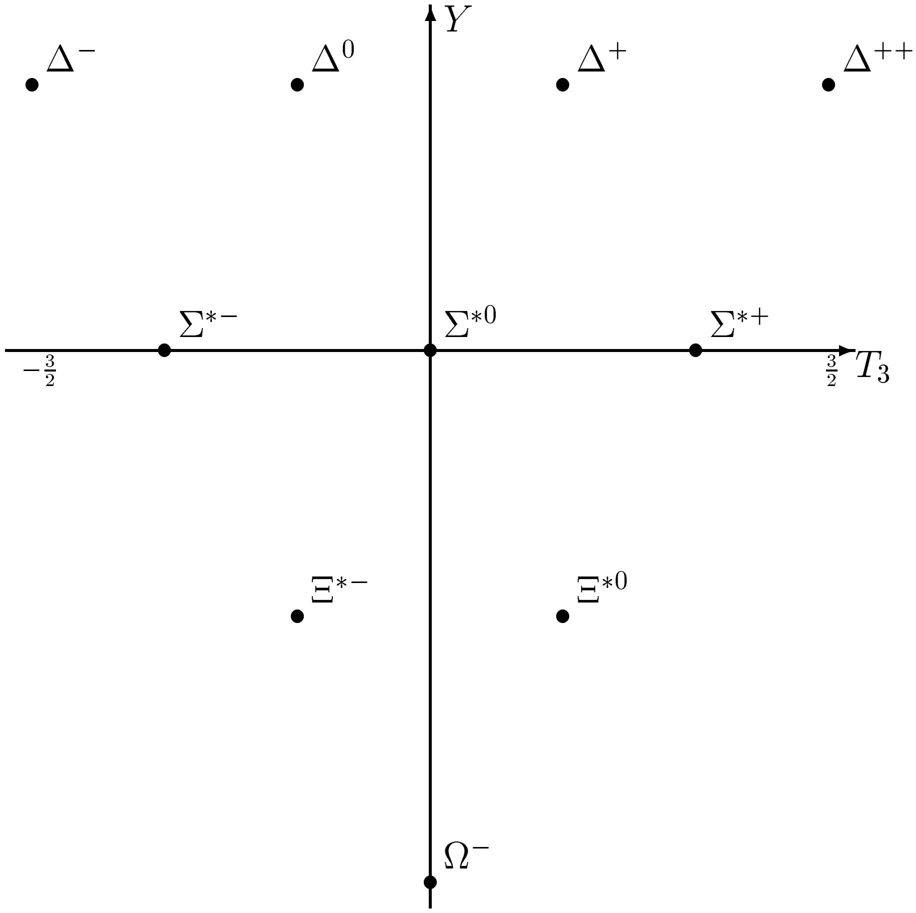


Figure 11: The resonances of the decuplet.

enhancement of the πp cross section at center of mass energies around 1230 MeV, which was interpreted as the production of a resonance. This resonance occurred in all charge states, ++, +, 0 and −, suggesting an isospin–3/2 multiplet. Comparing with our pictorial representations of the various $SU(3)$–multiplets, we find, that this suggests the **10**, the decuplet, see Fig.11. In fact, all particles displayed have been observed quite a time ago. However, when $SU(3)$ was suggested first, the Ω^- has not been seen. Gell–Mann was then able to predict not only its existence, but also its mass. Let us sketch the mass prediction once more.
We stay with our assumption, that $\mathcal{H}_{MS}$ transforms like the **8**–component of the octet. Thus, we need to compute $\mathbf{10} \otimes \mathbf{8}$. We find

$$(1,\,1) \otimes (3,\,0) = (4,\,1) \oplus (2,\,2) \oplus (1,\,1) \oplus (3,\,0)\,. \tag{506}$$

The important thing is, that we find an unique $\mathbf{10} = (3,\,0)$ in the product, therefore we can compute the matrix element $\langle \mathbf{10} | 8 | \mathbf{10} \rangle$ in terms of one single reduced matrix element : all matrix elements of the octet operators are then

baryon	mass [MeV]	Width [MeV]	flavors
Δ^{++}, Δ^{+}, Δ^{0}, Δ^{-}	1232	150	uuu, uud, udd, ddd
Σ^{+}, Σ^{0}, Σ^{-}	1385	37	uus, uds, dds
Ξ^{0}, Ξ^{-}	1533	9	uss, dss
Ω^{-}	1672	$\tau = 0.8 \cdot 10^{-10}$	sss

Table 8: Some of the properties of the decuplet baryons. Only average masses and widths are displayed to give the reader a flavor of the characteristic numbers. In contrast to the other resonances, the Ω^{-} is rather long–lived due to the fact, that it decays only weakly.

proportional. But we already know the matrix elements of one set of octet operators – namely the group generators themselves. In particular, we know, that T_8 is just proportional to hypercharge. So the mass splitting between the isospin multiplets should be just proportional to the hypercharge difference between them. Consequently, we predict

$$m_{\Sigma^*} - m_{\Delta} = m_{\Xi^*} - m_{\Sigma^*} = m_{\Omega^-} - m_{\Xi^*} \,. \tag{507}$$

We now compare this with the masses of the resonances listed in Table 8. Again, it works ! The success of this prediction convinced almost all particle physicists, that $SU(3)$ is an useful approximate symmetry in the realm of light baryons and mesons.

7.2.4 Quarks

We can think of all these strong interacting particles as being built out of quarks. They were first invented by Gell–Mann, and they served more as mnemonic device. Clearly, if we arrange (anti–) quarks in an (anti–) **3** of $SU(3)$ (see Fig.12), we can think of mesons as being built from one quark anti-quark pair. To see this, just remember, that

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1} \,. \tag{508}$$

Similarly, the baryons are built from three quarks, since

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1} \,. \tag{509}$$

Note, that we need two spin–1/2 particles to have scalars and vectors for the mesons and three such particles to have a fermion again. So we can invent three quarks – sometimes dubbed flavors – which were called u, d and s for "up", "down" and "strange". The electrical charges of these guys are given by

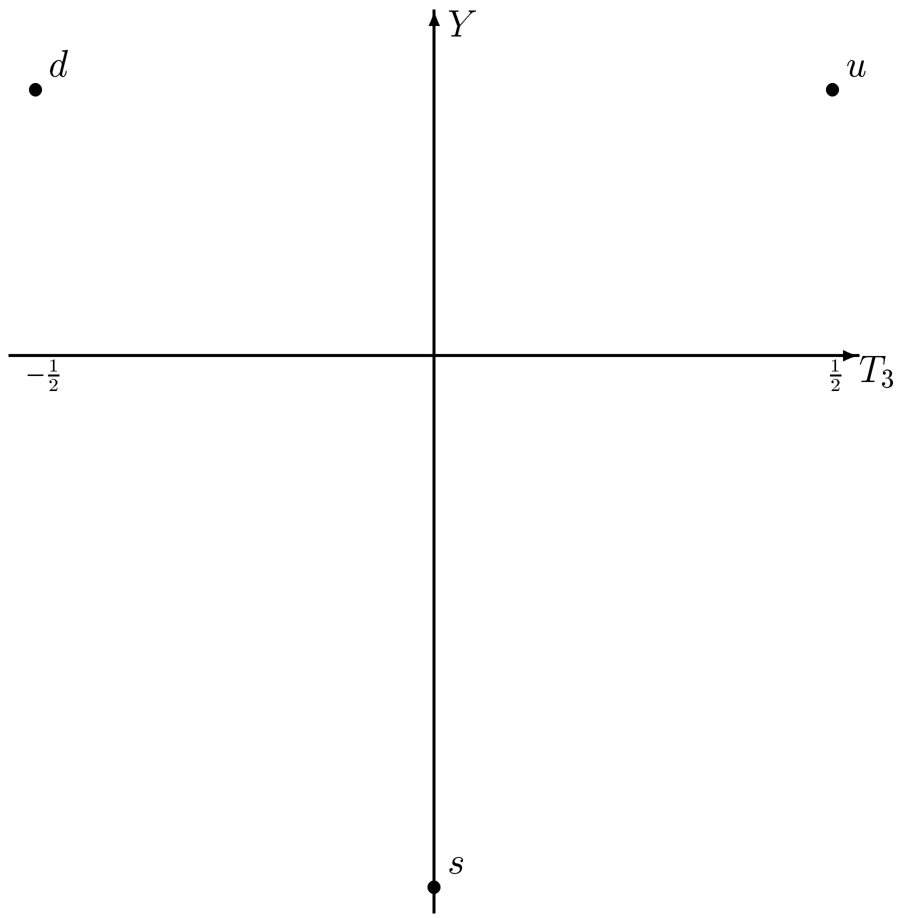


Figure 12: The quarks

$$Q = T_3 + \frac{Y}{2} = H_1 + \frac{H_2}{\sqrt{3}}\,, \tag{510}$$

and hence we have charges $+2/3$ for the up and $-2/3$ for down and strange quarks.

7.3 $SU(3)$ of color

7.3.1 Introducing color as a quantum number

We have to think of quarks as spin–1/2 fermions, since they form spin–0, 1, ... mesons when coupled as $q\bar{q}$–pairs as well as spin 1/2, 3/2, ... baryons when coupled as qqq–combinations. Consider now the Δ^{++} resonance with spin 3/2. Clearly in its ground state, the three up–quarks are coupled with parallel spin and therefore the supposedly symmetric wave function violates the Fermi–Dirac statistics. What can we do to save the apparently fundamental connection of half–integer spin and Fermi–Dirac statistics for the quarks ?

And what kind of peculiar dynamics delivers only $q\bar{q}$ and qqq states ? The resolutions of these two puzzles are in fact based in one common framework, which marks one of the most exiting realizations of $SU(3)$ to particle physics. Let us start with the first problem. The first naive guess might consist in introducing a new, so far hidden label – a new quantum number – in order to anti-symmetrize the wave functions in this label and to maintain the correct quantum numbers of the other symmetries. The most economic idea, which in fact has the most physical and group theoretical significance, is to introduce this new label as an $SU(3)$ index. The complete antisymmetric state in such indices thus is an $SU(3)$–singlet. This new quantum number carried by the quarks is called *color*, and the indices are often dubbed red, blue and green for reasons to become apparent. Let us stress here, that this new $SU(3)$ has nothing in common with the Gell–Mann $SU(3)$, which is an approximate symmetry in the three light quarks u, d and s. As far as we understand it, the color $SU(3)$ is an exact symmetry of the world, related to the fact that all observable hadronic matter consists of either three quarks or a quark–anti-quark pair. So, we assume now, that the quarks and anti-quarks transform as a $\mathbf{3}$ and a $\bar{\mathbf{3}}$ under this $SU(3)$. Now, the quark wave functions behave as tensors, i.e. like

$$q^i\,, \quad i = 1,\, 2,\, 3\,, \tag{511}$$

with i the color index standing for red, blue or green. The "cheapest" color singlet states can be formed via

$$q^i \bar{q}_j \delta_i^j \quad \text{or} \quad q^i q^j q^k \epsilon_{ijk} \tag{512}$$

and we found again our mesons or baryons constructed out of quarks.

7.3.2 Bound states

But one question remains : Why do we only find color singlets in nature ? To answer this question, we briefly go to the realm of atoms. There we find electrons with negative electrical charge bound to the positively charged nuclei. The strength of the coupling between charges is proportional to its product, and it becomes weaker with rising distance. Maybe, we can construct some generalization of this concept. Then, the electrical charges are replaced by the colors – to be more specific, by the $SU(3)$ generators.
Let us now construct such a model. We start with a state consisting of two particles A and B transforming according to some representation of $SU(3)$,

$$\begin{aligned} T_a^A |r, A\rangle &= \left[T_a^A\right]_{sr} |s, A\rangle \\ T_b^B |x, B\rangle &= \left[T_b^B\right]_{yx} |y, B\rangle\,. \end{aligned} \tag{513}$$

Here r, s and x, y label the color of the states A and B, respectively. Consequently, the two particle state is a tensor product described by

$$|v, AB\rangle = v_{rx}|r, A\rangle|x, B\rangle \,. \tag{514}$$

In analogy to the electromagnetic force we now assume, that the force between the two color charges is proportional to the product of the two $SU(3)$ generators

$$T_a^A T_a^B \,. \tag{515}$$

First, let us show, that this product is consistent with the transformation behavior under $SU(3)$. The color generators for the tensor product space are

$$T_a^{AB} = T_a^A + T_a^B \,, \tag{516}$$

and they commute with the product, since

$$\begin{aligned} \left[T_a^A + T_a^B, T_b^A T_b^B\right] &= \left[T_a^A, T_b^A\right] T_b^B + T_b^A \left[T_a^B, T_b^B\right] \\ &= i f_{abc} \left(T_c^A T_b^B + T_b^A T_c^B\right) = 0 \,. \end{aligned} \tag{517}$$

This is good news, because it tells us, that such product terms are allowed in an Hamiltonian without destroying our $SU(3)$ symmetry. Furthermore, the eigenstates of T_a^{AB} will be irreducible representations of $SU(3)$. We can rewrite

$$T_a^A T_a^B = \frac{1}{2}\left(T_a^A + T_a^B\right)^2 - \frac{1}{2}T_a^{A^2} - \frac{1}{2}T_a^{B^2} = \frac{1}{2}\left(T_a^{AB^2} - T_a^{A^2} - T_a^{B^2}\right) \,. \tag{518}$$

The object T_a^2 summed over a are called Casimir operators, since they commute with the $SU(3)$ generators. Therefore, in each irreducible representation, they are just numbers, which can be computed directly (in some sense, they are just the sums of squares of the weights of the representation). In a color singlet, these Casimir operators are of course zero, therefore, for given T_a^A and T_a^B Eq.518 is smallest, if T_a^{AB} describes a color singlet state. Then the force between the color charges A and B is the strongest – the minus sign reflects the fact, that opposite charges attract. For a $q\bar{q}$ combination, this configuration is easy to obtain, and it is equally easy to arrange the colors in a way, that they form a singlet. For a qqq state, we can not use Eq.518 directly. Instead we take a detour by considering each pair qq of quarks separately. For every pair, there are two possible arrangements, namely a $\mathbf{6}$ and a $\bar{\mathbf{3}}$ for symmetrical and anti-symmetrical configurations, respectively. For the $\mathbf{6}$, $T_a^{AB^2}$ is 10/3, while for the $\bar{\mathbf{3}}$ it is 4/3. We can conclude immediately, that the $\bar{\mathbf{3}}$ state is more attractive, and that the qqq state is most attractive, if all pairs are in relative $\bar{\mathbf{3}}$'s. This is the case, if the colors of the three quarks are in a completely antisymmetric state. Then the qqq state is a singlet as well.

7.3.3 Introducing gluons

Before we try to motivate the choice of T_8 for the Gell–Mann Okubo relation, let us briefly dwell on the theory of strong interactions on the quantum level. When constructing our bound states as color singlets out of $q\bar{q}$ and qqq configurations, we have borrowed the concept of attracting opposite charges from electrodynamics. On the quantum level, this attractive force in QED (Quantum ElectroDynamics) is identified with the coupling of photons to charged particles and to the exchange of these photons. In the quantum theory of the strong interactions, dubbed QCD (Quantum ChromoDynamics, since the charges are colored), this force is associated with the coupling of so–called gluons to the color charges. But instead of one uncharged photon in QED we have eight gluons in QCD, which carry some color charges themselves [2] However, in addition to the larger coupling constant, this introduces another difference when comparing QCD with QED. Because of the fact, that the gluons carry color charges themselves, there are not only interactions of the gluons with the quarks, but also interactions among the gluons. This self–interaction finally gives rise to a property of QCD which is dubbed *confinement* and results in the observation that free color charges do not exist in nature. The idea behind this is, that in striking difference to electromagnetic forces, due to quantum effects the color forces become stronger with rising distance until they become constant at scales of length $\Lambda^2_{\rm QCD}/\hbar c$, where $\Lambda \approx 250$ MeV.
With this picture of strong interactions in mind, we can now try to understand why $SU(3)$ of flavor is a useful symmetry. Since the interaction of gluons with the quarks is independent of the flavor of the quarks (QCD is "flavor–blind"), the only difference lies in the different masses of the quarks. A possible mass term in the Hamiltonian looks like

$$\mathcal{H}_M = m_u u^\dagger u + m_d d^\dagger d + m_s s^\dagger s\,, \tag{519}$$

where the u, d, s and their "daggered" counterparts denote creation and annihilation operators, respectively. This expression can be rewritten as

$$\begin{aligned}\mathcal{H}_M \;&=\; \frac{m_u+m_d+m_s}{3}\left(u^\dagger u + d^\dagger d + s^\dagger s\right)\\ &\quad+\frac{m_u-m_d}{2}\left(u^\dagger u - d^\dagger d\right)\\ &\quad+\frac{2m_s-m_u-m_d}{6}\left(2s^\dagger s - u^\dagger u - d^\dagger d\right)\,.\end{aligned} \tag{520}$$

We see, that the first line of this equation is $SU(3)$–flavor invariant. The second one is associated with the breaking of isospin. But this term nearly

[2]Because they are identified with the generators of $SU(3)$, the gluons are in the adjoint representation. In turn, the quarks are in the fundamental representation of $SU(3)$.

vanishes due to the tiny mass difference of u and d–quarks. Thus, isospin is a very good symmetry. The third term is responsible for breaking Gell–Manns flavor $SU(3)$ and, consequently, for the mass splitting within the multiplets. It is still fairly small, since the difference of s–mass to the u and d masses is smaller than Λ_{QCD}. Note the particular form of the tensor operator built from the quark operators. Obviously, it transforms like T_8, our friend from the Gell–Mann Okubo relation.

A Tedious proofs

A.1 Matrix elements of tensor operators

A.1.1 Scalar operators

Consider the matrix element $\langle Njm|\mathcal{S}|N'j'm'\rangle$ of a scalar operator between states characterized by the eigenvalues j of J^2, m of J_z and N for all additional quantum numbers. Since $\mathcal{S}$ commutes with the angular momentum operator, see Eq.271, J^2 and any component of J operate identically on $|Njm\rangle$ and $\mathcal{S}|Njm\rangle$. In other words, $\mathcal{S}|Njm\rangle$ transforms according to the m–th column of the j–th irreducible representation of $SO(3)$. Thus, we have

$$\langle Njm|\mathcal{S}|N'j'm'\rangle = \delta_{jj'}\delta_{mm'}\langle Njm|\mathcal{S}|N'jm\rangle\,. \tag{521}$$

Using now the ladder operators as

$$\begin{aligned} |Njm\rangle &= \frac{1}{\hbar\sqrt{j(j+1)-m(m-1)}}J_+|Njm-1\rangle \\ \langle Njm|J_+ &= \langle Njm-1|\sqrt{j(j+1)-m(m-1)}\hbar \end{aligned} \tag{522}$$

We can write

$$\begin{aligned} \langle Njm|\mathcal{S}|N'jm\rangle &= \frac{1}{\hbar\sqrt{j(j+1)-m(m-1)}}\langle Njm|\mathcal{S}J_+|N'jm-1\rangle \\ &= \langle Njm-1|\mathcal{S}|N'jm-1\rangle\,. \end{aligned} \tag{523}$$

This shows, that the matrix element is independent of the actual value of m.

A.1.2 Clebsch–Gordan series

We proof the relation Eq.295,

$$\begin{aligned} &D^{(j_1)}_{m_1m_1'}(R)D^{(j_2)}_{m_2m_2'}(R) = \\ &\qquad\sum_{j,m,m'}\langle j_1j_2;m_1m_2|j_1j_2;jm\rangle\langle j_1j_2;m_1'm_2'|j_1j_2;jm'\rangle D^{(j)}_{mm'}(R)\,. \end{aligned} \tag{524}$$

The left–hand side of this equation is

$$\begin{aligned} \langle j_1j_2;m_1m_2|D(R)|j_1j_2;m_1'm_2'\rangle &= \langle j_1m_1|D(R)|j_1m_1'\rangle\langle j_2m_2|D(R)|j_2m_2'\rangle \\ &= D^{(j_1)}_{m_1m_1'}(R)D^{(j_2)}_{m_2m_2'}(R)\,, \end{aligned} \tag{525}$$

but in fact we could compute the same matrix element by inserting a complete set of states in the (j,m)–basis, yielding

$$\begin{aligned}
&\langle j_1 j_2; m_1 m_2 | D(R) | j_1 j_2; m_1' m_2' \rangle \\
&= \sum_{j,m,j',m'} \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle \langle j_1 j_2; j'm' | j_1 j_2; m_1' m_2' \rangle \\
&\quad \times \ \langle j_1 j_2; jm | D(R) | j_1 j_2; j'm' \rangle \\
&= \sum_{j,m,j',m'} \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle \langle j_1 j_2; m_1' m_2' | j_1 j_2; j'm' \rangle D^{(j)}_{mm'} \delta_{jj'} \ , \qquad (526)
\end{aligned}$$

but this is exactly the right–hand side of Eq.295.

A.2 Commutation relations for the Runge–Lenz vector

A.2.1 Invariance of $\vec{M}'$

First, we will show, that the Runge–Lenz vector in fact is a constant of motion. We can show this on the classical level most easily by derivation with respect to time,

$$\begin{aligned}
\dot{\vec{M}}' &= \frac{1}{2\mu}\left[\dot{\vec{p}} \times \vec{L} - \vec{L} \times \dot{\vec{p}}\right] - \frac{Ze^2}{r}\left[\dot{\vec{r}} - \frac{\vec{r}(\vec{r}\cdot\vec{p})}{\mu r^2}\right] \\
&= \frac{1}{\mu}\left[\dot{\vec{p}} \times (\vec{r} \times \vec{p})\right] - \frac{Ze^2}{\mu r}\left[\vec{p} - \frac{\vec{r}(\vec{r}\cdot\vec{p})}{r^2}\right] \\
&= \frac{1}{\mu}\left[\vec{r}\left(\dot{\vec{p}}\cdot\vec{p}\right) - \vec{p}\left(\dot{\vec{p}}\cdot\vec{r}\right)\right] - \frac{Ze^2}{\mu r}\left[\vec{p} - \frac{\vec{r}(\vec{r}\cdot\vec{p})}{r^2}\right] \\
&= 0\,, \qquad (527)
\end{aligned}$$

where we have used, that $\vec{p} = \mu\vec{r}$, $\vec{L} = \vec{r} \times \vec{p}$ and $\dot{\vec{p}} = Ze^2\vec{r}/r^3$.
On the quantum mechanical level, things are not that easy. Instead we have to evaluate the commutator of the Runge–Lenz vector with the Hamiltonian of the system. For this purpose we need the following little helpers (remember, $\mathcal{H} = \vec{p}^2/(2\mu) - Ze^2/r$)

$$\begin{aligned}
[p_i, \mathcal{H}] &= -Ze^2\left[p_i, \frac{1}{r}\right] = iZe^2\hbar\left(\partial_i \frac{1}{r} - \frac{1}{r}\partial_i\right) = -iZe^2\hbar\frac{x_i}{r^3} \\
\left[\frac{x_i}{r}, \vec{p}^2\right] &= -\hbar^2\left(\frac{x_i}{r}\nabla^2 - \nabla^2\frac{x_i}{r}\right) = 2\hbar^2\left(-\frac{x_i}{r^3} + \frac{1}{r}\partial_i - \frac{x_i}{r^3}x_k\partial_k\right) , \qquad (528)
\end{aligned}$$

where we have used, that

$$\partial_k\partial_k\frac{x_i}{r} = \partial_k\left(\frac{\delta_{ik}}{r} - \frac{x_i x_k}{r^3} + \frac{x_i}{r}\partial_k\right) . \qquad (529)$$

Analogously, we have

$$\left[p_n, \frac{x_k}{r^3}\right] = -i\hbar \left(\frac{\delta_{kn}}{r^3} - \frac{3x_k x_n}{r^5}\right) . \tag{530}$$

This enables us to work out

$$\left[\vec{M}', \mathcal{H}\right] = \left[\frac{1}{2\mu}\left(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}\right) - \frac{Ze^2\vec{r}}{|\vec{r}|}, \mathcal{H}\right] \tag{531}$$

We see, that the second commutator,

$$\left[\frac{Ze^2\vec{r}}{|\vec{r}|}, \mathcal{H}\right] \quad = \quad Ze^2\left[\frac{x_i}{r}, \vec{p}^2\right] \tag{532}$$

is up to a constant factor already evaluated with our second helper. So, let's concentrate on the first commutator. In component notation we have

$$\begin{aligned} \epsilon_{ijk}\left[p_j L_k - L_j p_k\right]\mathcal{H} &= \epsilon_{ijk}\left[p_j \mathcal{H} L_k - L_j \mathcal{H} p_k + iZe^2\hbar L_j \frac{x_k}{r^3}\right] \\ &= \epsilon_{ijk}\left[\mathcal{H} p_j L_k - \mathcal{H} L_j p_k + iZe^2\hbar\left(L_j\frac{x_k}{r^3} - \frac{x_j}{r^3}L_k\right)\right] \end{aligned} \tag{533}$$

and we see, that the commutator's only survivor is the term in the round brackets. Using now, that $L_n = \epsilon_{lmn} x_l p_m$ we find

$$\begin{aligned} \epsilon_{ijk}\left[p_j L_k - L_j p_k, \mathcal{H}\right] &= iZe^2\hbar\epsilon_{ijk}\left(\epsilon_{jmn} x_m p_n \frac{x_k}{r^3} - \epsilon_{kmn}\frac{x_j}{r^3} x_m p_n\right) \\ &= -(i\hbar)^2 Ze^2 \epsilon_{ijk}\left[\left(\epsilon_{jmn}\frac{x_m x_k}{r^3} - \epsilon_{kmn}\frac{x_j x_m}{r^3}\right)\partial_n \right. \\ &\qquad \left. + \epsilon_{jmk}\frac{x_m}{r^3} + 3\epsilon_{jmn}\frac{x_k x_m x_n}{r^3}\right] \end{aligned} \tag{534}$$

Obviously, the last term vanishes, since $\epsilon_{jmn} x_m x_n \to 0$, and the first term can be rewritten using the fact, that we can exchange indices j and k. This first term can be further evaluated as

$$\begin{aligned} -2Ze^2(i\hbar)^2\epsilon_{ijk}\epsilon_{jmn}\frac{x_m x_k}{r^3}\partial_n &= -2Ze^2(i\hbar)^2\left(\delta_{km}\delta_{in} - \delta_{kn}\delta_{im}\right)\frac{x_m x_k}{r^3}\partial_n \\ &= 2Ze^2(i\hbar)^2\left(\frac{x_i}{r^3}x_k\partial_k - \frac{1}{r}\partial_i\right) . \end{aligned} \tag{535}$$

The second term is

$$-Ze^2(i\hbar)^2\epsilon_{ijk}\epsilon_{jmn}\frac{x_m}{r^3} = -Ze^2(i\hbar)^2\left(\frac{x_i}{r^3} - 3\frac{x_i}{r^3}\right) = 2Ze^2(i\hbar)^2\frac{x_i}{r^3}\,. \quad (536)$$

Collecting terms, we have

$$\frac{1}{2\mu}\epsilon_{ijk}\left[p_jL_k - L_jp_k, \mathcal{H}\right] = \frac{Ze^2}{\mu}(i\hbar)^2\left(\frac{x_i}{r^3} - \frac{1}{r}\partial_i + \frac{x_i}{r^3}x_k\partial_k\right) \quad (537)$$

and find, that indeed, $\vec{M}'$ commutes with $\mathcal{H}$.

A.2.2 $\vec{L}\cdot\vec{M}'$

Let us start with

$$\begin{aligned}\vec{L}\cdot\frac{\vec{r}}{r} &= \frac{L_ix_i}{r} = \epsilon_{imn}\frac{x_mp_nx_i}{r} \\ &= \epsilon_{imn}\frac{p_nx_mx_i + i\hbar\delta_{mn}x_i}{r} = 0 \end{aligned} \quad (538)$$

and continue with

$$\begin{aligned}\vec{L}\cdot\left(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}\right) &= \epsilon_{ijk}L_i\left(p_jL_k - L_jp_k\right) \\ &= \epsilon_{ijk}L_i\left(\epsilon_{kmn}p_jx_mp_n - \epsilon_{jmn}x_mp_np_k\right) \\ &= L_i\left(p_jx_ip_j - p_jx_jp_i + x_jp_ip_j - x_ip_jp_j\right) \\ &= 2i\hbar L_ip_i = 2i\hbar\epsilon_{ijk}p_ix_jp_k = 0\,. \end{aligned} \quad (539)$$

Therefore,

$$\vec{L}\cdot\vec{M}' = 0\,. \quad (540)$$

A.2.3 $[M_i', L_j]$

First, we determine

$$[L_i, p_j] = i\hbar\epsilon_{ijk}p_k\,. \quad (541)$$

Therefore

$$\begin{aligned}\frac{1}{2}\left[\left(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}\right)_x, L_y\right] &= \left[p_yL_z - L_yp_z, L_y\right] \\ &= i\hbar\left(-p_yL_x + L_yp_x\right)\,. \end{aligned} \quad (542)$$

Similarly,

$$\begin{aligned}
\frac{1}{2}\left[\left(\vec{p}\times\vec{L}-\vec{L}\times\vec{p}\right)_x, L_x\right] &= [p_y L_z - L_y p_z, L_x] \\
&= i\hbar\left(p_y L_y - L_y p_y\right) = 0\,. \qquad (543)
\end{aligned}$$

We are left with the terms

$$\begin{aligned}
\left[\frac{x}{r}, L_y\right] &= \hbar\frac{z}{r} \\
\left[\frac{x}{r}, L_x\right] &= 0\,, \qquad (544)
\end{aligned}$$

and gluing everything together, we recover the commutation relations

$$[M'_x, L_y] = i\hbar M'_z \;\text{ and }\; [M'_x, L_x] = 0\,. \qquad (545)$$

A.2.4 $\left[M'_i, M'_j\right]$ and M'^2

The commutator is essentially a brute force calculation in the spirit of the previous proof, so we leave it for the fearless reader. Instead, we calculate the Casimir–like operator M'^2. We start this by noting, that

$$\begin{aligned}
\left(\vec{L}\times\vec{p}\right)_i &= \epsilon_{ijk}L_j p_k = \epsilon_{ijk}\left(p_k L_j + \epsilon_{jkm}p_m\right) \\
&= -\left(\vec{p}\times\vec{L} + 2i\hbar\vec{p}\right)_i \qquad (546)
\end{aligned}$$

Hence, we can rewrite

$$\begin{aligned}
M'^2 &= \left[\frac{1}{\mu}\left(\vec{p}\times\vec{L} - i\hbar\vec{p}\right) - Ze^2\frac{\vec{r}}{|\vec{r}|}\right]^2 \\
&= \frac{1}{\mu^2}\left[\vec{p}^2\vec{L}^2 - 2(i\hbar)^2\vec{p}^2 - \hbar^2\vec{p}^2\right] \\
&\quad -\frac{Ze^2}{\mu}\left[\frac{2\vec{L}^2 - i\hbar\vec{r}\cdot\vec{p} + i\hbar\vec{r}\cdot\vec{p}}{r}\right] + Z^2e^4 \\
&= \frac{1}{\mu^2}\left[\vec{p}^2(\vec{L}^2+\hbar^2)\right] - \frac{ZE^2}{\mu}\frac{2(\vec{L}^2+\hbar^2)}{r} + Z^2E^4 \\
&= \frac{2\mathcal{H}}{\mu}\,(\vec{L}^2+\hbar^2) + Z^2e^4\,. \qquad (547)
\end{aligned}$$

B Young tableaux

This section is devoted to the introduction of another notation for the irreducible representations of $SU(3)$, which is extremely easy to handle and to calculate with. For $SU(3)$ these Young tableaux are nothing more than a very elegant diagrammatic notation incorporating the language of tensor products. The main virtues are, that we find an easy algorithm for the Clebsch–Gordan decomposition of tensor products of two representations, and that it generalizes straight forward to the case of $SU(N)$.

B.1 The algorithm for $SU(3)$

The algorithm is based on the observation, that in $SU(3)$ the $\bar{\mathbf{3}}$ is an antisymmetric combination of two $\mathbf{3}$'s. Ergo, we can write an arbitrary representation as a tensor product of $\mathbf{3}$'s with appropriate symmetry properties. To illustrate this point, let us consider the $(n,\, m)$ representation. Basically it is a tensor with components

$$A^{i_1 i_2 \ldots i_n}_{j_1 j_2 \ldots j_m} \,, \tag{548}$$

symmetric in upper and lower indices and traceless. We can now raise the lower indices with the help of a bunch of ϵ tensors and obtain an object with $n+2m$ upper indices, namely

$$\tilde{A}^{i_1 i_2 \ldots i_n k_1 k_2 \ldots k_m l_1 l_2 \ldots l_m} = A^{i_1 i_2 \ldots i_n}_{j_1 j_2 \ldots j_m} \epsilon^{j_1 k_1 l_1} \epsilon^{j_2 k_2 l_2} \ldots \epsilon^{j_m k_m l_m} \,. \tag{549}$$

Obviously, the $\tilde{A}$ is antisymmetric in each pair of indices $k_x \leftrightarrow l_x$ with $x = 1, \ldots, m$. Let us explore the symmetry properties of this object by considering the states

$$\tilde{A}^{i_1 i_2 \ldots i_n k_1 k_2 \ldots k_m l_1 l_2 \ldots l_m} \, |_{i_1 i_2 \ldots i_n k_1 k_2 \ldots k_m l_1 l_2 \ldots l_m}\rangle \tag{550}$$

forming the $(n,\, m)$ representation. We remember, that the state $|_1\rangle$ has the highest weight $(1/2,\, 1/2\sqrt{3})$ in the $\mathbf{3}$ representation. The state with next highest weight in this representation is $|_1\rangle$ with weight $(0,\, -1/\sqrt{3})$. Ergo, the highest weight in the $(n,\, m)$ represenstation is obtained, when all i's equal 1, and each pair of k, l takes values 1 and 3. For the state with highest weight non–zero tensor components $\tilde{A}$ are obtained from Eq.549 by setting all i's and all k's to 1 and all l's to 3 and by anti–symmetrization in all k, l pairs. Starting from this state, we can reach all other states by applying lowering operators. Let us now translate this into Young tableaux. In this formulation, all indices are associated with boxes, which we arrange as

$$\begin{array}{|c|c|c|c|c|c|}\hline k_1 & \dots & k_m & i_1 & \dots & i_n \\ \hline l_1 & \dots & l_m \\ \cline{1-3}\end{array} \tag{551}$$

This Young tableau then translates into our tensor by first symmetrizing in each row and by finally antisymmetrizing the indices in each column. After such a treatment, the resulting tensor corresponds to a state in the $(n, ,m)$ representation. This antisymmetrizing in the columns has the direct consequence, that columns with four or more boxes vanish, since there are only three indices to antisymmetrize in. Moreover, every column of three boxes simply yields an ϵ tensor in the three corresponding indices and gives a factor of 1,

$$\epsilon^{ijk} = \begin{array}{|c|}\hline i \\ \hline j \\ \hline k \\ \hline\end{array} \tag{552}$$

Thus, tableaux of the form

$$\begin{array}{|c|c|c|c|c|c|c|c|c|}\hline & & & k_1 & \dots & k_m & k_1 & \dots & k_m \\ \hline & & & l_1 & \dots & l_m \\ \cline{1-6} & & \\ \cline{1-3}\end{array} \tag{553}$$

are equivalent to the one depicted in Eq.551.
With help of the Young tableaux we can easily perform the Clebsch–Goredan decomposition of tensor products. The algorithm is as follows :

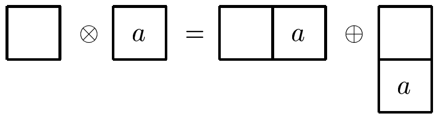
1. We start by multiplying two tableaux A and B corresponding to the product of representations $\alpha \otimes \beta$. So, B is to the left of A.

2. Denote all boxes in the top row of B by b_1 and all boxes in the second row by b_2. Remember, that an eventual third row is obsolete.

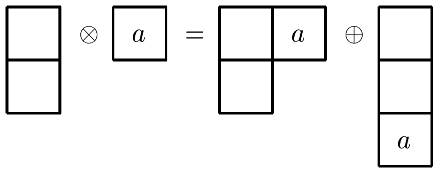
3. Take the boxes of B's top row and add them to A. Here, each of the b_1's has to be placed in a different column. Repeat the same procedure for the second row of B.

4. Neglect all of the newly formed tableaux, in which, reading from right to left or from top to bottom the number of b_1's is smaller than the number of b_2's. With this step, double counting of tensors is prevented.

Let us do some trivial examples :

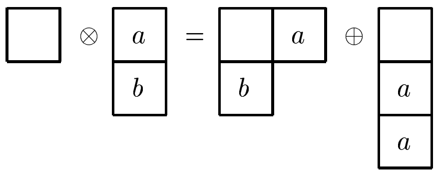


$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}$$



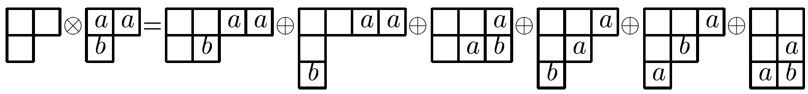
$$\bar{\mathbf{3}} \otimes \mathbf{3} = \mathbf{8} \oplus \mathbf{1} \tag{554}$$

A little bit less trivial, but with the same result as the last product is



$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1} \tag{555}$$

Finally, let us do something more involved,



$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{10} \oplus \overline{\mathbf{10}} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1} \tag{556}$$

Note, that the two $\mathbf{8}$'s occuring in this last example are different, since they have a different pattern of a's and b's filled into the boxes. Note, thast using the formalism of Young tableaux, it is quite easy to show, that

$$\mathbf{d} \in \mathbf{8} \otimes \mathbf{d} \ \text{ for all } \ \mathbf{d}\,. \tag{557}$$

C $SU(6)$ and the quark model for hadrons

C.1 Introducing $SU(6)$

Consider once more the lowest lying baryons, which are organized in one octet of spin–1/2 and one decuplet of spin–3/2 particles. The splittings between these two multiplets are not much larger than the mass difference within the multiplets themselves. This suggests, that we might combine the octet and the decuplet in one irreducible representation of some group, which is larger than $SU(3)$. Because the octet and the decuplet have different spin, this alleged big group can not commute with angular momentum. This means, that we mix up internal symmetries – i.e. flavor – with spacetime symmetries, and we should expect this to cause severe problems. In fact, these problems occur, but only, when we try to formulate this within a relativistic theory. But this is beyond the scope of the lecture anyhow. In the language of quark models, we have all together 56 baryon states, (2×8 for the octet with spin either up or down, and 4×10 for the decuplet with four possible spin configurations). These states correspond to 56 completely symmetric combinations of three spin–1/2 quarks. The approximate degeneracy of the octet and the decuplet suggests, that the forces that bind the quarks into the baryons are not only approximately $SU(3)$–invariant, but also approximately spin independent. This leads us to the idea, that we should look into the group $SU(6)$, where its six–dimensional representation consists of six quark states. In the tensor notation we can label these states as products of $SU(3)$– and $SU(2)$–indices like

$$\left\{ |u\rangle,\ |d\rangle,\ |s\rangle \right\} \otimes \left\{ \left|\frac{1}{2}\right\rangle \, ; \left|-\frac{1}{2}\right\rangle \right\} . \tag{558}$$

Thus, the generators are products of the $SU(3)$ 3×3 matrices and the 2×2 matrices in spin space. They include the Gell–Mann matrices and the spin angular momentum generators with identity in the respective complementary part understood. In other words, this $SU(6)$ has an $SU(3)$ subalgebra under which the **6** transforms like 3 **2**'s, and an $SU(2)$ subalgebra, under which the **6** transforms like 2 **3**'s. The generators of both subalgebras commute with each other. Hence, the $SU(6)$ has an $SU(3) \otimes SU(2)$ subalgebra, under which the **6** transforms as (3, 2). The rest of the generators are the products

$$\frac{1}{2}\lambda_a \sigma_j \, , \tag{559}$$

for a total of $8 \cdot 3 + 8 + 3$ generators.

Let us start to check our findings. The first, trivial check is that there really is a 56 dimensional representation with the wanted properties. The completely symmetric combination of 3 **6**'s is indeed a **56**. We can combine the 3 **3**'s and the 3 **2**'s symmetrically into a (10, 4) or we can have antisymmetry in one pair of both $SU(3)$ and $SU(2)$ indices yielding a (8, 2). All together we find

$$(3,\,2) \otimes (3,\,2) \otimes (3,\,2)_{\mathrm{sym.}} = (10,\,4) \oplus (8,\,2) = \mathbf{40} + \mathbf{16}\,. \tag{560}$$

Let us talk about the states now. We will use tensor product notation. In this notation, the decuplet states are particularly simple, symmetric in both $SU(3)$ and $SU(2)$ indices. For instance

$$\begin{aligned}
\left|\Delta^{++},\,\frac{3}{2}\right\rangle &= |uuu\rangle\,|\uparrow\uparrow\uparrow\rangle \\
\left|\Delta^{+},\,\frac{1}{2}\right\rangle &= \{|uud\rangle + |udu\rangle + |duu\rangle\} \cdot \{|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle\} \\
\left|\Sigma^{*0},\,\frac{1}{2}\right\rangle &= \frac{1}{3\sqrt{2}}\,\{|uds\rangle + 5\,\mathrm{perms.}\} \\
&\qquad \times \{|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle\}\,.
\end{aligned} \tag{561}$$

The octet states are somewhat tricky. To get a completely symmetric state, we have to multiply an $SU(3)$ state, which is antisymmetric in a pair of indices with a similarly antisymmetric $SU(2)$ state and add all cyclic permutations. For example

$$\begin{aligned}
\left|\Lambda,\,\frac{1}{2}\right\rangle &= \frac{1}{2\sqrt{3}}\,[(|uds\rangle - |dus\rangle) \cdot (|\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle) \\
&\qquad (|sud\rangle - |sdu\rangle) \cdot (|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle) \\
&\qquad (|dsu\rangle - |usd\rangle) \cdot (|\downarrow\uparrow\uparrow\rangle - |\uparrow\uparrow\downarrow\rangle)]\,.
\end{aligned} \tag{562}$$

Alternatively, we might take an $SU(3)$ state, which is symmetric in one pair of indices, and multiply it with an $SU(2)$ state symmetric in the same pair but with total spin 1/2 and add all cyclic permutations. Then we have

$$\begin{aligned}
\left|p,\,\frac{1}{2}\right\rangle &= \frac{1}{3\sqrt{2}}\,\Big[|uud\rangle \cdot (2|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle) \\
&\qquad + \mathrm{cycl.}\Big]\,.
\end{aligned} \tag{563}$$

C.2 Magnetic moments of baryons

What is our reasoning so far good for ? Let us take our $SU(6)$ and try to figure out something about the magnetic moments of the baryons. Let us start with the magnetic moment of a point-like quark. Given its mass m and charge Q, the magnetic moment is

$$\vec{\mu} = \frac{1}{2m} \cdot Q\vec{\sigma}\,. \tag{564}$$

But this is a generator, and we can conclude, that $\vec{\mu}$ transforms like the 35 dimensional adjoint representation of $SU(6)$. Thus, we are interested in matrix elements

$$\langle \mathbf{56}|\mathbf{35}|\mathbf{56}\rangle\,. \tag{565}$$

Note, that $\mathbf{35} \otimes \mathbf{56}$ contains one unique $\mathbf{56}$, so these matrix elements are determined up to an overall factor. Thus, they all are proportional to

$$\langle \mathbf{56}|Q\vec{\sigma}|\mathbf{56}\rangle\,. \tag{566}$$

We can use this to calculate the ratio of the magnetic moments of the proton and the neutron, μ_p and μ_n. Since neutrons are obtained from the proton states, Eq.563, by replacing u by d and vice versa (and ignoring possible irrelevant overall phases), they are eigenstates of σ_3. Thus, we only need to calculate the expectation values of $Q\sigma_3$. The other components will vanish. Furthermore, since the baryons are products in our tensor language, the generator acts on them as a sum over the three quark states. For example,

$$\begin{aligned} Q\sigma_3\left|p,\frac{1}{2}\right\rangle \quad = \quad & \frac{1}{3\sqrt{2}}\left[\frac{2}{3}|uud\rangle\cdot(2|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle)\right. \\ & +\frac{2}{3}|uud\rangle\cdot(2|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle) \\ & -\frac{1}{3}|uud\rangle\cdot(-2|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle) \\ & \left. + \text{cycl.}\right]\,. \end{aligned} \tag{567}$$

This leads to

$$\begin{aligned} \left\langle p,\frac{1}{2}\right| Q\sigma_3 \left|p,\frac{1}{2}\right\rangle \quad &= \quad 1\,, \\ \left\langle n,\frac{1}{2}\right| Q\sigma_3 \left|n,\frac{1}{2}\right\rangle \quad &= \quad -\frac{2}{3}\,, \end{aligned} \tag{568}$$

and we should have the ratio

$$\frac{\mu_p}{\mu_n} = -\frac{3}{2}\,. \tag{569}$$

Experimentally we find (in nuclear magnetons)

$$\mu_p = 2.79\ , \quad \mu_n = -1.91\ , \quad \frac{\mu_p}{\mu_n} = -1.46\,. \tag{570}$$

It works. But actually, it was not really big fun to calculate this within the $SU(6)$ formulation. From the quarks we took only the $SU(6)$ transformation law of the magnetic moment operators. Then we used the Wigner–Eckart theorem to show, that we can in fact calculate the ratio of these operators when sandwiched between the **56** states of $SU(6)$. Finally, we used the explicit form of the $SU(6)$ wave functions of the baryons to evaluate the ratios.

C.3 Quark model answer

In turn, employing the quark model alone, we just have to perform one step based on physics. We just trust on the physical existence of quarks and assume the baryon to be a bound state of three such quarks with no angular momentum. In this picture, the magnetic moment of the baryon is just the sum of the magnetic moments of the quarks. So, in tensor product language we have

$$\mu = \sum_q \frac{1}{2m_q}\, Q_q \vec{\sigma}\,, \tag{571}$$

which is just, what we had before plus the overall scale of the magnetic moment given by the quark masses. Assuming $m_u \approx m_d = m$ and $3m \approx m_p$ we have

$$\frac{\mu_p}{2m_p} = \left\langle p, \frac{1}{2} \right| \frac{1}{2m} \sum_q Q\vec{\sigma} \left| p, \frac{1}{2} \right\rangle = \frac{1}{2m}\,. \tag{572}$$

In this picture,

$$\mu_p = \frac{m_p}{m} \approx 3\,, \tag{573}$$

which is not too bad given the simplicity of our arguments. In turn, this simple formulation of the quark model predicts

$$\mu_n \approx -2 \tag{574}$$

yielding the same ratio as our involved $SU(6)$ calculation plus the correct order of magnitude of the individual magnetic moments.
Let us dwell on the predictive power of the quark model a little longer. Looking at the Λ we find, that

$$\left\langle \Lambda, \frac{1}{2} \middle| Q\sigma_3 \middle| \Lambda, \frac{1}{2} \right\rangle \quad = \quad -\frac{1}{3}\,, \tag{575}$$

so the $SU(6)$ prediction for μ_Λ is

$$\frac{\mu_\Lambda}{\mu_p} = -\frac{1}{3} \quad \rightarrow \quad \mu_\Lambda = -0.93\,. \tag{576}$$

In turn, the quark model predicts

$$\frac{\mu_\Lambda}{2m_\Lambda} = \left\langle \Lambda, \frac{1}{2} \middle| \sum_q \frac{1}{2m_q} Q_q \vec{\sigma} \middle| \Lambda, \frac{1}{2} \right\rangle \tag{577}$$

and we have to determine the strange quark mass. Assuming

$$m_\Lambda = 2m + m_s \;\text{ and }\; m_p = 3m \tag{578}$$

we have

$$m_s = \frac{3m_\Lambda - 2m_p}{3} \approx 500\text{MeV}\,. \tag{579}$$

Closer inspection of the Λ's wave function tells us, that only the s quark contributes – the contributions of the u and the d quark cancel out,

$$\frac{\mu_\Lambda}{2m_\Lambda} = -\frac{1}{3}\frac{1}{2m_s}\,. \tag{580}$$

This tells us, that

$$\mu_\Lambda = -\frac{1}{3}\frac{m_\Lambda}{m_s} \approx 0.74\,, \tag{581}$$

which is actually in better agreement with the measured value of $\mu_\Lambda = 0.93$.