Symmetries in Physics

Nicolas Borghini
Nicolas Borghini
Universität Bielefeld, Fakultät für Physik
Homepage: http://www.physik.uni-bielefeld.de/~borghini/
Email: borghini at physik.uni-bielefeld.de
I Basics of group theory

I.1 Group structure

I.2 Examples of groups

I.2.1 Finite groups

I.2.2 Infinite groups

I.3 Subgroups

I.3.1 Definition

I.3.2 Generating set

I.3.3 Cosets of a subgroup

I.4 Conjugacy and conjugacy classes

I.4.1 Equivalence relation

I.4.2 Conjugacy

I.4.3 Examples of conjugacy classes

I.5 Normal subgroups

I.5.1 Definition

I.5.2 Quotient group

I.5.3 Examples

I.5.4 Further definitions and results

I.6 Direct product

I.6.1 Cartesian product of groups

I.6.2 Inner direct product

I.7 Group homomorphisms

I.7.1 Definitions

I.7.2 Properties of homomorphisms

I.7.3 Inner automorphisms

II Representation theory

II.1 Group action

II.2 Group representation: first definitions and results

II.3 Examples of group representations

II.3.1 Trivial representation

II.3.2 Representations of the dihedral group $D_3$

II.3.3 Three-dimensional rotations

II.3.4 Direct sum and tensor product of representations

II.4 Classifying representations

II.4.1 Equivalent representations

II.4.2 Reducible and irreducible representations

II.5 Schur’s lemma

III Representations of finite groups

III.1 Full reducibility of the representations

III.2 Orthogonality relations

III.3 Irreducible representations of a finite group

III.3.1 First consequences of the orthogonality relations

III.3.2 Character theory
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>III.4</td>
<td>Reduction of a representation</td>
<td>40</td>
</tr>
<tr>
<td>III.4.1</td>
<td>Decomposition of a representation into irreducible representations</td>
<td>40</td>
</tr>
<tr>
<td>III.4.2</td>
<td>Regular representation</td>
<td>41</td>
</tr>
<tr>
<td>III.4.3</td>
<td>Reduction of a tensor-product representation</td>
<td>45</td>
</tr>
<tr>
<td>III.5</td>
<td>Representations of the symmetric group $S_n$</td>
<td>46</td>
</tr>
<tr>
<td>III.5.1</td>
<td>Irreducible representations of $S_n$</td>
<td>46</td>
</tr>
<tr>
<td>III.5.2</td>
<td>Product of irreducible representations of the symmetric group</td>
<td>49</td>
</tr>
<tr>
<td>IV</td>
<td>Physical applications</td>
<td>52</td>
</tr>
<tr>
<td>IV.1</td>
<td>Electric dipole of a molecule</td>
<td>52</td>
</tr>
<tr>
<td>IV.2</td>
<td>Molecular vibrations</td>
<td>52</td>
</tr>
<tr>
<td>IV.3</td>
<td>Crystal field splitting</td>
<td>53</td>
</tr>
<tr>
<td>V</td>
<td>Lie groups and Lie algebras</td>
<td>58</td>
</tr>
<tr>
<td>V.1</td>
<td>Lie groups</td>
<td>58</td>
</tr>
<tr>
<td>V.1.1</td>
<td>Definition and first results</td>
<td>58</td>
</tr>
<tr>
<td>V.1.2</td>
<td>Examples</td>
<td>59</td>
</tr>
<tr>
<td>V.2</td>
<td>Lie algebras</td>
<td>63</td>
</tr>
<tr>
<td>V.2.1</td>
<td>General definitions</td>
<td>63</td>
</tr>
<tr>
<td>V.2.2</td>
<td>Lie algebra associated to a Lie group</td>
<td>65</td>
</tr>
<tr>
<td>V.2.3</td>
<td>Lie algebras associated to the classical Lie groups</td>
<td>68</td>
</tr>
<tr>
<td>VI</td>
<td>The groups SO(3) and SU(2) and their representations</td>
<td>69</td>
</tr>
<tr>
<td>VI.1</td>
<td>The group SO(3)</td>
<td>69</td>
</tr>
<tr>
<td>VI.1.1</td>
<td>Rotations in three-dimensional space</td>
<td>69</td>
</tr>
<tr>
<td>VI.1.2</td>
<td>The group SO(3) and its Lie algebra $\mathfrak{so}(3)$</td>
<td>71</td>
</tr>
<tr>
<td>VI.2</td>
<td>The group SU(2)</td>
<td>74</td>
</tr>
<tr>
<td>VI.2.1</td>
<td>The group SU(2) and its Lie algebra $\mathfrak{su}(2)$</td>
<td>74</td>
</tr>
<tr>
<td>VI.2.2</td>
<td>Relation between the Lie groups SU(2) and SO(3)</td>
<td>75</td>
</tr>
<tr>
<td>VI.3</td>
<td>Representations of the Lie groups SU(2) and SO(3)</td>
<td>76</td>
</tr>
<tr>
<td>VI.3.1</td>
<td>Representations of SU(2) and of the Lie algebra $\mathfrak{su}(2)$</td>
<td>76</td>
</tr>
<tr>
<td>VI.3.2</td>
<td>Unitary representations of SU(2)</td>
<td>78</td>
</tr>
<tr>
<td>VI.3.3</td>
<td>Tensor products of irreducible representations of SU(2)</td>
<td>83</td>
</tr>
<tr>
<td>VI.4</td>
<td>Behavior of quantum-mechanical observables under rotations</td>
<td>85</td>
</tr>
<tr>
<td>VI.4.1</td>
<td>Rotation of a quantum-mechanical system</td>
<td>85</td>
</tr>
<tr>
<td>VI.4.2</td>
<td>Irreducible tensor operators</td>
<td>87</td>
</tr>
<tr>
<td>VI.4.3</td>
<td>Wigner–Eckart theorem</td>
<td>89</td>
</tr>
<tr>
<td>VII</td>
<td>Representations of GL($n$) and its continuous subgroups</td>
<td>91</td>
</tr>
<tr>
<td>VII.1</td>
<td>Low-dimensional irreducible representations of GL($n$, $\mathbb{C}$)</td>
<td>91</td>
</tr>
<tr>
<td>VII.1.1</td>
<td>Representations of dimension</td>
<td>91</td>
</tr>
<tr>
<td>VII.1.2</td>
<td>Vector representation</td>
<td>91</td>
</tr>
<tr>
<td>VII.1.3</td>
<td>Complex conjugate representation</td>
<td>91</td>
</tr>
<tr>
<td>VII.1.4</td>
<td>Dual representation</td>
<td>91</td>
</tr>
<tr>
<td>VII.2</td>
<td>Tensor representations of GL($n$, $\mathbb{C}$)</td>
<td>92</td>
</tr>
<tr>
<td>VII.2.1</td>
<td>Tensor products of the vector representation</td>
<td>92</td>
</tr>
<tr>
<td>VII.2.2</td>
<td>Operations on the indices of a tensor</td>
<td>93</td>
</tr>
<tr>
<td>VII.2.3</td>
<td>Irreducible representations built from the vector representation</td>
<td>94</td>
</tr>
<tr>
<td>VII.2.4</td>
<td>Tensor products of the dual and conjugate representations</td>
<td>97</td>
</tr>
<tr>
<td>VII.2.5</td>
<td>Representations of SL($n$, $\mathbb{C}$) and SL($n$, $\mathbb{R}$)</td>
<td>99</td>
</tr>
<tr>
<td>VII.2.6</td>
<td>Representations of SU($n$)</td>
<td>100</td>
</tr>
</tbody>
</table>
CHAPTER I
Basics of group theory

An intro is missing

I.1 Group structure

Definition I.1. A group \((G, \cdot)\) consists of a set \(G\) and a binary operation \(\cdot\), the group law, that fulfills the following axioms:

- **closure**: for all \(g, g' \in G\), the result \(g \cdot g'\) is also in \(G\); \((G1)\)
- **associativity**: for all \(g, g', g'' \in G\), \((g \cdot g') \cdot g'' = g \cdot (g' \cdot g'')\); \((G2)\)
- there exists a neutral element \(e\) such that for every \(g \in G\), \(e \cdot g = g \cdot e = g\); \((G3)\)
- for every \(g \in G\) there is an inverse element \(g^{-1} \in G\) such that \(g^{-1} \cdot g = g \cdot g^{-1} = e\). \((G4)\)

Remarks:

* The closure axiom \((G1)\) is often omitted, at the cost of calling the group law an “internal composition law”.

* The associativity property \((G2)\) means that parentheses are not needed — as long as one only considers the group law.

* The axioms \((G3)\) and \((G4)\) can be replaced by weaker, yet equivalent versions, namely the existence of a left neutral element \(e\) such that \(e \cdot g = g\) for every \(g \in G\) and that of a left inverse \(g^{-1} \in G\) to \(g\) such that \(g^{-1} \cdot g = e\). These are then automatically also right neutral element and right inverse, respectively.

* The postulated existence of a neutral element, also called unit or identity element, guarantees that the set \(G\) is not empty.

* In the following, the result \(g \cdot g'\) will also be written as \(gg'\), and referred to as the “product” or the “composition” of the elements \(g\) and \(g'\). Accordingly, \(g \cdot g\) will be written as \(g^2\) and more generally \(g \cdot g^n\) with \(n \in \mathbb{N^*}\) as \(g^{n+1}\).

* Similarly, in keeping with a widespread habit I shall regularly refer to “a group \(G\)”, without specifying the group law. This assumes either that the latter is obvious, or that it is the generic group law of an unspecified group.

Theorem I.2. Every group \(G\) contains a single neutral element, and each element \(g \in G\) has a unique inverse element \(g^{-1}\).
Theorem I.3. For every triplet $g$, $g'$, $g''$ of elements of a group $(G, \cdot)$, the equality $g \cdot g' = g \cdot g''$ (or $g' \cdot g = g'' \cdot g$) implies $g' = g''$.

Definition I.4. A group $(G, \cdot)$ is called Abelian\(^{(a)}\) if the group law is commutative, i.e. if for all $g, g' \in G$, $g \cdot g' = g' \cdot g$.

Remark: The group law of an Abelian group is often denoted with $+$.  

Definition I.5. The number of elements in the set $G$ of a group $(G, \cdot)$ is called the order of the group and will be hereafter denoted by $|G|$.

Depending on the finiteness of the group order one distinguishes between finite groups ($|G| \in \mathbb{N}^*$) and infinite groups. In turn, infinite groups can either be countable or not countable, where the latter can further be split into compact and non-compact groups.\(^{(1)}\)

In the case of a finite group $(G, \cdot)$, one can (in principle) write down all possible compositions $g \cdot g' = gg'$ in a Cayley\(^{(b)}\) table, i.e. as

\[
\begin{array}{ccccccc}
  & g_1 & g_2 & \cdots & g_j & \cdots & g_n \\
 g_1 & (g_1)^2 & g_1g_2 & \cdots & g_1g_j & \cdots & g_1g_n \\
g_2 & g_2g_1 & (g_2)^2 & \cdots & g_2g_j & \cdots & g_2g_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
g_i & g_ig_1 & g_ig_2 & \cdots & g_ig_j & \cdots & (g_i)^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
g_n & g_ng_1 & g_ng_2 & \cdots & g_ng_j & \cdots & (g_n)^2 \\
\end{array}
\tag{1.6}
\]

with $G = \{g_1, g_2, \ldots, g_n\}$, where the ordering of the group elements is irrelevant — although the “first” element in the examples listed hereafter (§ I.2.1a) will always be the identity element of the group. Note that the entry in the $i$-th line and $j$-th column is by convention the product $g_ig_j$.

One easily proves that each element $g \in G$ appears only once in every line and every column of the group (Cayley) table. This is a particular case of the following result, valid for both finite and infinite groups:

Theorem I.7 (Group rearrangement theorem). Let $G$ be a group and $g \in G$ one of its element. Defining a set $gG = \{gg' \mid g' \in G\}$, one has $gG = G$.

Proof: Let us show that both inclusions $gG \subseteq G$ and $gG \supseteq G$ are fulfilled, starting with the first one. Consider an arbitrary $g' \in G$. Then $gg' \in G$ and therefore $gG \subseteq G$.

Conversely, take an arbitrary $g' \in G$. The element $g^{-1}g'$ is again in $G$ and so is $gg^{-1}g' = g'$, which shows $gG \supseteq G$.

\(^{(1)}\)These terms will be defined and discussed in further detail in Chapter V

\(^{(a)}\)N. H. Abel, 1802–1829 \hspace{1em} \(^{(b)}\)A. Cayley, 1821–1895
I.2 Examples of groups

We now list a few examples of groups, starting with finite ones (Sec. I.2.1) and then going on to infinite ones (Sec. I.2.2). Further examples will be given later in this chapter as well as in the next chapters.

I.2.1 Finite groups

I.2.1 a Small finite groups

Order 1

One may define a group structure on any set \( S \) consisting of a single element \( a \), i.e. \( S = \{a\} \), by defining an internal composition law \( \cdot \) on \( S \) by \( a \cdot a = a \). \((S, \cdot)\) trivially fulfills the 4 group axioms \((G)\).

As “symmetry group” of the transformations that leave a physical or mathematical object invariant, a group of order 1 is obviously quite boring and corresponds to an object having “no symmetry”, since the only transformation leaving it unchanged is the identity transformation.

Order 2

A first example of a group of order 2 consists of the set \( \{0, 1\} \), where 0 and 1 denote the “usual” natural numbers, while the group law is the addition modulo 2, i.e. such that \( 1 + 1 = 0 \). The corresponding Cayley table is

\[
\begin{array}{c|cc}
+ & 0 & 1 \\
0 & 0 & 1 \\
1 & 1 & 0 \\
\end{array}
\]  

(I.8)

and the group is traditionally called \((\mathbb{Z}_2, +)\).

A second example is that of the set \( \{1, -1\} \) with the usual multiplication of integers, yielding the Cayley table

\[
\begin{array}{c|ccc}
\times & 1 & -1 \\
1 & 1 & -1 \\
-1 & -1 & 1 \\
\end{array}
\]  

(I.9)

A further example can be built as follows. Let \( \mathcal{F}(\mathbb{R}^3) \) denote a set of functions defined on \( \mathbb{R}^3 \) — the target set of the functions is irrelevant. One defines two operators \( \hat{I} \) and \( \hat{P} \) acting on this set, i.e. associating to a function \( f \in \mathcal{F}(\mathbb{R}^3) \) another function \( g \in \mathcal{F}(\mathbb{R}^3) \), by

\[
\hat{I} : f(\vec{r}) \mapsto g(\vec{r}) = f(\vec{r}) \quad \text{and} \quad \hat{P} : f(\vec{r}) \mapsto g(\vec{r}) = f(-\vec{r})
\]  

(I.10)

respectively. That is, \( \hat{I} \) is the identity operator, while \( \hat{P} \) reverses the argument of the function it acts upon. As internal composition law on the set \( \{\hat{I}, \hat{P}\} \) one considers the successive operation, i.e. the composition (of operators) often denoted by \( \circ \). One then at once sees that applying \( \hat{P} \) twice to a function \( f \) gives back the same function, i.e. \( \hat{P}^2 = \hat{P} \circ \hat{P} = \hat{I} \). More generally, one finds that \( \{(\hat{I}, \hat{P}), \circ\} \) is a group of order 2 with the Cayley table

\[
\begin{array}{c|ccc}
\circ & \hat{I} & \hat{P} \\
\hat{I} & \hat{I} & \hat{P} \\
\hat{P} & \hat{P} & \hat{I} \\
\end{array}
\]  

(I.11)

Inspecting the Cayley tables (I.8), (I.9), and (I.11) reveals that they are all similar, namely of the form

\[
\begin{array}{c|ccc}
\cdot & e & a \\
e & e & a \\
a & a & e \\
\end{array}
\]  

(I.12)
with \( e \) the neutral element of the group. This table describes an abstract group \( \{e, a\} \) with an internal composition law \( \cdot \) such that \( a^2 = e \), which is Abelian.

In turn, the groups \((\mathbb{Z}_2, +)\), \((\{1, -1\}, \times)\), and \((\{\mathcal{I}, \mathcal{G}\}, \circ)\) are various representations of this abstract group law — which happens to be the only possible one for a set with two elements, as is easily checked.

**Order 3**

Considering now groups of order 3, we shall start with a set \( \{e, a, b\} \) and show that there is only one possibility to define an internal composition law that turns it into a group, actually an Abelian group, with \( e \) as the identity element.

Writing down the Cayley table, the line and column associated with the identity element \( e \) follow automatically from axiom (G3):

\[
\begin{array}{ccc}
  & e & a & b \\
e & e & a & b \\
a & a & & \\
b & b & & \\
\end{array}
\]

Turning now to the second line, associated to \( a \), we know from the result preceding Theorem (I.7) that it should still contain \( e \) and \( b \), and that \( b \) may not appear in the third column, which already contains \( b \), so that necessarily \( a^2 = b \). A similar reasoning gives for the third line \( b^2 = a \) — which is quite normal, \( a \) and \( b \) play symmetric roles —, and thus

\[
\begin{array}{ccc}
  & e & a & b \\
e & e & a & b \\
a & a & b & \\
b & b & a & \\
\end{array}
\]

Eventually, the remaining entries can be filled in by realizing that \( e \) is still missing from the second and third lines, leading eventually to \( ab = ba = e \) and therefore to the table

\[
\begin{array}{ccc}
  & e & a & b \\
e & e & a & b \\
a & a & b & e \\
b & b & e & a \\
\end{array}
\]  

(I.13)

In particular, both \( a \) and \( b \) obey \( a^3 = b^3 = e \).

A possible representation of the abstract group (I.13) consists of the three rotations through angles \( 2k\pi/3 \) with \( k \in \{0, 1, 2\} \) — that with \( k = 0 \) being obviously the identity transformation — about an arbitrary point, e.g. the origin \( O \), of a (two-dimensional!) plane, with the composition of geometrical transformations as group law. Altogether, these three rotations constitute the symmetry group of an equilateral triangle centered on the origin with oriented sides, as in Fig. [I.1]

![Figure I.1](image-url)

As alternative groups of order 3, one may consider the set of complex numbers \( \{1, e^{2i\pi/3}, e^{4i\pi/3}\} \) with the usual product of complex numbers as group law, or the set \( \mathbb{Z}_3 \equiv \{0, 1, 2\} \) of integers with the addition modulo 3.
Order 4
Consider now a set of four elements \( \{e, a, b, c\} \). Proceeding as in the previous paragraph, one can investigate internal composition laws on this set that would turn it into a group of order 4 with identity element \( e \), by systematically building the Cayley table of the group. It turns out that there are now two different possible group laws, up to renaming of the elements.

A first possibility is that of a group with \( a^2 = b \) and \( a^3 = c \) or equivalently \( c^2 = b \) and \( c^3 = a \), i.e. in which two of the elements (a and c) play symmetric roles, while the third non-identity element (b) behaves differently since \( b^2 = e \) and thus \( b^3 = b \):

\[
\begin{array}{cccc}
  \cdot & e & a & b & c \\
  e & e & a & b & c \\
  a & a & b & c & e \\
  b & b & c & e & a \\
  c & c & e & a & b \\
\end{array}
\]

\( (I.14) \)

Note that this is again an Abelian group.

A possible representation of the abstract group \( (I.14) \) are the four rotations through angles \( k\pi/2 \) with \( k \in \{0, 1, 2, 3\} \) about an arbitrary point \( O \) of a plane, with the composition of geometrical transformations as group law: this is the symmetry group of a square with oriented sides centered on \( O \) (Fig. I.2). Alternatively, one may consider the set of complex numbers \( \{1, i, -1, -i\} \) with the usual product of complex numbers as group law, or the set \( \mathbb{Z}_4 \equiv \{0, 1, 2, 3\} \) of integers with the addition modulo 4.

There is a second possible group structure on a set with four elements, in which the three non-identity elements play symmetric roles: \( a \cdot b = b \cdot a = c, a \cdot c = c \cdot a = b, b \cdot c = c \cdot b = b \), and \( a^2 = b^2 = c^2 = e \):

\[
\begin{array}{cccc}
  \cdot & e & a & b & c \\
  e & e & a & b & c \\
  a & a & e & c & b \\
  b & b & c & e & a \\
  c & c & b & a & e \\
\end{array}
\]

\( (I.15) \)

This is again an Abelian group, known as the \textit{Klein} \((c)\) \textit{group} or \textit{Vierergruppe} (“four-group”) and often denoted by \( V_4 \).

A possible representation consists of geometrical transformations in a plane: the identity transformation, a (“two-dimensional”) rotation through \( \pi \) about a point \( O \), and the reflections across two perpendicular axes going through \( O \). This is the symmetry group of a non-equilateral rectangle (Fig. I.3). Alternatively, the reflections can be viewed as rotations through \( \pi \) about the two axes in three-dimensional space, while the “two-dimensional” rotation is now seen as a third “three-

\( ^{(c)} \) F. Klein, 1849–1925
dimensional” rotation about an axis perpendicular to the other two axes, restoring the symmetry between the three transformations.

I.2.1b Four families of finite groups

More generally, one can identify several families of finite groups, four of which we now list.

Cyclic groups $C_n$

The groups with respective Cayley tables (I.12), (I.13), (I.14) obviously have an important feature in common, namely the existence of a (non-identity) element $a$ such that the $n = 2, 3$ or 4 elements of the group can be expressed as $a, a^2, \ldots, a^{n-1}$, and $a^n = e$.

More generally, one defines such a group with $n \in \mathbb{N}^*$ elements, $C_n = \{a, a^2, \ldots, a^{n-1}, a^n = e\}$ (2) called the cyclic group of order $n$, with the Cayley table

\[
\begin{array}{cccccccc}
\cdot & e & a & a^2 & \cdots & a^{n-1} \\
e & e & a & a^2 & \cdots & a^{n-1} \\
a & a & a^2 & a^3 & \cdots & e \\
a^2 & a^2 & a^3 & a^4 & \cdots & a \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
a^{n-1} & a^{n-1} & e & a & \cdots & a^{n-2} \\
\end{array}
\]

(I.16)

These groups are clearly Abelian for any $n$.

One can generalize the representations given in § I.2.1a in the cases $n = 2, 3$ or 4 to representations of $C_n$. For instance, the rotations through angles $2k\pi/n$ with $k \in \{0, 1, \ldots, n-1\}$ about a fixed point constitute, with the composition of rotations, such a representation, which is the symmetry group of a regular polygon with $n$ oriented sides.

Alternatively, the abstract group $C_n$ can be represented by the set $\mathbb{Z}_n \equiv \{0, 1, 2, \ldots, n-1\}$ of integers with the addition modulo $n$, or by the set of complex numbers $\{1, e^{2\pi i/n}, e^{4\pi i/n}, \ldots, e^{2k\pi i/n}\}$ with the usual multiplication of complex numbers.

Dihedral groups $D_n$

Consider an regular polygon with $n \geq 3$ sides (3) centered on a point $O$. The geometric transformations leaving the polygon invariant are on the one hand the rotations $r_k$ through $2k\pi/n$ about the center $O$, with $k \in \{0, 1, \ldots, n-1\}$, and on the other hand $n$ reflections $s_k$ across lines going through $O$ and through either a vertex of the polygon or the midpoint of a side.

If $n$ is odd, the $n$ symmetry axes are indeed joining a vertex of the $n$-gon to the midpoint of the opposite side. If $n$ is even, there are $n/2$ diagonals joining opposite vertices, and $n/2$ lines joining the midpoints of opposite sides.

The composition of two rotations or two reflections gives a rotation, while the composition of a reflection and a rotation yield a reflection, and all these transformations together form a group. Denoting $s$ one of the reflections, any one of them, then all other reflections are of the form $sr_k$.

---

(2) The group is also denoted $\mathbb{Z}_n$.
(3) In contrast to the triangle and square pictures in Figs. I.1 I.2 the sides are not oriented.
In addition, \( r_0 \) is obviously the identity transformation \( e \) and denoting \( r \equiv r_1 \) one has \( r_k = r^k \) for \( k \geq 1 \), so that the group eventually is
\[
\{e, r, r^2, \ldots, r^{n-1}, s, sr, sr^2, \ldots, sr^{n-1}\} \quad \text{with} \quad r^n = s^2 = e. \quad (I.17)
\]
This is a finite group of order \( 2n \), called dihedral group of order \( 2n \) and often denoted \( D_n \).

**Remark:** Unfortunately, the dihedral group of order \( 2n \) is equally frequently denoted \( D_{2n} \) in the literature. If the subscript is an odd integer there is no ambiguity; otherwise, let the reader beware!

One generalizes the construction to the cases \( n = 1 \) and \( n = 2 \), for which the geometric representation is meaningless. Thus \( D_1 = \{e, s\} \) with \( s^2 = e \) is actually the same as the cyclic group \( C_2 \) of order 2. In turn, \( D_2 = \{e, r, s, sr\} \) has the same structure as the Klein group \( V_4 \). Both \( D_1 \) and \( D_2 \) are thus Abelian, which is not true of the higher-order dihedral groups \( D_n \) with \( n \geq 3 \).

**Symmetric groups \( S_n \)**

Consider a set \( \mathcal{X} \) with \( n \in \mathbb{N}^* \) elements, for instance \( \mathcal{X} = \{1, 2, \ldots, n\} \). The bijections \( \mathcal{X} \to \mathcal{X} \) are also called *permutations* of \( \mathcal{X} \). Since the composition of two such permutations is again a permutation, one easily checks the following theorem:

**Theorem & Definition I.18.** The permutations on a set of \( n \) elements, together with the composition of functions, form a group \( S_n \), called the symmetric group of degree \( n \).

**Property I.19.** The symmetric group \( S_n \) is of order \( n! \).

A usual notation for a permutation \( \sigma \in S_n \) consists of two lines: the elements of \( \mathcal{X} \) are listed in the first row, and for each element its image under \( \sigma \) below it in the second row:
\[
\sigma = \begin{pmatrix}
1 & 2 & \cdots & n \\
\sigma(1) & \sigma(2) & \cdots & \sigma(n)
\end{pmatrix}. \quad (I.20)
\]
For instance, the identity element of \( S_n \) takes the form
\[
e = \begin{pmatrix}
1 & 2 & \cdots & n \\
1 & 2 & \cdots & n
\end{pmatrix}. \quad (I.21)
\]
The order of the elements of \( \mathcal{X} \) in the first row is irrelevant, so that the permutation \( \sigma \) of Eq. (I.20) may equivalently be written
\[
\sigma = \begin{pmatrix}
i_1 & i_2 & \cdots & i_n \\
\sigma(i_1) & \sigma(i_2) & \cdots & \sigma(i_n)
\end{pmatrix}, \quad (I.22)
\]
where \((i_1, \ldots, i_n)\) is a permutation(!) of \((1, \ldots, n)\). Using this property, the inverse of \( \sigma \) reads
\[
\sigma^{-1} = \begin{pmatrix}
\sigma(1) & \sigma(2) & \cdots & \sigma(n) \\
1 & 2 & \cdots & n
\end{pmatrix} \quad (I.23)
\]
since obviously \( 1 = \sigma^{-1}(\sigma(1)) \), \( 2 = \sigma^{-1}(\sigma(2)) \) and so on. The trick is also useful to construct the “product” of two permutations, i.e. their composition. Writing
\[
\tau = \begin{pmatrix}
1 & 2 & \cdots & n \\
\tau(1) & \tau(2) & \cdots & \tau(n)
\end{pmatrix} = \begin{pmatrix}
\sigma(1) & \sigma(2) & \cdots & \sigma(n) \\
\tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(n))
\end{pmatrix},
\]
one has
\[
\tau \sigma \equiv \tau \circ \sigma = \begin{pmatrix}
\sigma(1) & \sigma(2) & \cdots & \sigma(n) \\
\tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(n))
\end{pmatrix} \begin{pmatrix}
1 & 2 & \cdots & n \\
\sigma(1) & \sigma(2) & \cdots & \sigma(n)
\end{pmatrix} \quad (I.24)
\]
\[
= \begin{pmatrix}
1 & 2 & \cdots & n \\
\tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(n))
\end{pmatrix} = \begin{pmatrix}
\tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(n)) \\
\tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(n))
\end{pmatrix}. \quad (I.25)
\]
In general, \( \tau \sigma \neq \sigma \tau \), i.e. the group \( S_n \) is non Abelian — this is the case for \( n \geq 3 \).
Decomposition of a permutation into independent cycles

Consider an arbitrary permutation \( \sigma \) of \( S_n \), for instance

\[
\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 3 & 9 & 6 & 1 & 7 & 5 & 8 & 2 \end{pmatrix} \in S_9.
\]

One remarks that the three sets of elements \( \{1, 4, 5, 6, 7\} \), \( \{2, 3, 9\} \), and \( \{8\} \) “evolve” independently of each other under the application of \( \sigma \):

- \( 1 \xrightarrow{\sigma} 4 \xrightarrow{\sigma} 6 \xrightarrow{\sigma} 7 \xrightarrow{\sigma} 5 \xrightarrow{\sigma} 1 \);
- \( 2 \xrightarrow{\sigma} 3 \xrightarrow{\sigma} 9 \xrightarrow{\sigma} 2 \);
- \( 8 \xrightarrow{\sigma} 8 \).

These three sets form respective cycles or cyclic permutations, for which one introduces a shorter, one-line notation as e.g.

\[
(1 \ 4 \ 6 \ 7 \ 5) \equiv (1 \ 4 \ 6 \ 7 \ 5 \ 1) \equiv (1 \ 4 \ 6 \ 7 \ 5 \ 1 \ 2 \ 3 \ 8 \ 9) \equiv (1 \ 4 \ 6 \ 7 \ 5 \ 2 \ 3 \ 8 \ 9),
\]

which only affects the elements \( \{1, 4, 5, 6, 7\} \), leaving the other ones unchanged. Using this notation, the permutation of Eq. (I.26) can be written as

\[
\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 3 & 9 & 6 & 1 & 7 & 5 & 8 & 2 \end{pmatrix} = (1 \ 4 \ 6 \ 7 \ 5 \ 2 \ 3 \ 8 \ 9).
\]

As a last simplification, one drops the cycles of length one — here \( (8) \) —, with the convention that non-written elements belong to such cycles \(^{(4)}\)

\[
\sigma = (1 \ 4 \ 6 \ 7 \ 5 \ 2 \ 3 \ 9) = (1 \ 4 \ 6 \ 7 \ 5 \ 2 \ 3 \ 9) \cdot (8).
\]

The permutation \( \sigma \) has now be decomposed into the product of independent (or “disjoint”) cycles. Since the latter affect different elements, their product actually commute:

\[
\sigma = (1 \ 4 \ 6 \ 7 \ 5 \ 2 \ 3 \ 9) = (2 \ 3 \ 9) \cdot (1 \ 4 \ 6 \ 7 \ 5).
\]

The above procedure can be applied to any permutation, which proves the following theorem:

**Theorem I.27.** Every permutation of \( n \) elements can be expressed as the (commutative) product of independent cycles.

**Definition I.28.** A cycle with two elements is called a transposition of those elements.

For any \( k \in \mathbb{N}^* \) with \( k \geq 2 \), and any \( k \)-tuple of numbers \( (i_1, i_2, \ldots, i_k) \), one checks the identity

\[
(i_1 \ i_2 \ i_3 \ \cdots \ i_k) = (i_1 \ i_k) \cdots (i_1 \ i_3)(i_1 \ i_2),
\]

which proves the

**Theorem I.30.** Any cycle can be written as a product of transpositions.

Note that the transpositions on the right-hand side of Eq. (I.29) are not independent of each other.

By combining the theorems (I.27) and (I.30), one finds that any permutation can be written as a product of transpositions. The number of such transpositions is not fixed, yet one can show that for a given permutation \( \sigma \in S_n \), the number of transpositions in any decomposition always has the same parity, i.e. is either always even or always odd. This leads to the consistency of the following definition:

\(^{(4)}\)The notation becomes singular in the case of the identity permutation, which only consists of cycles of length 1!
**Definition I.31.** The parity or signature of a permutation $\sigma \in S_n$ is the parity, characterized by a number $\varepsilon(\sigma) \in \{-1, +1\}$, of the number of transpositions in all decompositions of $\sigma$.

**Example:** Using for instance the decomposition (I.29), one finds that a cycle with $k$ elements has parity $(-1)^{k-1}$.

Let us eventually introduce a last definition:

**Definition I.32.** A permutation that leaves no element unchanged is called a derangement.

The decomposition into disjoint cycles of a derangement contains no cycle of length 1.

---

**Alternate groups $A_n$**

Eventually, one easily checks the following results:

**Theorem & Definition I.33.** The permutations of $n \in \mathbb{N}^*$ elements with even parity form a group, called the alternating group $A_n$.

**Property I.34.** The alternating group $A_n$ is of order $n!/2$.

One checks that only $A_1$, $A_2$ and $A_3$ are Abelian — they consist of only 1, 2 or 3 elements, respectively. For $n \geq 4$, $A_n$ is non Abelian.

---

**I.2.2 Infinite groups**

Let us now give a few examples of infinite groups.

An important one is the set of integers with the addition, $(\mathbb{Z}, +)$. This group is countably infinite, as is $\mathbb{Z}$, and Abelian since $n + m = m + n$ for all $m, n \in \mathbb{Z}$.

Turning now to non-countably infinite groups, one may quote the set of real numbers with the addition, $(\mathbb{R}, +)$, or the set of non-zero real numbers with the usual multiplication $(\mathbb{R}^*, \times)$ — where leaving aside the 0 is necessary to ensure the existence of the inverse for every element of the group. Both these groups are clearly Abelian.

Another example of non-countably infinite, Abelian group is the set of complex numbers with modulus 1 with the multiplication as group law, $(\{e^{i\varphi} | \varphi \in [0, 2\pi]\}, \times)$.

Let now $\mathcal{X}$ be an arbitrary infinite set and $\text{Bij}(\mathcal{X})$ denote the set of bijections of $\mathcal{X}$ on itself. Using the composition of functions as internal law, $(\text{Bij}(\mathcal{X}), \circ)$ is always a group — which generalizes to the infinite case the symmetric groups $S_n$. In general, $(\text{Bij}(\mathcal{X}), \circ)$ is non-Abelian.

Taking as set $\mathcal{X}$ an $n$-dimensional real or complex vector space, say $\mathbb{R}^n$ or $\mathbb{C}^n$, and considering the bijective linear applications on that vector space, the latter can be represented (after choosing a basis) by regular $n \times n$ matrices with real resp. complex entries. The set of such matrices, with the usual matrix product as internal composition law, forms a group $\text{GL}(n, \mathbb{R})$ resp. $\text{GL}(n, \mathbb{C})$.

---

**I.3 Subgroups**

**I.3.1 Definition**

**Definition I.35.** Let $(\mathcal{G}, \cdot)$ be a group. A non-empty set $\mathcal{H} \subseteq \mathcal{G}$ is called a subgroup of $\mathcal{G}$ if it is itself a group under the same composition law $\cdot$ as that of $\mathcal{G}$.

For any group $\mathcal{G}$ with identity element $e$, both $\mathcal{G}$ itself and $\{e\}$ are (trivial) subgroups of $\mathcal{G}$.

**Definition I.36.** A proper subgroup of a group $\mathcal{G}$ is a subgroup $\mathcal{H}$ which differs from both $\mathcal{G}$ and $\{e\}$, i.e. $\{e\} \subsetneq \mathcal{H} \subsetneq \mathcal{G}$. 

Examples:

* The Klein group $V_4$ has three proper subgroups $\{e, a\}$, $\{e, b\}$, and $\{e, c\}$, as can be read off its Cayley table.\[1.15\]

* The set $2\mathbb{Z}$ of even numbers is a proper subgroup of the group $$(\mathbb{Z}, +)$$ of integers, which is itself a subgroup of $$(\mathbb{R}, +)$$. More generally, the set $p\mathbb{Z}$ of the multiples of a number $p \in \mathbb{N}^*$ is a proper subgroup of $$(\mathbb{Z}, +)$$.

**Theorem I.37.** Let $(G, \cdot)$ be a group. A non-empty subset $H \subset G$ is a subgroup of $G$ if and only if for all $g, h \in H$, $g \cdot h^{-1} \in H$. If $G$ is finite, then the condition $g \cdot h \in H$ for all $g, h \in H$ is sufficient.

### I.3.2 Generating set

**Theorem I.38.** Let $(G, \cdot)$ be a group. If $H$ and $K$ are two subgroups of $G$, then their intersection $H \cap K$ is also a subgroup of $G$.

**Theorem & Definition I.39.** Let $S$ be a subset of a group $(G, \cdot)$. The intersection of the subgroups $H$ of $G$ that include $S$ is again a subgroup including $S$, denoted by $\langle S \rangle$.

**Property I.40.** $\langle S \rangle$ is the “smallest” subgroup including $S$, in that it is included in every subgroup $H$ of $G$ including $S$. Note, however, that two arbitrary subgroups $H, K$ of $G$ may not be comparable, i.e. neither $H \subseteq K$ nor $K \subseteq H$ holds.\[5\]

**Definition I.41.** A group $G$ is called *finitely generated* if there exists a finite set $S = \{a_1, \ldots, a_n\}$ of elements of $G$ such that $G = \langle S \rangle$, which is also denoted $G = \langle a_1, \ldots, a_n \rangle$.

Examples:

* Every finite group is obviously finitely generated. For instance, the Klein group is generated by any two of its non-identity elements: $V_4 = \{a, b\} = \{a, c\} = \{b, c\}$.

* The group $(\mathbb{Z}, +)$ or integers is finitely generated: $\mathbb{Z} = \langle 1 \rangle$.

**Definition I.42.** A group is called *cyclic* if it is generated by a single element.

**Property I.43.** A cyclic group is always Abelian.

**Definition I.44.** The *order* of an element $g$ of a group $G$ is the order of the subgroup $\langle g \rangle$ generated by $g$.

**Property I.45.** The order of an element $g$ is either infinite, or it is the smaller positive integer $s$ such that $g^s = e$.

### I.3.3 Cosets of a subgroup

**Definition I.46.** Let $H$ denote a subgroup of a group $(G, \cdot)$ and $g \in G$. The set $gH = \{g \cdot h \mid h \in H\}$ is called left coset of $H$ in $G$ with respect to $g$. Similarly, one defines the right coset $Hg = \{h \cdot g \mid h \in H\}$.

In the remainder of this section, we give results for the left cosets of a subgroup in a group only. The reader can state and prove equivalent results for the right cosets.

\[5\] Stated differently, inclusion is not a total order on the set of subgroups (or more generally subsets) of $G$. 

I.4.1 Equivalence relation

Definition I.55. An equivalence relation \( \sim \) on a set \( S \) is a binary relation on \( S \) which is reflexive, symmetric, and antisymmetric, that is:

- **reflexivity**: for all \( a \in S \), \( a \sim a \)

- **symmetry**: for all \( a, b \in S \), \( a \sim b \) if and only if \( b \sim a \)

- **transitivity**: for all \( a, b, c \in S \), if \( a \sim b \) and \( b \sim c \) then \( a \sim c \)

\[ \text{(E1)} \]
\[ \text{(E2)} \]
\[ \text{(E3)} \]

\(^{(6)}\) A partition of a set \( S \) is a family \( \{A_i\}_{i \in T} \) of subsets \( A_i \subseteq S \) such that the intersection of any two different \( A_i \) is empty (\( A_i \cap A_j = \emptyset \) for \( i \neq j \)), while the union of all \( A_i \) is \( S \), \( \bigcup_j A_i = S \).

\(^{(a)}\) J.-L. Lagrange, 1738–1813

Theorem I.47. The various left cosets of a subgroup \( H \) in a group \( G \) with respect to all elements \( g \in G \) are either identical or disjoint, and they provide a partition \(^{(6)}\) of \( G \).

Proof: let \( g_1, g_2 \in G \) be such that the left cosets \( g_1H \) and \( g_2H \) have (at least) a common element. Then there exist \( h_1, h_2 \in H \) such that \( g_1h_1 = g_2h_2 \). Multiplying this identity left by \( g_2^{-1} \) and right by \( h_1^{-1} \) yields \( g_2^{-1}g_1 = h_2h_1^{-1} \), where the latter product is an element of \( H \). Multiplying both members of this identity right with any \( h \in H \) yields \( g_2^{-1}g_1h = h_2h_1^{-1}h \in H \), i.e. \( g_2^{-1}g_1H \subset H \), which leads at once to \( g_1H \subset g_2H \).

Coming back to \( g_1h_1 = g_2h_2 \) and multiplying now left by \( g_1^{-1} \) and right by \( h_2^{-1} \) amounts to exchanging the roles of \( g_1 \) and \( g_2 \), i.e. leads in turn to \( g_2H \subset g_1H \), so that \( g_1H = g_2H \): both cosets are identical as soon as they have one common element.

Since the left cosets are disjoint, we only have to show that their union \( \bigcup_g gH \), which is necessarily included in \( G \), actually equals \( G \) to prove that they form a partition of the whole group. Now, from \( e \in H \) follows \( g \in gH \) for all \( g \in G \), leading to \( G \subset \bigcup_g gH \).

\[ \square \]

Definition I.48. Let \( H \) denote a subgroup of a group \( G \). The number of distinct left cosets of \( H \) in \( G \) is called the index of \( H \) in \( G \) and denoted \( [G:H] \).

Theorem I.49. Let \( H \) denote a subgroup of a group \( G \). If two of the numbers \( |G|, |H| \) and \( |G:H| \) are finite, then so is the third and they are then related by

\[ |G| = |G:H||H|. \]  \[ \text{(I.50)} \]

Proof: all left cosets have the same order \( |gH| = |H| \) for every \( g \in G \), which leads to the wanted result.

In the case of finite groups, the theorem \(^{(a)}\) leads to several straightforward consequences:

Corollary I.51 (Lagrange’s \(^{(a)}\) theorem). The order \( |H| \) of every subgroup \( H \) of a finite group \( G \) divides the order \( |G| \) of the group.

Corollary I.52. The order \( |\langle g \rangle| \) of every element \( g \) of a finite group \( G \) divides the order \( |G| \) of the group.

Corollary I.53. For every element \( g \) of a finite group \( \langle G, \cdot \rangle \), \( g|G| = e \).

This follows from the property \(^{(a)}\) of the order of an element and from the previous corollary.

Corollary I.54. If \( G \) is a finite group, whose order \( |G| \) is a prime number, then \( G \) is cyclic.

I.4 Conjugacy and conjugacy classes
**Definition I.56.** Let $\sim$ be an equivalence relation on a set $S$. The subset of $S$ consisting of the elements $b$ such that $a \sim b$ is called the equivalence class of $a$ in $S$ with respect to $\sim$.

**Theorem I.57.** Let $\sim$ be an equivalence relation on a set $S$. The various equivalence classes with respect to $\sim$ provide a partition of $S$.

### I.4.2 Conjugacy

**Definition I.58.** Let $(G, \cdot)$ be a group. An element $b \in G$ is said to be conjugate to an element $a \in G$ if there exists $g \in G$ such that $b = g \cdot a \cdot g^{-1}$. $g$ is then called conjugating element.

**Remark:** The conjugating element is not necessarily unique.

**Theorem & Definition I.59.** Let $G$ be a group. The conjugacy relation “element $a$ is conjugate to element $b$ in $G$” is an equivalence relation on $G$, whose equivalence classes are called conjugacy classes.

Accordingly, the conjugacy classes of a group $G$ provide a partition of $G$; the elements of a given conjugacy class share many properties, as will now be illustrated on a few examples.

**Remark I.60.** One sees at once that the identity element $e$ of a group will always form a conjugacy class by itself.

### I.4.3 Examples of conjugacy classes

#### I.4.3a Abelian groups

Coming back to definition (I.58), one sees that every element $a$ of an Abelian group $G$ is only conjugate to itself, since for all $g \in G$, $g \cdot a \cdot g^{-1} = a$. That is, the conjugacy classes of an Abelian group are singletons, i.e. sets with a single element.

In particular, since the cyclic group $C_n$ is Abelian for any $n \in \mathbb{N}^*$, each of its element forms a conjugacy class by itself.

#### I.4.3b Symmetric group $S_n$

Let $\sigma \in S_n$ be a permutation of $n$ objects. The conjugacy class of $\sigma$ consists of the permutations of $S_n$ which have the same cycle structure as $\sigma$.

To prove this result, let us first write the decomposition of $\sigma$ into $r$ disjoint cycles of respective sizes $\ell_1, \ldots, \ell_r$ with $1 \leq \ell \leq n$, including the cycles of length 1:

$$\sigma = (i_1 \ i_2 \ \ldots \ i_{\ell_1})(j_1 \ \ldots \ j_{\ell_2})(k_1 \ \ldots \ k_{\ell_3}) \ldots$$  \hspace{1cm} (I.61)

with $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_r \geq 1$ and $\ell_1 + \ell_2 + \cdots + \ell_r = n$. In this decomposition, the $n$ numbers $i_1, i_2, \ldots, j_{\ell_2}, \ldots \in \{1, \ldots, n\}$ are all different from each other.

Any permutation $\tau \in S_n$ can be written as

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix} = \begin{pmatrix} i_1 & i_2 & \cdots & i_{\ell_1} & j_1 & \cdots & j_{\ell_2} & \cdots \\ \tau(i_1) & \tau(i_2) & \cdots & \tau(i_{\ell_1}) & \tau(j_1) & \cdots & \tau(j_{\ell_2}) & \cdots \end{pmatrix}$$

with the inverse permutation $\tau^{-1}$ [cf. Eq. (I.23)]

$$\tau^{-1} = \begin{pmatrix} i_1 & i_2 & \cdots & i_{\ell_1} & j_1 & \cdots & j_{\ell_2} & \cdots \\ \tau(i_1) & \tau(i_2) & \cdots & \tau(i_{\ell_1}) & \tau(j_1) & \cdots & \tau(j_{\ell_2}) & \cdots \end{pmatrix}.$$

A straightforward calculation then gives

$$\tau \sigma \tau^{-1} = (\tau(i_1) \ \tau(i_2) \ \ldots \ \tau(i_{\ell_1})) (\tau(j_1) \ \ldots \ \tau(j_{\ell_2})) (\tau(k_1) \ \ldots \ \tau(k_{\ell_3})) \ldots,$$  \hspace{1cm} (I.62)
Conjugacy and conjugacy classes

i.e. a permutation with a decomposition into \( r \) cycles with respective sizes \( \ell_1, \ldots, \ell_r \), similar to that of \( \sigma \).

Conversely, one easily checks that two permutations of \( S_n \) having the same cycle structure are conjugate to each other — Equations (I.61) and (I.62) show how one can find a conjugating element relating them.

All in all, the various conjugacy classes of the symmetric group \( S_n \) are in one-to-one correspondence with the different integer partitions\(^{(7)}\) of \( n \), i.e. of the possible ways of writing \( n \) as a sum of positive integers (up to the order of the summands), namely as

\[
 n = \ell_1 + \ell_2 + \cdots + \ell_r \quad \text{with} \quad \ell_1 \geq \ell_2 \geq \cdots \geq \ell_r \geq 1 \quad \text{and} \quad \ell_1, \ell_2, \ldots, \ell_r \in \mathbb{N}^*. \quad (I.63)
\]

This will now be illustrated with the cases of \( S_2 \), \( S_3 \) and \( S_4 \).

The two permutations of \( S_2 \), namely the identity — with cycle decomposition \((1)(2)\) — and the transposition of the two elements \((1 \ 2)\), have different cycle structures and thus belong to different conjugacy classes. This is actually consistent with the fact that \( S_2 \) is Abelian, so that each of its elements builds a conjugacy class by itself.

The corresponding partitions of 2 are 2 = 1 + 1 and 2 = 2.

The six permutations of \( S_3 \) are distributed into three conjugacy classes. A first class consists of the identity permutation \((1)(2)(3)\), in agreement with remark \( I.60 \). Then come the three permutations with a single transposition of two elements, namely \((1 \ 2)(3)\), \((1 \ 3)(2)\), and \((2 \ 3)(1)\). Eventually there remain the two circular permutations \((1 \ 2 \ 3)\) and \((1 \ 3 \ 2)\).

The corresponding partitions of 3 are 3 = 1 + 1 + 1 and 3 = 2 + 1, and 3 = 3, respectively.

Coming now to \( S_4 \), one finds five conjugacy classes, which respectively correspond to the partitions \( 4 = 1 + 1 + 1 + 1, 4 = 2 + 1 + 1, 4 = 2 + 2, 4 = 3 + 1, \) and \( 4 = 4 \). First, the identity permutation \((1)(2)(3)(4)\) forms its own class. There are then 6 permutations consisting of a single transposition: \((1 \ 2)(3)(4), (1 \ 3)(2)(4), (1 \ 4)(2)(3), (2 \ 3)(1)(4), (2 \ 4)(1)(3), (3 \ 4)(1)(2)\). Three permutations can be decomposed into two disjoint transpositions, namely \((1 \ 2)(3 \ 4), (1 \ 3)(2 \ 4), \) and \((1 \ 4)(2 \ 3)\). With a cycle of length 3, one finds the 8 permutations \((1 \ 2 \ 3)(4), (1 \ 3 \ 2)(4), (1 \ 2 \ 4)(3), (1 \ 4 \ 2)(3), (1 \ 3 \ 4)(2), (1 \ 4 \ 3)(2), (2 \ 3 \ 4)(1), \) and \((2 \ 4 \ 3)(1)\). Eventually, there are 6 distinct circular permutations of all 4 elements: \((1 \ 2 \ 3 \ 4), (1 \ 2 \ 4 \ 3), (1 \ 3 \ 2 \ 4), (1 \ 3 \ 4 \ 2), (1 \ 4 \ 2 \ 3), (1 \ 4 \ 3 \ 2)\).

**Young diagrams**

A convenient way to depict a partition of the integer \( n \) consists of a Young\(^{(e)}\) diagram, in which \( n \) boxes (\( \square \)) are arranged in \( r \) rows, with \( \ell_s \) boxes on the \( s \)-th row.

Thus, the two partitions of 2 are respectively represented by

\[
\begin{array}{c}
\,
\end{array}
\quad \text{and} \quad
\begin{array}{c}
\,
\end{array}.
\quad (I.64)
\]

the three partitions of 3 by

\[
\begin{array}{c}
\,
\end{array},
\begin{array}{c}
\,
\end{array}, \quad \text{and} \quad
\begin{array}{c}
\,
\end{array},
\quad (I.65)
\]

and the five partitions of 4 by

\[
\begin{array}{c}
\,
\end{array},
\begin{array}{c}
\,
\end{array},
\begin{array}{c}
\,
\end{array},
\begin{array}{c}
\,
\end{array}, \quad \text{and} \quad
\begin{array}{c}
\,
\end{array}.
\quad (I.66)
\]

\(^{(7)}\)Note that a partition of an integer \( n \) is not the same notion as the partition of a set as defined in footnote (6).

\(^{(e)}\)A. Young, 1873–1940
Remark: The number of permutations of \( n \) elements in the conjugacy class corresponding to a decomposition with \( r_1 \) cycles of length 1, \( r_2 \) cycles of length 2, \ldots, \( r_n \) cycles of length \( n \) — where the \( r_i \) can now take their values in \( \{0, \ldots, n\} \) — is given by

\[
n! \quad \frac{r_1! \cdot r_2! \cdots r_n!}{1! \cdot 2! \cdots n!},
\]

with the usual convention \( 0! = 1 \). Setting \( r_1 = n \) and all other \( r_i = 0 \), one again finds that the identity permutation is alone in its class. In turn, with \( r_n = 1 \) and all other \( r_i = 0 \), one sees that there are \( (n-1)! \) different circular permutations of all \( n \) elements — as was indeed found for \( n = 2, 3, 4 \).

I.4.3c Three-dimensional rotations

Considering now rotations about axes going through a fixed point \( O \) of three-dimensional space, let us denote by \( R_{\vec{n}}(\alpha) \) the rotation through an angle \( \alpha \) about the axis with direction \( \vec{n} \).

If \( R' \) is any arbitrary rotation about an axis going through \( O \), one checks that the composition of transformations \( R' R_{\vec{n}}(\alpha) R'^{-1} \) is again a rotation through \( \alpha \), now about the axis with direction \( \vec{n}' = R' \vec{n} \). Conversely, one finds that two rotations through the same angle \( \alpha \) are conjugate to each other.

I.4.3d Invertible matrices

In the group of regular real-valued resp. complex-valued \( n \times n \) matrices \( GL(n, \mathbb{R}) \) resp. \( GL(n, \mathbb{C}) \), the conjugacy relation is the so-called matrix similarity: two similar matrices actually represent the same linear application in two different bases.

I.5 Normal subgroups

In the previous two sections, we have introduced two different partitions of a group \( \mathcal{G} \): first, by considering a subgroup \( \mathcal{H} \) of \( \mathcal{G} \) and the collection of its left cosets \( g \mathcal{H} \) (Sec. I.3.3); then, by looking at the conjugacy classes of \( \mathcal{G} \) (Sec. I.4). Since the conjugacy class of the identity is always a singleton, which is not the case of the other classes if the group is non-Abelian, while all cosets of a subgroup have the same number of elements, the conjugacy classes are in general not the cosets of a subgroup — the exception being the case of the conjugacy classes of an Abelian group, which are the cosets of the trivial subgroup \( \{e\} \).

Nevertheless, one can identify specific subgroups with a well-defined behavior under conjugacy, whose cosets can be provided with a group structure.

I.5.1 Definition

Definition I.67. A subgroup \( \mathcal{N} \) of a group \( \mathcal{G} \) is called normal (or invariant) in \( \mathcal{G} \) if the conjugate of every element of \( \mathcal{N} \) by any element \( g \) of \( \mathcal{G} \) is still in \( \mathcal{N} \), i.e. \( g \mathcal{N} g^{-1} \subseteq \mathcal{N} \) or, in less compact notation, for all \( n_1 \in \mathcal{N} \) and for all \( g \in \mathcal{G} \), there exists \( n_2 \in \mathcal{N} \) such that \( n_2 = gn_1g^{-1} \).

Remarks:
* One actually even has the equality \( g \mathcal{N} g^{-1} = \mathcal{N} \) for all \( g \in \mathcal{G} \).
* The trivial subgroups \( \{e\} \) and \( \mathcal{G} \) of a group \( \mathcal{G} \), where \( e \) is the identity element of \( \mathcal{G} \), are clearly always normal in \( \mathcal{G} \).

Notation: If \( \mathcal{N} \) is a proper normal subgroup of \( \mathcal{G} \), one writes \( \mathcal{N} \triangleleft \mathcal{G} \).

Property I.68. If \( \mathcal{G} \) is an Abelian group, then all its subgroups are normal.
I.5.2 Quotient group

From definition [I.67] follows at once the following result.

**Property I.69.** If \( \mathcal{N} \) is a normal subgroup of a group \( \mathcal{G} \), then the left cosets of \( \mathcal{N} \) in \( \mathcal{G} \) equal the right cosets, i.e. \( g\mathcal{N} = \mathcal{N}g \) for all \( g \in \mathcal{G} \).

This property is actually the key to the proper definition an internal composition law \( \cdot \) on the set \( \{g\mathcal{N} \mid g \in \mathcal{G}\} \) of the (left) cosets of \( \mathcal{N} \) in \( \mathcal{G} \).

**Theorem & Definition I.70.** Let \( \mathcal{N} \) be a normal subgroup of a group \( \mathcal{G} \). The set \( \mathcal{G}/\mathcal{N} = \{g\mathcal{N} \mid g \in \mathcal{G}\} \) of the cosets of \( \mathcal{N} \) in \( \mathcal{G} \), together with the internal composition law \( \cdot \) defined by

\[
(g\mathcal{N}) \cdot (g'\mathcal{N}) = (gg')\mathcal{N},
\]

forms a group, called the quotient group (or factor group) of \( \mathcal{G} \) by \( \mathcal{N} \).

**Property I.72.** The identity element of \( \mathcal{G}/\mathcal{N} \) is \( e\mathcal{N} = \mathcal{N} \) and the inverse element of \( g\mathcal{N} \) in \( \mathcal{G}/\mathcal{N} \) is \( g^{-1}\mathcal{N} \).

To prove the theorem [I.70] one must first show that Eq. (I.71) indeed defines a composition law on the set of cosets. This already requires two logical steps.

Assume first that \( g, g' \in \mathcal{G} \) are given. Then for all \( n_1, n_2 \in \mathcal{N} \), \( gn_1 \) resp. \( g'n_2 \) is in the left coset \( g\mathcal{N} \) resp. \( g'\mathcal{N} \), and the associativity of the group law on \( \mathcal{G} \) gives \( (gn_1)(g'n_2) = g(n_1g'n_2) \). In the latter right member, \( n_1g' \) is an element of the right coset \( \mathcal{N}g' \), which according to property I.69 equals the left coset \( g'\mathcal{N} \); that is, there exists \( n_3 \in \mathcal{N} \) such that \( n_1g' = g'n_3 \). One may thus write \( (gn_1)(g'n_2) = g(n_1g'n_2) = g(g'n_3)n_2 = (g'g)n_2 \), where associativity in \( \mathcal{G} \) was again used: \( n_3n_2 \) is an element of \( \mathcal{N} \), so that \( (g'g)(n_3n_2) \in (g'g)\mathcal{N} \): for all \( n_1, n_2 \in \mathcal{N} \), \( (gn_1)(g'n_2) \in (g'g)\mathcal{N} \), so that Eq. (I.71) makes sense.

To conclude on the consistency of Eq. (I.71) as a well-defined internal law, one still needs to check that the product is independent of the group element \( g \) chosen to label a given coset \( g\mathcal{N} \). That is, if \( g_1\mathcal{N} = g_2\mathcal{N} \) and \( g'_1\mathcal{N} = g'_2\mathcal{N} \) with \( g_2 \neq g_1 \), \( g'_2 \neq g'_1 \), then \( (g_1\mathcal{N}) \cdot (g'_1\mathcal{N}) = (g_2\mathcal{N}) \cdot (g'_2\mathcal{N}) \).

**Remark I.73.** The order of the quotient group of \( \mathcal{G} \) by \( \mathcal{N} \) equals the index of \( \mathcal{N} \) in \( \mathcal{G} \):

\[
|\mathcal{G}/\mathcal{N}| = |\mathcal{G} : \mathcal{N}|.
\]

I.5.3 Examples

As a first example, consider the group \( \mathcal{G} = \{0, 1, 2, 3, 4, 5\} \) with the addition modulo 6 as group law — which is in fact a representation of the cyclic group \( \mathbb{C}_6 \). Since the group is Abelian, all its subgroups are automatically normal (property I.68), as for instance \( \mathcal{N} = \{0, 3\} \). The corresponding cosets are the sets \( g + \mathcal{N} \) with \( g \in \mathcal{G} \), namely \( \mathcal{N} = 3 + \mathcal{N} \), \( 1 + \mathcal{N} = 4 + \mathcal{N} = \{1, 4\} \), and \( 2 + \mathcal{N} = 5 + \mathcal{N} = \{2, 5\} \). The quotient group of \( \mathcal{G} \) by \( \mathcal{N} \) is thus

\[
\mathcal{G}/\mathcal{N} = \{\{0, 3\}, \{1, 4\}, \{2, 5\}\}
\]

and examples of group additions on \( \mathcal{G}/\mathcal{N} \) are \( \{0, 3\} + \{1, 4\} = \{1, 4\} \) (since \( \{0, 3\} = \mathcal{N} \) is the identity element) or \( \{1, 4\} + \{2, 5\} = \{0, 3\} \), meaning that \( \{1, 4\} \) is the inverse of \( \{2, 5\} \).

Next, one shows using the decomposition of permutation into transpositions that the alternating group \( \mathbb{A}_n \) is normal in the symmetric group \( \mathbb{S}_n \) — for every \( \sigma \in \mathbb{S}_n \), \( \sigma \) and its inverse \( \sigma^{-1} \) have the same parity.

Taking \( \mathcal{G} = \mathbb{Z} \), with the addition of integer numbers as group law, and \( \mathcal{N} = 2\mathbb{Z} \), i.e. the set of even numbers, one finds that there are two cosets \( 2\mathbb{Z} \) itself and the set \( 1 + 2\mathbb{Z} \) of odd numbers.
The quotient group \( G/N = \mathbb{Z}/2\mathbb{Z} \) thus contains two elements only — it may thus be identified \( \mathbb{Z}_2 = \{0, 1\} \).

**I.5.4 Further definitions and results**

**Definition I.74.** A group \( G \) is called a *simple group* if it contains no proper normal subgroup.

**Remark:** The classification of the finite simple groups has been achieved in 2008. For example, all cyclic groups \( C_p \) with a prime order \( p \) are simple, since they include no proper subgroup, as are the alternating groups \( A_n \) with \( n \geq 5 \) (as well as those with \( n \leq 3 \), which are Abelian).

**Composition series**

**Definition I.75.** A *composition series* of a group \( G \) is a finite series of the type

\[
\{e\} = \mathcal{H}_0 < \mathcal{H}_1 < \cdots < \mathcal{H}_r = G
\]

such that each quotient group \( \mathcal{H}_{i+1}/\mathcal{H}_i \) with \( i \in \{0, \ldots, r-1\} \) is simple. The quotient groups \( \mathcal{H}_{i+1}/\mathcal{H}_i \) are called *composition factors*.

**Remark:** A series of the type \( \{e\} \), and more generally a similar sequence of subgroups of infinite length, is called a *subnormal series*.

**Example:** The symmetric group \( S_3 \) admits the alternating group \( A_3 \) as proper normal subgroup, and \( A_3 \) itself is simple:

\[
\{e\} < A_3 < S_3,
\]

where \( e \) denotes the identity permutation of three elements. Now, \( A_3/\{e\} \) is (isomorphic\(^8\) to) \( A_3 \) itself, and is thus simple. In turn, \( S_3/A_3 \) is of order 2, i.e. will be (isomorphic to) \( C_2 \) and thus again simple.

More generally, every finite group has a composition series. However, this is not necessarily true for infinite groups, as e.g. for \( (\mathbb{Z}, +) \): for any proper (normal\(^9\)) subgroup \( \mathcal{N} \) of \( \mathbb{Z} \), one easily finds a proper normal subgroup \( \mathcal{N}' \) of \( \mathcal{N} \), for instance \( 2\mathcal{N} \), such that \( \mathcal{N}'/\mathcal{N} \) is simple, so that one cannot find a subnormal series of finite length.

If a group has a composition series, then it is basically unique:

**Theorem I.77** (Jordan\(^f\)–Hölder\(^g\) theorem). *The composition factors \( \mathcal{H}_{i+1}/\mathcal{H}_i \) of two composition series \( \{e\} \) of a given group \( G \) are the same, up to a permutation of the factors.*

**I.6 Direct product**

**I.6.1 Cartesian product of groups**

**Theorem & Definition I.78.** Given two groups \( G_1 \) and \( G_2 \), their *direct Cartesian product* \( G_1 \times G_2 \) is the set of ordered pairs \( (g_1, g_2) \) with \( g_1 \in G_1, g_2 \in G_2 \):

\[
G_1 \times G_2 = \{(g_1, g_2) \mid g_1 \in G_1, g_2 \in G_2\}.
\]

(1.78a)

The internal composition law \( \cdot \) defined by

\[
(g_1, g_2) \cdot (g'_1, g'_2) = (g_1 g'_1, g_2 g'_2)
\]

for all \( g_1, g'_1 \in G_1, g_2, g'_2 \in G_2 \),

(1.78b)

provides \( G_1 \times G_2 \) with a group structure.

\(^8\)The notion of being “isomorphic” will be defined in Sec. I.7.1

\(^9\)(\(\mathbb{Z}, +\)) is Abelian, so that all its subgroups are normal.

\(^f\)C. Jordan, 1838–1922 \(^g\)O. Hölder, 1859–1937
I.6.2 Inner direct product

Remark I.79. The identity element of \( G_1 \times G_2 \) is \((e_1, e_2)\), where \(e_1\) and \(e_2\) are the respective identity elements of \(G_1\) and \(G_2\). In turn, the inverse of \((g_1, g_2) \in G_1 \times G_2\) is \((g_1^{-1}, g_2^{-1})\).

Property I.80. The subsets \( \tilde{G}_1 = G_1 \times \{ e_2 \} \) and \( \tilde{G}_2 = \{ e_1 \} \times \tilde{G}_2 \), where \(e_1\) and \(e_2\) are the respective identity elements of \(G_1\) and \(G_2\), obey the following properties:

- \( \tilde{G}_1 \) and \( \tilde{G}_2 \) are normal subgroups of the direct product \( G_1 \times G_2 \);
- the intersection of \( \tilde{G}_1 \) and \( \tilde{G}_2 \) contains a single element, namely the identity element of the direct product, \( \tilde{G}_1 \cap \tilde{G}_2 = \{ e \} \);
- every element of \( \tilde{G}_1 \) commutes with every element of \( \tilde{G}_2 \).

Remark: Note that the latter property holds even if the groups \( \tilde{G}_1 \) and \( \tilde{G}_2 \) are not Abelian.

Given \( s \geq 2 \) groups \( G_1, G_2, \ldots, G_s \), their direct Cartesian product \( G_1 \times G_2 \times \cdots \times G_s \) follows from a straightforward generalization of definition I.78. The order of this direct product is then

\[
|G_1 \times G_2 \times \cdots \times G_s| = \prod_{i=1}^{s} |G_i|.
\]

One easily generalizes the property I.80 to the Cartesian direct product of more than two groups.

I.6.2 Inner direct product

Definition I.82. A group \( G \) is called (inner) direct product group of normal subgroups \( N_1, \ldots, N_s \) if the following properties hold:

- \( G = N_1 \cdots N_s \); \hfill (I.82a)
- \( N_i \cap (N_1 \cdots N_{i-1}N_{i+1} \cdots N_s) = \{ e \} \) for all \( i \in \{1, \ldots, s\} \), \hfill (I.82b)

with \( e \) the identity element of \( G \).

Equation (I.82b) means that every element \( g \in G \) can be written as

\[
g = n_1n_2\cdots n_s
\]

with \( n_i \in N_i \) for all \( i \in \{1, \ldots, s\} \).

Property I.84. Up to a permutation of the factors \( n_i \) (see property I.85), the decomposition (I.83) is unique.

Property I.85. The elements of \( N_i \) commute with those of \( N_j \) for \( i \neq j \).

Example I.86. Let \( m \) and \( n \) be coprime integers. The group \((\mathbb{Z}_{mn}, +)\) contains two proper subgroups consisting of the multiples modulo \( mn \) of \( m \) and of those of \( n \): \( N_1 = \{0, m, 2m, \ldots, (n-1)m\} \) and \( N_2 = \{0, n, 2n, \ldots, (m-1)n\} \). As \( \mathbb{Z}_{mn} \) is Abelian, both \( N_1 \) and \( N_2 \) are normal (property I.68). Since \( m \) and \( n \) are coprime, each number \( k \in \mathbb{Z}_{mn} \) can be written as the sum modulo \( mn \) of an element of \( N_1 \) and an element of \( N_2 \) \[\text{(10)}\] The intersection of these subgroups reduces to the singleton \( \{0\} \). All in all, \( \mathbb{Z}_{mn} \) is thus the direct product of \( N_1 \) and \( N_2 \), which are respectively isomorphic \[\text{(8)}\] to \( \mathbb{Z}_m \) and \( \mathbb{Z}_n \).

Property I.87. If \( G \) is the direct product of normal subgroups \( N_1 \) and \( N_2 \), \( G = N_1 \times N_2 \), then \( N_1 \) is isomorphic \[\text{(8)}\] to the quotient group \( G/N_2 \).

The converse does not hold! For instance, \[\text{(11)}\] the dihedral group \( D_3 \) contains a normal subgroup, consisting of the rotations through \( 2k\pi/3 \) with \( k = 0, 1, 2 \), which is isomorphic to \( C_3 \). The quotient group \( D_3/C_3 \), consisting of two elements, is isomorphic to \( C_2 \). Yet the direct product group \( C_2 \times C_3 \), which is Abelian, cannot be isomorphic to the non-Abelian \( D_3 \).

---

\[\text{(10)}\] The proof follows from Bézout’s \[\text{(6)}\] lemma applied to the integers \( mk \) and \( nk \), whose greatest common divisor is \( k \).

\[\text{(11)}\] If she wishes, the reader may replace \( D_3 \) by the symmetric group \( S_3 \) and \( C_3 \) by the alternating group \( A_3 \) in this example.

\[\text{(b)}\] E. Bézout, 1730–1783
I.7 Group homomorphisms

I.7.1 Definitions

Definition I.88. Let $(G, \cdot)$ and $(G', \ast)$ be two groups. A mapping $f : G \to G'$ is called a (group) homomorphism if it is compatible with the group structure, i.e. if

$$f(g_1 \cdot g_2) = f(g_1) \ast f(g_2) \text{ for all } g_1, g_2 \in G.$$  \hfill (I.89)

Definition I.90. A bijective homomorphism is called an isomorphism.

Definition I.91. A homomorphism from a given group $G$ into itself is called an endomorphism of $G$.

Definition I.92. A bijective endomorphism of $G$ is called an automorphism of $G$.

Theorem & Definition I.93. If there exists an isomorphism $f : G \to G'$, then the groups $G$ and $G'$ are said to be isomorphic. This defines an equivalence relation between groups, denoted $G \cong G'$.

Examples:

* Considering on the one hand the group $(\mathbb{Z}_n, +)$ and on the other hand the group of the $n$-th roots of the unity $(\{e^{2ik\pi/n}, k \in \{0, 1, \ldots, n-1\}\}, \times)$, there is the obvious isomorphism $k \mapsto e^{2ik\pi/n}$.

* For any $a \in \mathbb{R}$, the linear mapping $x \mapsto ax$ is an endomorphism of $(\mathbb{R}, +)$. If $a \neq 0$, it is an isomorphism.

* The well-known property $\ln(xy) = \ln x + \ln y$ expresses that the logarithm $\ln$ is a homomorphism from the group $(\mathbb{R}^*_+, \times)$ of positive real numbers (with the multiplication as internal law) into the group $(\mathbb{R}, +)$ of real numbers with the addition as group law.

Conversely, the exponential is a homomorphism from $(\mathbb{R}, +)$ into $(\mathbb{R}^*_+, \times)$, since $e^{x+y} = e^x e^y$.

* In turn, the identity $\det(AB) = (\det A)(\det B)$ means that the determinant is a homomorphism from the group $\text{GL}(n, \mathbb{C})$ of regular $n \times n$ matrices with complex entries on the group $(\mathbb{C}\setminus\{0\}, \times)$ of non-zero complex numbers.

* The signature $\varepsilon$ (definition I.31) is a homomorphism from the symmetric group $S_n$ (with $n \in \mathbb{N}^*$) into the group $(\{1, -1\}, \times)$.

Theorem I.94 (Cayley theorem). Every (finite) group of order $n \in \mathbb{N}^*$ is isomorphic to a subgroup of the symmetric group $S_n$.

To prove the theorem, one only needs to realize that each line of the Cayley table of the group is precisely a permutation of the group elements.

I.7.2 Properties of homomorphisms

Let $G$ and $G'$ be two groups with respective identity elements $e$ and $e'$, and let $f : G \to G'$ be a homomorphism. One easily checks that the following results hold for the examples of group homomorphisms given above.

Theorem I.95. The homomorphism $f$ maps the identity element $e \in G$ to the identity element $e' \in G'$:

$$f(e) = e'.$$ \hfill (I.95)

Theorem I.96. For every element $g \in G$, the homomorphism $f$ maps the inverse of $g$ to the inverse of the image $f(g)$:

$$f(g^{-1}) = (f(g))^{-1}.$$ \hfill (I.96)
Theorem & Definition I.97. Denoting $\text{im } f$ the image by the homomorphism $f$ of the group $G$, 

$$\text{im } f = \{ f(g) \mid g \in G \},$$  

(I.97)

$\text{im } f$ is a subgroup of $G'$. 

For instance, one may consider that the homomorphism $k \mapsto e^{2i\pi/n}$ maps $\mathbb{Z}_n$ into a subgroup of the group $C\setminus\{0\}$. 

Theorem & Definition I.98. The kernel $\ker f$ of the homomorphism $f$, i.e. the set of elements $g \in G$ which are mapped to the identity element $e'$, 

$$\ker f = \{ g \in G \mid f(g) = e' \},$$  

(I.98)

is a normal subgroup of $G$. 

Example: The kernel of the signature $\varepsilon : S_n \to \{1, -1\}$ is the alternating group $A_n$. 

Remark: Conversely, every normal subgroup of a group $G$ can be used as kernel of a homomorphism, as exemplified in the proof of the isomorphism theorem I.101. 

Theorem I.99. A homomorphism $f$ is injective if and only if $\ker f = \{e\}$. 

Theorem I.100. More generally, the image by the homomorphism $f$ of a subgroup $H \subset G$ is a subgroup of $G'$, while the “preimage” by $f$ of a subgroup $H' \subset G'$ is a subgroup of $G$. 

Since $\ker f$ is a normal subgroup of $G$, one can define a quotient group $G/\ker f$. 

Theorem I.101 (Isomorphism theorem). If $f : G \to G'$ is a group homomorphism, then its image is isomorphic to the quotient group $G/\ker f$: 

$$\text{im } f \cong G/\ker f.$$  

(I.101)

To prove this theorem, one has to check that the mapping $F : G/K \to \text{im } f$, which maps the coset $gK$ to $F(gK) = f(g)$ with $g \in G$ is an isomorphism, where for the sake of brevity we introduce the shorter notation $K = \ker f$. 

(more details later) 

The results listed in this section strongly constrain the form of homomorphisms, as we now illustrate on a simple example. Consider the homomorphisms from the symmetric group $S_3$ (or equivalently the dihedral group $D_3$ since they are isomorphic) into another group $G$. Theorem I.98 already restricts the kernel of such a homomorphism $f$, which has to be a normal subgroup of $S_3$, leaving only three possibilities: $\ker f = \{e\}$ (the identity permutation), the alternating group $A_3$, or $S_3$ itself. 

If $\ker f = \{e\}$, then $f$ is injective, so that its image is isomorphic to $S_3$, i.e. $f$ is (equivalent to) an isomorphism of $S_3$. In turn, if $\ker f = S_3$, then $f$ is a constant mapping since its image reduces to the identity element of the target group $G$. 

Eventually, if $\ker f = A_3$, the only proper normal subgroup of $S_3$, then the isomorphism theorem I.101 tells us that the image of $f$ is isomorphic to the quotient group $S_3/A_3$, which consists of two elements only. More precisely, $f$ is necessarily (equivalent to) the signature $\varepsilon$. 

I.7.3 Inner automorphisms 

Definition I.102. Let $G$ be a group and $a \in G$ one of its elements. The automorphism $\phi_a$ defined by 

$$\phi_a : \{ G \to G \\ g \mapsto \phi_a(g) = aga^{-1} \}$$  

(I.103)

is called an inner automorphism of $G$. 

**Theorem I.104.** The set of inner automorphisms of $\mathcal{G}$ is a subgroup $\text{Inn}(\mathcal{G})$ of the group of automorphisms of $\mathcal{G}$.

**Proof:** For all $a \in \mathcal{G}$ one has $(\phi_a)^{-1} = \phi_{a^{-1}}$ and thus for all $a, b, g \in \mathcal{G}$

$$\phi_b \circ (\phi_a)^{-1}(g) = b(a^{-1}ga)b^{-1} = ba^{-1}gab^{-1} = (ba^{-1})g(ba^{-1})^{-1} = \phi_{ba^{-1}}(g),$$

i.e. $\phi_b \circ (\phi_a)^{-1} = \phi_{ba^{-1}} \in \text{Inn}(\mathcal{G})$.

**Remark:** Coming back to definition I.67, one sees that a normal subgroup of a group $\mathcal{G}$ is a subgroup which is invariant under $\text{Inn}(\mathcal{G})$, which justifies the alternative denomination “invariant subgroup”.
CHAPTER II
Representation theory

An intro is missing

II.1 Group action

**Definition II.1.** An action of a group \( G \) on a set \( X \) is a correspondence that associates to each element \( g \in G \) a mapping \( \phi_g : X \rightarrow X \) in such a way that

- for all \( g_1, g_2 \in G \), \( \phi_{g_1 g_2} = \phi_{g_1} \circ \phi_{g_2} \) \hspace{1cm} (II.1a)
- \( \phi_e \) is the identity mapping on \( X \). \hspace{1cm} (II.1b)

Instead of \( \phi_g(x) \) with \( x \in X \), one often writes \( gx \).

**Theorem & Definition II.2.** The action of a group \( G \) on a set \( X \) induces an equivalence relation on the latter: two elements \( x_1, x_2 \in X \) are equivalent if there exists \( g \in G \) such that \( x_2 = gx_1 \). The equivalence class of an element \( x \in X \) is called the orbit of \( x \) under the action of \( G \).

II.2 Group representation: first definitions and results

**Definition II.3.** Let \( V \) be a vector space over a base field \( K \). A representation of a group \( G \) on \( V \) is a homomorphism \( \hat{\rho} \) of \( G \) into the group of the regular operators of \( V \) to itself.

Coming back to the definition I.88 of a homomorphism, the operators \( \hat{\rho}(g) \) with \( g \in G \) obey the following properties:

- for all \( g_1, g_2 \in G \), \( \hat{\rho}(g_1 g_2) = \hat{\rho}(g_1) \hat{\rho}(g_2) \), \hspace{1cm} (II.4a)

where the composition of operators is denoted as a product;

- \( \hat{\rho}(e) = \hat{1}_V \), \hspace{1cm} (II.4b)

where \( e \) is the identity element of \( G \) and \( \hat{1}_V \) the identity operator on \( V \); and

- for all \( g \in G \), \( \hat{\rho}(g^{-1}) = \hat{\rho}(g)^{-1} \), \hspace{1cm} (II.4c)

which explains why the operators \( \hat{\rho}(g) \) have to be regular.

**Remark:** Comparing with definition II.1, one sees that a group representation is an instance of group action on the vector space \( V \).

**Definition II.5.** The vector space \( V \) is the representation space and its dimension \( \dim V \) is called the dimension or degree if the representation.

Accordingly, one talks of an \( s \)-dimensional representation, often abbreviated “\( s \)-rep”.

**Definition II.6.** A representation is said to be real resp. complex when the base field of the vector space is \( K = \mathbb{R} \) resp. \( K = \mathbb{C} \).
Definition II.7. If the mapping $G \rightarrow \hat{G}(G)$ is an isomorphism, then the representation $\hat{G}$ is called faithful.

Definition II.8. A linear representation is a representation whose operators $\hat{G}(g)$ with $g \in G$ are all linear, i.e. $\hat{G}(g) \in \text{GL}(\mathcal{V})$ for all $g \in G$, where $\text{GL}(\mathcal{V})$ precisely denotes the ("General Linear") group of regular linear operators on $\mathcal{V}$.

Definition II.9. A representation is called unitary resp. orthogonal if

- $\mathcal{V}$ is a Hilbert space, i.e. in particular a vector space with a scalar product;
- the operators $\hat{G}(g)$ with $g \in G$ belong to the group $\text{U}(\mathcal{V})$ of unitary operators resp. to the group $\text{O}(\mathcal{V})$ of orthogonal operators on $\mathcal{V}$, i.e. $\text{im} \hat{G} \subset \text{U}(\mathcal{V})$ resp. $\text{im} \hat{G} \subset \text{O}(\mathcal{V})$.

In the case of a finite-dimensional linear representation, one can choose a basis on $\mathcal{V}$ to represent the operators $\hat{G}(g)$ as non-singular square matrices — which we shall denote $\hat{G}(g) —$, resulting in a matrix representation of the group. Denoting $\text{dim} \mathcal{V} = s$, $\hat{G}(g) \in \text{GL}(s, \mathbb{C})$ for all $g \in G$, and the group law is mapped to the product of matrices.

Eventually, let us quote two results:

Theorem II.10. If $\mathcal{H}$ is a subgroup of a group $G$, then every representation of $G$ is also a representation of $\mathcal{H}$.

The proof follows at once from the existence of a natural homomorphism (the "canonical injection") $\mathcal{H} \rightarrow G$ which maps $g$ on itself, and from the fact that the composition of two group homomorphisms is again a homomorphism.

Theorem II.11. For every representation $\hat{G}$ of an Abelian group $G$, $\hat{G}(g_1)\hat{G}(g_2) = \hat{G}(g_2)\hat{G}(g_1)$ for all $g_1, g_2 \in G$.

II.3 Examples of group representations

II.3.1 Trivial representation

Theorem & Definition II.12. For any group $G$, the mapping which maps every element $g \in G$ to the identity operator $\hat{I}_C$ of $C$ is a representation of $G$, called the trivial representation.

Since $C$ is a one-dimensional complex vector space, the trivial representation is one-dimensional. For any scalar product $(\cdot, \cdot)$ on $C$, the identity $(\hat{I}_C(z), \hat{I}_C(z')) = (z, z')$ holds for all $z, z' \in C$ — by definition, $\hat{I}_C(z) = z$, so that the trivial representation is also unitary.

In any basis of the representation space $C$, the operator $\hat{I}_C$ is represented by the (unitary!) $1 \times 1$-matrix $(1)$, which explains why one often says that the trivial representation "maps every group element to 1".

When the group $G$ contains more than one element, the trivial representation is not faithful.

II.3.2 Representations of the dihedral group $D_3$

Let us now consider the non-Abelian dihedral group $D_3$, i.e. the group of the geometrical transformations that leave an equilateral triangle invariant, which is isomorphic to the symmetric group $S_3$. We call $A$, $B$, $C$ the corners of the triangle and introduce for further convenience a set of

---

(12) A unitary operator $\hat{U}$ resp. orthogonal operator $\hat{O}$ is an operator that preserves the scalar product $(\cdot, \cdot)$ on $\mathcal{V}$, i.e. $\forall x \in \mathcal{V}, (\hat{U}(x), \hat{U}(x)) = (x, x)$ resp. $(\hat{O}(x), \hat{O}(x)) = (x, x)$.

(13) D. HILBERT, 1862–1943
orthonormal basis vectors in the plane of the triangle, taking the origin of coordinates at the center of the triangle and $A$ along the $x$-direction (Fig. II.1).

The elements of the group are

- the identity transformation, which we denote by $\text{Id}$;
- the three reflections with respect to the medians of the triangle, we which denote $(A B)$, $(A C)$, $(B C)$ according to which pair of points they exchange;
- the rotations about $O$ through $2\pi/3$ and $4\pi/3$, respectively denoted $(A B C)$ and $(A C B)$.

Note that the notations reflect the isomorphism between the geometrical transformations and the permutations of the three corners.

We now introduce three different matrix representations of $D_3$.

**II.3.2 a Trivial representation**

As discussed in Sec. II.3.1, one may represent all six group elements by the same matrix

$$\mathcal{D}^{(1)}(g) = (1) \quad \forall g \in D_3,$$

which provides a non-faithful, unitary one-dimensional representation of $D_3$.

**II.3.2 b Sign representation**

Another one-dimensional unitary representation of $D_3$ consist of mapping the identity transformation and the two rotations to the matrix $(1)$:

$$\mathcal{D}^{(\varepsilon)}(\text{Id}) = \mathcal{D}^{(\varepsilon)}((A B C)) = \mathcal{D}^{(\varepsilon)}((A C B)) = (1)$$

and the three reflections to the matrix $(-1)$:

$$\mathcal{D}^{(\varepsilon)}((A B)) = \mathcal{D}^{(\varepsilon)}((A C)) = \mathcal{D}^{(\varepsilon)}((B C)) = (-1).$$

By noticing that this is the same thing as considering the signature $\varepsilon$ of the associated permutations, which is a homomorphism from $S_3 \cong D_3$ to the group $\{1, -1\}$, one checks that this is indeed a representation, called the *sign*, *signature* or *alternating representation*.

**Remark:** This representation can be introduced for any of the symmetric groups $S_n$. For $n \geq 2$ it “differs from” — that is, anticipating on a notion introduced in Sec. II.4 — it is not equivalent to — the trivial representation.

**II.3.2 c Standard representation**

Introducing Cartesian coordinates in the plane of the triangle, one can associate to each corner $A$, $B$, $C$ its coordinates $(x_A, y_A)$, $(x_B, y_B)$, $(x_C, y_C)$, and then represent each element of $D_3$ by the $2 \times 2$-matrix that performs the proper transformation of these coordinates. Accordingly we define six matrices

$$\mathcal{D}^{(s)}(\text{Id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
\[
\mathcal{D}^{(s)}((A \ B)) = \begin{pmatrix}
\frac{1}{2} & \frac{\sqrt{3}}{2} \\
\frac{\sqrt{3}}{2} & \frac{1}{2}
\end{pmatrix}, \quad \mathcal{D}^{(s)}((A \ C)) = \begin{pmatrix}
\frac{1}{2} & -\frac{\sqrt{3}}{2} \\
-\frac{\sqrt{3}}{2} & \frac{1}{2}
\end{pmatrix}, \quad \mathcal{D}^{(s)}((B \ C)) = \begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\]

and
\[
\mathcal{D}^{(s)}((A \ B \ C)) = \begin{pmatrix}
-\frac{1}{2} & -\frac{\sqrt{3}}{2} \\
\frac{\sqrt{3}}{2} & -\frac{1}{2}
\end{pmatrix}, \quad \mathcal{D}^{(s)}((A \ C \ B)) = \begin{pmatrix}
-\frac{1}{2} & \frac{\sqrt{3}}{2} \\
-\frac{\sqrt{3}}{2} & -\frac{1}{2}
\end{pmatrix}.
\]

One can check that these unitary matrices indeed form a group, which is isomorphic to D₃, and thus constitutes a faithful, two-dimensional unitary representation of D₃.

### II.3.3 Three-dimensional rotations

As last example (for the moment), let us consider the group of three-dimensional rotations, where a typical rotation will be denoted \(\mathcal{R}\). The example will also help us illustrate why group representations are of paramount importance in physics and how/where they play a role in a quantum-mechanical problem (§II.3.3b).

#### II.3.3a Standard representation

An obvious matrix representation of the group of three-dimensional rotations is that involving “rotation matrices”, i.e. real \(3 \times 3\)-matrices \(R\) such that \(RR^\dagger = R^\dagger R = 1_3\) and with determinant 1. These matrices constitute a group called SO(3).

#### II.3.3b Representations of the group of rotations in a quantum-mechanical problem

Consider a physical system \(\Sigma\), whose position with respect to a fixed reference frame is described by a vector \(\vec{r}\). In a quantum-mechanical context, and using the language of wave mechanics, the state of \(\Sigma\) is represented by a wave function \(\psi(\vec{r})\), where for the sake of simplicity we omit any time dependence.

Let us rotate the system \(\Sigma\), so that its new position is now described by a vector \(\vec{r}' = \mathcal{R}\vec{r}\). The state of the system is now represented by a new wave function \(\psi'(\vec{r})\), such that the probability density \(|\psi'(\vec{r})|^2\) has been rotated with respect to \(|\psi(\vec{r})|^2\).

**Remark:** The “physical” rotation \(\mathcal{R}\) of the system, say through an angle \(\alpha\) about direction \(\vec{n}\), which is considered here is a so-called active transformation. An alternative point of view consists in leaving \(\Sigma\) at the same place, while using a new reference frame, rotated about the same direction \(\vec{n}\) through the opposite angle \(-\alpha\), i.e. corresponding to \(\mathcal{R}^{-1}\); this is a passive transformation. Both points of view lead to the same mathematical equation (II.16).

Now, intuitively the new wave function \(\psi'\) is such that it has at the “rotated” position \(\vec{r}'\) the same value as the old wave function \(\psi\) at the original position \(\vec{r}\), i.e. \([13]\)

\[\psi'(\vec{r}') = \psi(\vec{r}) \quad \text{with} \quad \vec{r}' = \mathcal{R}\vec{r}.
\]

Inverting the transformation between \(\vec{r}\) and \(\vec{r}'\), we have \(\vec{r} = \mathcal{R}^{-1}\vec{r}'\) and thus

\[\psi'(\vec{r}') = \psi(\mathcal{R}^{-1}\vec{r}'),
\]

valid for all \(\vec{r}'\). In this equation, \(\vec{r}'\) is only a dummy variable appearing on both sides, which can also be written \(\vec{r}\), leading to

\[\psi(\vec{r}) = \psi(\mathcal{R}^{-1}\vec{r}).
\]

Summarizing, the rotation \(\vec{r} \rightarrow \vec{r}' = \mathcal{R}\vec{r}\) of the system induces a transformation

\[\psi \rightarrow \psi' = \mathcal{D}(\mathcal{R})\psi
\]

\([13]\) In truth, the only statement which can be made is that the probability density at the new position \(|\psi'(\vec{r}')|^2\) should be equal to that at the old position \(|\psi(\vec{r})|^2\).
of the wave function, where \( \psi' \) obeys Eq. (II.16). The operator \( \hat{D}(\mathcal{R}) \) acts on the functional space of wave functions, and \( \mathcal{D} \) is a representation of the group of rotations on that functional space.

The latter is in general infinite-dimensional, yet by introducing a properly chosen (Hilbert) basis, one can in practical cases work in a finite-dimensional subspace, i.e. with a finite-dimensional representation of the group of rotations.

As an illustration, consider a quantum-mechanical problem involving a single particle, and thus only one position-vector \( \vec{r} \), in a central potential \( V(\vec{r}) = V(r) \) with \( r \equiv |\vec{r}| \). A stationary solution of the problem is an energy-eigenfunction, which may be written after introducing spherical coordinates as \( \psi_{n\ell m}(\vec{r}) = R_{nt}(r)Y_{\ell m}(\theta, \varphi) \) with \( R_{nt} \) a problem-specific radial function and \( Y_{\ell m} \) a spherical harmonic.

If the system is rotated, it still remains in a stationary state, described by a new wave function \( \psi'_{n'\ell'm'}(\vec{r}) \). Let us discuss the relation of the “new” quantum numbers \( n', \ell', m' \) to the original ones \( n, \ell, m \).

- The principal quantum number \( n \) and the orbital quantum number \( \ell \in \mathbb{N} \) are related to the energy and the squared angular-momentum operator \( \hat{L}^2 \) of the system, which are both independent of the spatial orientation. Accordingly, a rotation will leave \( n \) and \( \ell \) unchanged, i.e. \( n' = n, \ell' = \ell \).

- In turn, the magnetic quantum number \( m \in \{-\ell, -\ell + 1, \ldots, \ell - 1, \ell\} \) quantifies the amount of angular momentum \( \hat{L}_z \) along a given direction, i.e. bears some relation to the orientation of the system: under a rotation, \( m \) will be modified.

Generally, an energy eigenfunction \( \psi'_{n'\ell'm'} \) for the rotated system with well-defined magnetic quantum number \( m' \) will be a linear combination of the energy eigenfunctions \( \psi_{n\ell m} \) of the non-rotated system with the same values of \( n \) and \( \ell \), yet with different \( m' \):

\[
\psi'_{n'\ell'm'} = \sum_{m=-\ell}^{\ell} \mathcal{D}_{m'm}^{(\ell)}(\mathcal{R}) \psi_{n\ell m}.
\]

The coefficients \( \mathcal{D}_{m'm}^{(\ell)}(\mathcal{R}) \) in this linear combination depend on the rotation \( \mathcal{R} \), and they are actually the matrix elements of a representation of the rotation group on the \( 2\ell+1 \)-dimensional space spanned by the \( \{\psi_{n\ell m}\} \) with fixed \( n \) and \( \ell \).

We shall come back to these representations of the rotation group in further detail later

### II.3.4 Direct sum and tensor product of representations

Let \( \hat{\mathcal{G}}^{(\alpha)} \), \( \hat{\mathcal{G}}^{(\beta)} \) be two linear representations of the same group \( \mathcal{G} \) on respective vector spaces \( \mathcal{V}^{(\alpha)} \), \( \mathcal{V}^{(\beta)} \).

#### II.3.4a Direct sum of representations

**Theorem & Definition II.18.** The mapping \( \hat{\mathcal{G}}^{(\alpha \oplus \beta)} \) from \( \mathcal{G} \) into the group of linear operators of the direct-sum space \( \mathcal{V}^{(\alpha)} \oplus \mathcal{V}^{(\beta)} \) to itself, such that for all \( g \in \mathcal{G}, x_\alpha \in \mathcal{V}^{(\alpha)}, x_\beta \in \mathcal{V}^{(\beta)} \)

\[
\hat{\mathcal{G}}^{(\alpha \oplus \beta)}(g)(x_\alpha + x_\beta) = \hat{\mathcal{G}}^{(\alpha)}(g)(x_\alpha) \oplus \hat{\mathcal{G}}^{(\beta)}(g)(x_\beta)
\]

(II.18)

is a (linear) representation of the group \( \mathcal{G} \), called the direct sum representation.

The dimension of the direct-space representation is the sum of the dimensions of the individual representations, namely \( \dim \mathcal{V}^{(\alpha)} + \dim \mathcal{V}^{(\beta)} \).

If the representations \( \hat{\mathcal{G}}^{(\alpha)}, \hat{\mathcal{G}}^{(\beta)} \) are finite-dimensional, by choosing a basis of \( \mathcal{V}^{(\alpha)} \oplus \mathcal{V}^{(\beta)} \) consisting of the union of a basis of \( \mathcal{V}^{(\alpha)} \) and a basis of \( \mathcal{V}^{(\beta)} \), one obtains a matrix representation of the operators \( \hat{\mathcal{G}}^{(\alpha \oplus \beta)}(g) \) of the block form.
II.4 Classifying representations

The two operations introduced in Sec. II.3.4 mean that the set of possible representations of a given group $G$ is infinitely large. To bring some order into that set, we now introduce two notions, diminishing the number of “independent” representations (Sec. II.4.1) and highlighting the role of “elementary building blocks” of specific representations (Sec. II.4.2).

II.4.1 Equivalent representations

Definition II.24. Two matrix representations $\hat{\mathcal{D}}^{(\alpha)}$, $\hat{\mathcal{D}}^{(\beta)}$ of the same dimension of a given group $G$ are called equivalent if there exists an invertible matrix $P$ such that for every element $g \in G$,

$$\hat{\mathcal{D}}^{(\alpha)}(g) = P \hat{\mathcal{D}}^{(\beta)}(g) P^{-1}. \quad (II.24)$$

L. Kronecker, 1823–1891
II.4 Classifying representations

That is, the matrices $D^{(\alpha)}(g)$ and $D^{(\beta)}(g)$ are similar for all $g \in G$, and the similarity matrix $P$ is the same for all pairs $(D^{(\alpha)}(g), D^{(\beta)}(g))$. Accordingly, for every $g \in G$ the linear operators $\hat{D}^{(\alpha)}(g)$ and $\hat{D}^{(\beta)}(g)$ are the same, yet represented in different bases of the representation space, with the same basis transformation for all $g \in G$.

**Example:** For instance, the two-dimensional rotation about a given point through an angle $\theta$ may be equivalently represented by the $2 \times 2$-matrices

\[
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}.
\]

Both representations are equivalent, since for all $\theta \in [0, 2\pi]$ one has

\[
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1}.
\]

“Being equivalent” obviously defines an equivalence relation on the set of representations of a given group, which is compatible with the group structure: for all $g, g' \in G$, one has

\[
D^{(\alpha)}(gg') = PD^{(\beta)}(g')P^{-1} = PD^{(\beta)}(g)D^{(\beta)}(g')P^{-1},
\]

where we used the group property of the matrices $D^{(\beta)}(g)$. Introducing an identity matrix written as $1 = P^{-1}P$, one then finds

\[
D^{(\alpha)}(g) = PD^{(\beta)}(g)P^{-1}PD^{(\beta)}(g')P^{-1} = D^{(\alpha)}(g)D^{(\alpha)}(g'),
\]

as it should be.

Instead of the representations, one investigates their equivalence classes, which means that two equivalent (matrix) representations are identified — which reduces the number of “different” representations.

**Remark:** Two one-dimensional representations of a group are equivalent if and only if they are indeed equal, since Eq. (II.24) is then an equality between numbers, which commute with each other.

### II.4.2 Reducible and irreducible representations

**Definition II.25.** A representation $\hat{\mathcal{D}}$ of a group $G$ is called **reducible** if there exists at least one proper subspace of the representation space which is invariant under all operators $\hat{\mathcal{D}}(g)$ with $g \in G$. A representation which is not reducible is said to be **irreducible**.

**Remark:** In the literature, one often finds “irrep” instead of “irreducible representation”.

**Property II.26.** A representation of dimension 1 is automatically irreducible, since the representation space admits no proper subspace.

When the reducible representation has a finite dimension $n_1 + n_2$, with $n_1$ the dimension of the invariant subspace, the corresponding matrices take the form

\[
\mathcal{D}(g) = \begin{pmatrix}
\mathcal{D}^{(1)}(g) & A(g) \\
0_{n_2 \times n_1} & \mathcal{D}^{(2)}(g)
\end{pmatrix}
\]

\[(14)\] A subspace $\mathcal{W}$ of a vector space $\mathcal{V}$ is said to be invariant under a linear operator $\hat{O}$ on $\mathcal{V}$ if $\hat{O}(\mathcal{W}) \subseteq \mathcal{W}$, i.e. if for all $x \in \mathcal{W}$, $\hat{O}(x) \in \mathcal{W}$.
in a basis whose first \(n_1\) vectors are a basis of the invariant subspace. \(\mathcal{D}^{(1)}(g)\) resp. \(\mathcal{D}^{(2)}(g)\) is a \(n_1 \times n_1\)- resp. \(n_2 \times n_2\)-matrix, \(A(g)\) a \(n_1 \times n_2\)-matrix, and \(0_{n_2 \times n_1}\) denotes the zero \(n_2 \times n_1\)-matrix.

The block product of two matrices \(\mathcal{D}(g)\) and \(\mathcal{D}(g')\) of the form \(\text{(II.26)}\), with \(g, g' \in \mathcal{G}\), gives

\[
\mathcal{D}(g) \mathcal{D}(g') = \begin{pmatrix}
\mathcal{D}^{(1)}(g) \mathcal{D}^{(1)}(g') & \mathcal{D}^{(1)}(g) A(g') + A(g) \mathcal{D}^{(2)}(g') \\
0_{n_2 \times n_1} & \mathcal{D}^{(2)}(g) \mathcal{D}^{(2)}(g')
\end{pmatrix},
\]

i.e. is again of the same form. On the other hand, the group law gives \(\mathcal{D}(g) \mathcal{D}(g') = \mathcal{D}(gg')\) with

\[
\mathcal{D}(gg') = \begin{pmatrix}
\mathcal{D}^{(1)}(gg') & A(gg') \\
0_{n_2 \times n_1} & \mathcal{D}^{(2)}(gg')
\end{pmatrix}.
\]

Comparing both expressions, one finds

\[
\mathcal{D}^{(1)}(g) \mathcal{D}^{(1)}(g') = \mathcal{D}^{(1)}(gg'), \quad \mathcal{D}^{(2)}(g) \mathcal{D}^{(2)}(g') = \mathcal{D}^{(2)}(gg'),
\]

which expresses the fact that \(\mathcal{D}^{(1)}\) and \(\mathcal{D}^{(2)}\) are smaller-dimensional (matrix) representations of \(\mathcal{G}\).

If \(\mathcal{D}^{(1)}\) and/or \(\mathcal{D}^{(2)}\) are themselves reducible, one can pursue the reduction in every invariant subspace, until one is only left with irreducible representations. The matrix representation eventually takes the “block triangular” form

\[
\mathcal{D}(g) = \begin{pmatrix}
\mathcal{D}^{(1)}(g) & A^{(1)}(g) \\
0 & \mathcal{D}^{(2)}(g) & A^{(2)}(g) \\
0 & 0 & \mathcal{D}^{(3)}(g) & A^{(3)}(g) \\
0 & 0 & 0 & \ddots & \ddots \\
0 & 0 & 0 & 0 & \mathcal{D}^{(r)}(g)
\end{pmatrix}
\]

for all \(g \in \mathcal{G}\), \(\text{(II.27)}\)

with \(\mathcal{D}^{(1)}, \ldots, \mathcal{D}^{(r)}\) irreducible matrix representations.

An even more favorable situation is when the latter are in fact fully decomposable, which we now discuss.

**Definition II.28.** A representation is called **fully reducible** or **decomposable** if to each invariant subspace there corresponds an invariant complementary subspace.

In that case, by choosing on the representation space a basis consisting in the union of bases of complementary subspaces, one can bring the matrices of the representation — if it is finite-dimensional — in a block-diagonal form:
II.4 Classifying representations

\[
D(g) = \begin{pmatrix}
D^{(1)}(g) & 0 \\
0 & D^{(2)}(g) \\
0 & 0 & D^{(3)}(g) \\
0 & 0 & 0 & \ddots & 0 \\
0 & 0 & 0 & 0 & D^{(r)}(g)
\end{pmatrix}
\]

for all \( g \in G \), \hspace{1cm} (II.29)

with \( D^{(1)}, \ldots, D^{(r)} \) irreducible matrix representations.

Comparing with Eq. (II.19), one sees that a decomposable representation is thus the direct sum of irreducible representations.

Accordingly, an important topic we shall deal with in the following chapters is the search for the irreducible representations of a given group. In addition, we shall learn how to decompose a fully reducible representation into irreducible representations.

Remark II.30. As illustrated in the following example, the reducibility of a representation may depend on the choice of the base field \( K \) of the representation space.

Example: Representations of the group of two-dimensional rotations

Consider the rotations about a fixed origin \( O \) in the \((y, z)\)-plane. Let \( R_\theta \) denote the rotation through a given angle \( \theta \in [0, 2\pi] \).

Viewing \( R_\theta \) as a three-dimensional rotation about the \( x \)-axis, one may represent it with the help of the \( 3 \times 3 \)-matrix

\[
D^{(3)}(R_\theta) = \begin{pmatrix}
1 & 0 & 0 \\
0 & \cos \theta & \sin \theta \\
0 & -\sin \theta & \cos \theta
\end{pmatrix},
\]

i.e. with an element of the matrix group \( SO(3) \).

Physically, it is clear that both the \((x, y)\)-plane and the \( z \)-axis, i.e. complementary subspaces, are invariant under all rotations \( R_\theta \), so that the three-dimensional representation \( D^{(3)} \) should be fully reducible. This is also clear when inspecting the matrix (II.31), which may be recast as

\[
D^{(3)}(R_\theta) = \begin{pmatrix}
1 & 0 & 0 \\
0 & \cos \theta & \sin \theta \\
0 & -\sin \theta & \cos \theta
\end{pmatrix},
\]

i.e. in the generic form (II.29) with \( D^{(1)}(R_\theta) \) a \( 1 \times 1 \)-matrix and \( D^{(2)}(R_\theta) \) a \( 2 \times 2 \)-matrix.

Since \( D^{(1)}(R_\theta) = (1) \) for every rotation \( R_\theta \), \( D^{(1)} \) is actually the trivial representation. In turn, the matrices

\[
D^{(2)}(R_\theta) = \begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix},
\]

of the matrix group \( SO(2) \) provide a faithful representation of the group of two-dimensional ro-
tions. This representation is irreducible on $\mathbb{R}$, yet considered on $\mathbb{C}$ it is reducible since the matrix \( \mathcal{D}' = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \), i.e. with complementary invariant subspaces, on which we have two inequivalent one-dimensional representations $\mathcal{D}'(\mathbb{R}) = (e^{i\theta})$ and $\mathcal{D}'(\mathbb{R})^* = (e^{-i\theta})$.

\section*{II.5 Schur’s lemma}

An important theorem, which is used to derive a lot of results on the irreducible representations of finite and compact groups, is that known as Schur's lemma:

\textbf{Theorem II.34} (Schur’s lemma).

1. Let $\hat{\mathcal{D}}, \hat{\mathcal{D}}'$ be two irreducible representations of a group $\mathcal{G}$ on respective vector spaces $\mathcal{V}, \mathcal{V}'$. Let $\hat{T} : \mathcal{V} \to \mathcal{V}'$ be a linear operator such that

\begin{equation}
\hat{T} \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}'(g) \circ \hat{T} \quad \text{for all } g \in \mathcal{G},
\end{equation}

Then either $\hat{T} = 0$, or $\hat{T}$ is an isomorphism.

2. Let $\hat{\mathcal{D}}'$ be a complex irreducible representation of a group $\mathcal{G}$ on a finite-dimensional vector space $\mathcal{V}$. If $\hat{T} : \mathcal{V} \to \mathcal{V}'$ is a linear operator such that

\begin{equation}
\hat{T} \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}'(g) \circ \hat{T} \quad \text{for all } g \in \mathcal{G},
\end{equation}

then $\hat{T}$ is a homothety, i.e. there exists $\lambda \in \mathbb{C}$ such that $\hat{T} = \lambda \mathbb{I}_\mathcal{V}$.

In this lemma, 0 denotes the zero operator, which maps every vector $x \in \mathcal{V}$ to the zero vector of the target space, while $\mathbb{I}_\mathcal{V}$ is the identity operator on $\mathcal{V}$.

To prove the first part of Schur’s lemma, let us note that Eq. (II.34a) implies two properties. First, the kernel of $\hat{T}$ is invariant by all operators $\hat{\mathcal{D}}(g)$ with $g \in \mathcal{G}$:

Let $x \in \ker \hat{T}$ and $g \in \mathcal{G}$. Then $\hat{\mathcal{D}}'(g) \circ \hat{T}(x) = \hat{\mathcal{D}}'(g)(\hat{T}(x)) = \hat{\mathcal{D}}'(g)(0_{\mathcal{V}'}) = 0_{\mathcal{V}'}$, with $0_{\mathcal{V}'}$ the zero vector of $\mathcal{V}'$. Eq. (II.34a) then gives $\hat{T} \circ \hat{\mathcal{D}}(g)(x) = 0_{\mathcal{V}'} = \hat{T}(\hat{\mathcal{D}}(g)(x))$, i.e. $\hat{\mathcal{D}}(g)(x) \in \ker \hat{T}$.

In turn the image of $\hat{T}$ is invariant by all operators $\hat{\mathcal{D}}'(g)$:

Let $g \in \mathcal{G}$ and $x' \in \im \hat{T}$: there exists $x \in \mathcal{V}$ such that $x' = \hat{T}(x)$. Applying Eq. (II.34a) to the vector $x$ gives

$$\hat{T} \circ \hat{\mathcal{D}}'(g)(x) = \hat{\mathcal{D}}'(g) \circ \hat{T}(x) \quad \text{i.e. } \hat{T}(\hat{\mathcal{D}}'(g)(x)) = \hat{\mathcal{D}}'(g)\hat{T}(x) = \hat{\mathcal{D}}'(g)(x'),$$

which shows that $\hat{\mathcal{D}}'(g)(x')$ is indeed an element of $\im \hat{T}$.

Now, since the representations $\hat{\mathcal{D}}$ resp. $\hat{\mathcal{D}}'$ is by assumption irreducible, $\ker \hat{T}$ resp. $\im \hat{T}$ has to be an improper subspace of $\mathcal{V}$ resp. $\mathcal{V}'$. If $\ker \hat{T}$ is the whole vector space $\mathcal{V}$, then $\hat{T}$ is the zero operator (and $\im \hat{T}$ reduces to the zero vector of $\mathcal{V}'$). Otherwise, it means that $\ker \hat{T}$ reduces to the zero vector of $\mathcal{V}$, i.e. $\hat{T}$ is injective. Since $\im \hat{T}$ is an improper subspace, and differs from $\{0_{\mathcal{V}'}\}$, it has to be $\mathcal{V}'$ itself, i.e. $\hat{T}$ is also surjective, and thereby bijective, i.e. it is an isomorphism.

\textbf{Remark:} If $\mathcal{V}$ and $\mathcal{V}'$ have the same dimension, then the existence of a bijective linear operator $\hat{T}$ obeying Eq. (II.34a) means that $\hat{\mathcal{D}}$ and $\hat{\mathcal{D}}'$ are equivalent.

\footnote{\(\hat{T}\) is called an \textit{intertwining operator}.}

\footnote{I. Schur, 1875–1941}
Let us now prove the second part of Schur’s lemma II.34. Since it now deals with operators on a complex finite-dimensional vector space, the characteristic polynomial \( \det(\hat{T} - \lambda \hat{1}_V) \) admits at least one root \( \lambda \in \mathbb{C} \), i.e. an eigenvalue of \( \hat{T} \) to which corresponds at least a non-zero eigenvector. Subtracting \( -\lambda \hat{D}(g) = -\lambda \hat{1}_V \circ \hat{D}(g) = \hat{D}(g) \circ (\hat{T} - \lambda \hat{1}_V) \) from Eq. (II.34b) yields
\[
(\hat{T} - \lambda \hat{1}_V) \circ \hat{D}(g) = \hat{D}(g) \circ (\hat{T} - \lambda \hat{1}_V)
\]
for all \( g \in \mathcal{G} \), from where follows that the kernel of \( \hat{T} - \lambda \hat{1}_V \) is invariant under all operators \( \hat{D}(g) \) with \( g \in \mathcal{G} \).

Since the representation \( \hat{D} \) is irreducible, \( \ker(\hat{T} - \lambda \hat{1}_V) \) is either reduced to the zero vector, which is excluded by the fact that \( \lambda \) is by construction eigenvalue to \( \hat{T} \), or the whole vector space \( V \), which precisely means \( \hat{T} = \lambda \hat{1}_V \).

A first, important consequence of Schur’s lemma is the following result.

**Corollary II.35.** The only irreducible finite-dimensional complex representations of an Abelian group are its representations of dimension one.

The operators of any representation \( \hat{D} \) of an Abelian group \( \mathcal{G} \) commute with each other: fixing \( g' \in \mathcal{G} \), \( \hat{D}(g) \hat{D}(g') = \hat{D}(g') \hat{D}(g) \) for all \( g \in \mathcal{G} \). If \( \hat{D} \) is a finite-dimensional complex irreps., the operator \( \hat{D}(g') \) obeys Eq. (II.34b) and the assumptions of the second part of Schur’s lemma are fulfilled, so that \( \hat{D}(g') \) is an homothety — and this holds for all \( g' \in \mathcal{G} \). Since every subspace of a vector space is invariant by an homothety, \( \hat{D} \) can only be irreducible if it is of dimension one, so that the representation space has no proper invariant space.

To finish this Section, let us give without proof a partial converse to Schur’s lemma:

**Theorem II.36.** Let \( \hat{D} \) be a fully reducible finite-dimensional representation of a group \( \mathcal{G} \) on a vector space \( V \). If the only linear operators on \( V \) that commute with all operators \( \hat{D}(g) \) with \( g \in \mathcal{G} \) are the homotheties of \( V \), then \( \hat{D} \) is irreducible.
In this chapter, we focus on the case of the representations of finite groups. We first argue in Sec. III.1 that such a representation is always decomposable, which displaces the focus towards irreducible representations. For the latter, we first derive in Sec. III.2 a set of relations. These are then exploited in Sec. III.3 to constrain a number of properties of the unitary irreducible representations, in particular their characters. In Sec. III.4, we discuss how one can determine which irreducible representations, and with which multiplicity, enter the decomposition of a given representation. In particular, we introduce the regular representation of a finite group as well as the notion of the group algebra. Eventually, Section III.5 deals with the special case of the representations of the symmetric groups $S_n$.

### III.1 Full reducibility of the representations

We first state a theorem which is valid not only for finite groups, but also for infinite ones.

**Theorem III.1.** Every reducible finite-dimensional unitary representation of an arbitrary group $G$ is decomposable.

Let $\hat{D}$ be a representation of a finite group $G$ on the vector space $V$, on which a scalar product $(\cdot, \cdot)$ has been defined. We shall first introduce a second scalar product $\langle \cdot, \cdot \rangle$ on $V$ for which $\hat{D}$ is unitary. If $\hat{D}$ is reducible, then there exists a proper subspace $\mathcal{W}$ of $V$ which is invariant under all operators $\hat{D}(g)$ with $g \in G$. Let $\mathcal{W}^\perp$ denote the orthogonal complement of $\mathcal{W}$; since $V$ is finite-dimensional, $\mathcal{W}^\perp$ is a complementary subspace of $\mathcal{W}$ in $V$. We shall show that $\mathcal{W}^\perp$ is invariant under all operators $\hat{D}(g)$ with $g \in G$, thereby proving the theorem.

For all $g \in G$, $x \in \mathcal{W}^\perp$ and $y \in \mathcal{W}$, consider the scalar product $\langle \hat{D}(g)(x), y \rangle$. Invoking the unitarity property of $\hat{D}(g)$, i.e. $\hat{D}(g)^{-1} = \hat{D}(g)^\dagger$, one has

$$\langle \hat{D}(g)(x), y \rangle = \langle x, \hat{D}(g)^{-1}(y) \rangle = \langle x, \hat{D}(g^{-1})(y) \rangle.$$ 

From the invariance of $\mathcal{W}$ under $\hat{D}$, i.e. in particular under $\hat{D}(g^{-1})$, follows $\hat{D}(g^{-1})(y) \in \mathcal{W}$, so that the rightmost scalar product in the above equation is zero. Thus $\langle \hat{D}(g)(x), y \rangle = 0$ for all $y \in \mathcal{W}$, i.e. $\hat{D}(g)(x) \in \mathcal{W}^\perp$, which shows the invariance of $\mathcal{W}^\perp$ under $\hat{D}(g)$.  

**Theorem III.2.** Any representation of a finite group is equivalent to a unitary representation.

Let $\hat{D}$ be a representation of a finite group $G$ on the vector space $V$, on which a scalar product $(\cdot, \cdot)$ has been defined. We shall first introduce a second scalar product $\langle \cdot, \cdot \rangle$ on $V$, for which every operator $\hat{D}(g)$ with $g \in G$ will be unitary. Then we shall find a representation $\hat{D}'$ equivalent to $\hat{D}$ and such that the operators $\hat{D}(g)$ are actually also unitary for the scalar product $(\cdot, \cdot)$.

For $x, y \in V$, the formula

$$\langle x, y \rangle = \frac{1}{|G|} \sum_{g \in G} \langle \hat{D}(g)(x), \hat{D}(g)(y) \rangle$$

defines a mapping which is bilinear and symmetric in $x$ and $y$, and which is positive definite since this is the case of the scalar product $(\cdot, \cdot)$; it is thus a scalar product (or inner product) on $V$. For
every \( g \in \mathcal{G} \) and \( x, y \in \mathcal{V} \), one has
\[
\langle \hat{\mathcal{D}}(g)(x), \hat{\mathcal{D}}(g)(y) \rangle = \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \langle \hat{\mathcal{D}}(g')(\hat{\mathcal{D}}(g)(x)), \hat{\mathcal{D}}(g')(\hat{\mathcal{D}}(g)(y)) \rangle
\]
i.e.
\[
\langle \hat{\mathcal{D}}(g)(x), \hat{\mathcal{D}}(g)(y) \rangle = \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \langle \hat{\mathcal{D}}(g') \circ \hat{\mathcal{D}}(g)(x), \hat{\mathcal{D}}(g') \circ \hat{\mathcal{D}}(g)(y) \rangle.
\]

The group property of the operators \( \hat{\mathcal{D}}(g) \) gives \( \hat{\mathcal{D}}(g') \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}(g'g) \). In addition, at fixed \( g \in \mathcal{G} \), a sum over all \( g' \in \mathcal{G} \) is equivalent to a sum over all \( g'' = g'g \in \mathcal{G} \) (group rearrangement theorem \[ I.7 \]), leading eventually to
\[
\langle \hat{\mathcal{D}}(g)(x), \hat{\mathcal{D}}(g)(y) \rangle = \frac{1}{|\mathcal{G}|} \sum_{g'' \in \mathcal{G}} \langle \hat{\mathcal{D}}(g'')(x), \hat{\mathcal{D}}(g'')(y) \rangle = \langle x, y \rangle.
\]

This identity, valid for all \( g \in \mathcal{G} \) and \( x, y \in \mathcal{V} \), means that the operator \( \hat{\mathcal{D}}(g) \) is unitary for the scalar product \( \langle \cdot, \cdot \rangle \).

Consider now two bases \( \{ e_i \} \) and \( \{ e_j' \} \) of the space \( \mathcal{V} \), which are respectively orthonormal for the scalar products \( \langle \cdot, \cdot \rangle \) resp. \( \langle \cdot, \cdot \rangle' \). There exist a non-singular linear operator \( \hat{M} \) such that \( \hat{M}(e_i) = e_i' \) for all \( i \) — this is precisely the linear operator corresponding to the transformation matrix of the change of basis. Now, using the definitions of \( \hat{M} \) and of the orthonormal bases, one can write for all \( i,j \)
\[
\langle \hat{M}(e_i), \hat{M}(e_j) \rangle = \langle e_i', e_j' \rangle = \delta_{ij} = \langle e_i, e_j \rangle
\]
with \( \delta_{ij} \) the usual Kronecker symbol. The leftmost and rightmost terms in this equation are bilinear, which leads for every \( x, y \in \mathcal{V} \) to
\[
\langle \hat{M}(x), \hat{M}(y) \rangle = \langle x, y \rangle. \tag{III.4}
\]

Let us define a new representation \( \hat{\mathcal{D}}' \) of \( \mathcal{G} \) by \( \hat{\mathcal{D}}'(g) = \hat{M}^{-1} \circ \hat{\mathcal{D}} \circ \hat{M} \) for all \( g \in \mathcal{G} \): by construction \( \hat{\mathcal{D}}' \) is equivalent to \( \hat{\mathcal{D}} \). For all \( g \in \mathcal{G} \) and \( x, y \in \mathcal{V} \) one has
\[
\langle \hat{\mathcal{D}}'(g)(x), \hat{\mathcal{D}}'(g)(y) \rangle = \langle \hat{M}^{-1} \circ \hat{\mathcal{D}} \circ \hat{M}(x), \hat{M}^{-1} \circ \hat{\mathcal{D}} \circ \hat{M}(y) \rangle = \langle \hat{\mathcal{D}}(g) \circ \hat{M}(x), \hat{\mathcal{D}}(g) \circ \hat{M}(y) \rangle\]
where we used Eq. \( \text{[III.4]} \). Using the unitarity of \( \hat{\mathcal{D}}(g) \) for the scalar product \( \langle \cdot, \cdot \rangle \) then yields
\[
\langle \hat{\mathcal{D}}'(g)(x), \hat{\mathcal{D}}'(g)(y) \rangle = \langle \hat{M}(x), \hat{M}(y) \rangle.
\]
Eventually, invoking again Eq. \( \text{[III.4]} \), one obtains
\[
\langle \hat{\mathcal{D}}'(g)(x), \hat{\mathcal{D}}'(g)(y) \rangle = \langle x, y \rangle,
\]
i.e. \( \hat{\mathcal{D}}'(g) \) is unitary for the scalar product \( \langle \cdot, \cdot \rangle \).

**Remark:** It is interesting to note where the finiteness of the group was used in the proof. The factor \( 1/|\mathcal{G}| \) in the definition \( \text{[III.3]} \) is actually irrelevant for the reasoning, as one easily checks. What is really crucial is the possibility to sum over the elements of the group: when we turn to compact infinite groups, this sum will be replaced by an integral, yet this is not possible for non-compact infinite groups.

Combining the theorems \( \text{[III.1]} \) and \( \text{[III.2]} \) leads to Maschke's \([4]\) theorem:

**Theorem III.5** (Maschke's theorem). All finite-dimensional reducible representations of a finite group are decomposable.

**Remark III.6.** One can also show that every irreducible representation of a finite group is in fact finite-dimensional.

All in all, we see that to find all representations of a finite group \( \mathcal{G} \), it is sufficient to find its unitary irreducible representations, which is the topic of the next two Sections.

\[^{[4]}\text{H. Maschke, 1853–1909}\]
### III.2 Orthogonality relations

**Theorem III.7.** Let \( \hat{\mathcal{D}} \) and \( \hat{\mathcal{D}}' \) be two irreducible representations of a finite group \( \mathcal{G} \) on respective vector spaces \( \mathcal{V}, \mathcal{V}' \). Given a linear operator \( \phi: \mathcal{V} \to \mathcal{V}' \), one defines an operator \( \mathcal{V} \to \mathcal{V}' \) by

\[
\hat{T}_\mathcal{G}(\phi) \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \hat{\mathcal{D}}'(g) \circ \phi \circ \hat{\mathcal{D}}(g^{-1}) \tag{III.7}
\]

- If \( \hat{\mathcal{D}} \) and \( \hat{\mathcal{D}}' \) are not equivalent, then \( \hat{T}_\mathcal{G}(\phi) = 0 \).
- If \( \hat{\mathcal{D}} = \hat{\mathcal{D}}' \), then \( \hat{T}_\mathcal{G}(\phi) \) is a homothety: \( \hat{T}_\mathcal{G}(\phi) = \frac{\text{Tr} \phi}{\dim \mathcal{V}} \hat{1}_\mathcal{V} \).

In this theorem, \( \hat{0} \) denotes the zero operator, which maps every vector \( x \in \mathcal{V} \) to the zero vector of the target space \( \mathcal{V}' \), while \( \hat{1}_\mathcal{V} \) is the identity operator on \( \mathcal{V} \).

**Proof:** One first checks that \( \hat{T}_\mathcal{G}(\phi) \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}'(g) \circ \hat{T}_\mathcal{G}(\phi) \) for all \( g \in \mathcal{G} \):

\[
\hat{T}_\mathcal{G}(\phi) \circ \hat{\mathcal{D}}(g) = \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \hat{\mathcal{D}}'(g') \circ \phi \circ \hat{\mathcal{D}}(g^{-1}) \circ \hat{\mathcal{D}}(g) = \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \hat{\mathcal{D}}'(g') \circ \phi \circ \hat{\mathcal{D}}(g^{-1}g),
\]

where we used the group law of the operators \( \hat{\mathcal{D}}(g) \). Now \( g^{-1}g = (g^{-1}g)^{-1} \) and we write

\[
\hat{T}_\mathcal{G}(\phi) \circ \hat{\mathcal{D}}(g) = \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \hat{\mathcal{D}}'(g') \hat{\mathcal{D}}(g^{-1}g') \circ \phi \circ \hat{\mathcal{D}}((g^{-1}g')^{-1})
\]

\[
\hat{T}_\mathcal{G}(\phi) \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}'(g) \frac{1}{|\mathcal{G}|} \sum_{g' \in \mathcal{G}} \hat{\mathcal{D}}'(g^{-1}g') \circ \phi \circ \hat{\mathcal{D}}((g^{-1}g')^{-1}).
\]

Letting \( g'' \equiv g^{-1}g' \), \( g'' \) runs over the whole group \( \mathcal{G} \) when \( g' \) does at fixed \( g \), which gives the announced result \( \hat{T}_\mathcal{G}(\phi) \circ \hat{\mathcal{D}}(g) = \hat{\mathcal{D}}'(g) \circ \hat{T}_\mathcal{G}(\phi) \).

Invoking Schur’s lemma \[\text{II.34}\] gives that either \( \hat{T}_\mathcal{G}(\phi) = 0 \), namely precisely when the two representations are inequivalent, or \( \hat{T}_\mathcal{G}(\phi) = \lambda \hat{1}_\mathcal{V} \) with some \( \lambda \in \mathbb{C} \), which we can compute when \( \hat{\mathcal{D}} = \hat{\mathcal{D}}' \). Writing that the right hand side of Eq. \[\text{II.7}\] with \( \hat{\mathcal{D}} = \hat{\mathcal{D}} \) equals \( \lambda \hat{1}_\mathcal{V} \) and computing the trace gives, with the linearity of the trace and \( \text{Tr} \hat{1}_\mathcal{V} = \dim \mathcal{V} \),

\[
\lambda \dim \mathcal{V} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \text{Tr} [\hat{\mathcal{D}}(g) \circ \phi \circ \hat{\mathcal{D}}(g^{-1})] = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \text{Tr} [\hat{\mathcal{D}}(g^{-1}) \circ \hat{\mathcal{D}}(g) \circ \phi],
\]

where we used the cyclicity of the trace. The sum on the right hand side of this equation is just \( |\mathcal{G}| \) times the trace \( \text{Tr} \phi \), which gives \( \lambda = \text{Tr} \phi / |\dim \mathcal{V}| \).

**Corollary III.8.** Let \( \mathcal{D}, \mathcal{D}' \) be two irreducible matrix representations of a finite group \( \mathcal{G} \).

- If \( \mathcal{D} \) and \( \mathcal{D}' \) are not equivalent, then

\[
\frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} [\mathcal{D}'(g)]_{ij} [\mathcal{D}(g^{-1})]_{kl} = \delta_{ij} \delta_{kl} \quad \forall i, j, k, l. \tag{III.8a}
\]

- If \( \mathcal{D} = \mathcal{D}' \), then

\[
\frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} [\mathcal{D}(g)]_{ij} [\mathcal{D}(g^{-1})]_{kl} = \delta_{il} \delta_{jk} \frac{\dim \mathcal{V}}{|\mathcal{G}|} \quad \forall i, j, k, l. \tag{III.8b}
\]

**Proof:** In matrix representation, the Eq. \[\text{III.7}\] reads

\[
[T_G(\phi)]_{ij} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \sum_{j', k'} [\mathcal{D}'(g)]_{ij'} [\phi]_{j'k'} [\mathcal{D}(g^{-1})]_{k'l} \quad \forall i, l \tag{III.9}
\]

where the indices \( i, j', k', l \) takes values between 1 and the relevant vector-space dimension. Choosing for \( \phi \) the operator whose matrix elements are \( [\phi]_{j'k'} = \delta_{j'j} \delta_{kk'} \), where \( j \) and \( k \) are arbitrary, and applying the theorem \[\text{III.7}\] yields the results \[\text{III.8a}\] and \[\text{III.8b}\].
III.3 Irreducible representations of a finite group

Let us from now on label the irreducible representations of a given finite group \( G \) — or more accurately, the various equivalence classes of irreducible representations of \( G \) — with a superscript bracketed Greek letter: \( \hat{\mathcal{D}}^{(\alpha)} \), \( \hat{\mathcal{D}}^{(\beta)} \), \ldots. The corresponding representation spaces will be similarly denoted \( V^{(\alpha)} \), \( V^{(\beta)} \), and so on.

Using this notation, both equations (III.8a)--(III.8b) can be gathered into a single one:

\[
\frac{1}{|G|} \sum_{g \in G} \left[ \hat{\mathcal{D}}^{(\alpha)}(g) \right]_{ij} \left[ \hat{\mathcal{D}}^{(\beta)}(g^{-1}) \right]_{kl} = \frac{\delta^{\alpha\beta} \delta_{il} \delta_{jk}}{\dim V^{(\alpha)}} \quad \forall i, j, k, l
\]  

(III.10a)

known as (Schur’s) orthogonality relations, where the Kronecker symbol \( \delta^{\alpha\beta} \) equals 1 if the representations are the same, and 0 if they are inequivalent.

Invoking theorem III.2, we may assume that the irreducible matrix representations we are dealing with are unitary. In that case, using \( \mathcal{D}^{(\beta)}(g^{-1}) = \mathcal{D}^{(\beta)}(g)^\dagger \), which follows from the group law, one has

\[
\mathcal{D}^{(\beta)}(g^{-1}) = \mathcal{D}^{(\beta)}(g)^\dagger,
\]

where \( ^\dagger \) denotes the Hermitian conjugate matrix, i.e. for the matrix elements

\[
\left[ \mathcal{D}^{(\beta)}(g^{-1}) \right]_{kl} = \left[ \mathcal{D}^{(\beta)}(g)^\dagger \right]_{lk} \quad \forall k, l.
\]

The orthogonality relations can thus be recast as

\[
\frac{1}{|G|} \sum_{g \in G} \left[ \mathcal{D}^{(\alpha)}(g) \right]_{ij} \left[ \mathcal{D}^{(\beta)}(g)^\dagger \right]_{lk} = \frac{\delta^{\alpha\beta} \delta_{il} \delta_{jk}}{\dim V^{(\alpha)}} \quad \forall i, j, k, l.
\]  

(III.10b)

In the following Section we shall exploit these relations to derive constraints obeyed by the irreducible representations of a finite group.

### III.3 Irreducible representations of a finite group

#### III.3.1 First consequences of the orthogonality relations

**Theorem & Definition III.11.** Let \( \mathcal{G} \) be a finite group. The expression

\[
(\psi, \varphi)_G = \frac{1}{|G|} \sum_{g \in \mathcal{G}} \varphi(g) \psi(g)^\dagger
\]  

with \( \varphi, \psi : \mathcal{G} \to \mathbb{C} \) defines a scalar product on the space of complex-valued functions on \( \mathcal{G} \).

**Proof:** One easily checks that the expression on the right-hand side is sesquilinear, positive, and definite.

**Remark:** We used the physicists’ convention for sesquilinearity, namely linearity resp. antilinearity in the second resp. first argument.

At fixed \( i \) and \( j \) resp. \( k \) and \( l \), the matrix elements \( \left[ \mathcal{D}^{(\alpha)}(g) \right]_{ij} \) and \( \left[ \mathcal{D}^{(\beta)}(g)^\dagger \right]_{lk} \) are the values in \( g \in \mathcal{G} \) of complex-valued functions \( \mathcal{D}^{(\alpha)} \), \( \mathcal{D}^{(\beta)} \) on \( \mathcal{G} \), and the term on the left hand side of the orthogonality relation (III.10b) is the scalar product of these functions as defined in Eq. (III.11):

\[
(\left[ \mathcal{D}^{(\beta)} \right]_{lk}, \left[ \mathcal{D}^{(\alpha)} \right]_{ij})_G = \frac{\delta^{\alpha\beta} \delta_{il} \delta_{jk}}{\dim V^{(\alpha)}} \quad \forall i, j, k, l.
\]  

(III.12)

Assume first \( \alpha = \beta \), i.e. identical representations, so that \( \delta^{\alpha\beta} = 1 \) and the four indices \( i, j, k, l \) can each take \( \dim V^{(\alpha)} \) different values. Unless \( i = l \) and \( j = k \), the scalar product of \( \left[ \mathcal{D}^{(\alpha)} \right]_{ij} \) and
[\mathcal{D}(\alpha)]_{ij} \text{ vanishes, i.e. the functions are orthogonal to each other. As } i, j \in \{1, \ldots, \dim \mathcal{V}(\alpha)\} \text{ vary, there are } (\dim \mathcal{V}(\alpha))^2 \text{ such mutually orthogonal functions.}

Considering now another irreducible representation \(\mathcal{D}(\beta)\) with \(\beta \neq \alpha\), the \((\dim \mathcal{V}(\beta))^2\) corresponding functions \([\mathcal{D}(\beta)]_{ij}\) are likewise mutually orthogonal, but relation \((\text{III.12})\) with \(\delta^{\alpha\beta}\) shows, that they are also orthogonal to the functions \([\mathcal{D}(\beta)]_{ij}\). Repeating the argument for all (classes of) inequivalent irreducible representations of the group \(\mathcal{G}\), one finds that their respective matrix elements provide a set of

$$\sum_{\text{irreps } \alpha} (\dim \mathcal{V}(\alpha))^2$$

pairwise orthogonal complex-valued functions on \(\mathcal{G}\). Now, such a function is defined by the \(|\mathcal{G}|\) values it can take, so that the space of these functions is a vector space on \(\mathbb{C}\) of dimension \(|\mathcal{G}|\). On a \(|\mathcal{G}|\)-dimensional vector space, there can be at most \(|\mathcal{G}|\) mutually orthogonal vectors. All in all, the dimensions of the various inequivalent irreducible representations of a finite group \(\mathcal{G}\) must obey the inequality

$$\sum_{\text{irreps } \alpha} (\dim \mathcal{V}(\alpha))^2 \leq |\mathcal{G}|.$$  \((\text{III.13})\)

An important consequence of this inequality is that the number of inequivalent irreducible representations of a finite group is finite, and their dimensions are bounded from above.

**Example:** In Sec. 1.3.2 we introduced three representations of the dihedral group \(D_3\): the trivial representation \(\mathcal{D}(1)\), the sign representation \(\mathcal{D}(\varepsilon)\), and the standard representation \(\mathcal{D}(s)\). The reader can check that the latter is actually irreducible (also in \(\mathbb{C}\)), as are the first two ones since they are one-dimensional. We have thus found three representations with respective dimensions 1, 1 and 2, for a group with 6 elements, thereby saturating the inequality \((\text{III.13})\).

Using the regular representation of \(\mathcal{G}\), we shall show in Sec. 1.4.2 that in fact the equality always holds in Eq. \((\text{III.13})\):

**Theorem III.14.** The sum of the squares of the dimensions of a complete set of inequivalent irreducible representations of a finite group \(|\mathcal{G}|\) equals the order of the group:

$$\sum_{\text{irreps } \alpha} (\dim \mathcal{V}(\alpha))^2 = |\mathcal{G}|.$$  \((\text{III.14})\)

### III.3.2 Character theory

#### III.3.2a Definition and first properties

**Definition III.15.** Let \(\hat{\mathcal{D}}\) be a finite-dimensional complex representation of a group \(\mathcal{G}\). The character of \(\hat{\mathcal{D}}\) is the complex-valued function \(\chi_{\hat{\mathcal{D}}}\) on \(\mathcal{G}\) defined by

$$\chi_{\hat{\mathcal{D}}}(g) = \text{Tr} [\hat{\mathcal{D}}(g)] \text{ for all } g \in \mathcal{G}. $$ \((\text{III.15})\)

**Property III.16.** For all \(g, h \in \mathcal{G}\), the cyclicity of the trace gives \(\chi_{\hat{\mathcal{D}}}(hgh^{-1}) = \chi_{\hat{\mathcal{D}}}(g)\), i.e. the character of a representation \(\hat{\mathcal{D}}\) is constant on a conjugacy class of the group \(\mathcal{G}\). Accordingly, \(\chi_{\hat{\mathcal{D}}}\) is referred to as a class function.

**Definition III.17.** A character \(\chi_{\hat{\mathcal{D}}}\) is called simple or irreducible if the corresponding representation \(\hat{\mathcal{D}}\) is itself irreducible. In contrast, the character of a reducible representation is called a compound character.
III.3 Irreducible representations of a finite group

Going back to the relevant definitions, one easily proves the following properties of characters:

- Equivalent representations \( \hat{D}, \hat{D}' \) of a group have the same character \( \chi_{\hat{D}} = \chi_{\hat{D}'} \).
- The character of a direct-sum representation \( \hat{D} \oplus \hat{D}' \) is the sum of the respective characters of \( \hat{D} \) and \( \hat{D}' \):
  \[
  \chi_{\hat{D} \oplus \hat{D}'} = \chi_{\hat{D}} + \chi_{\hat{D}'}.
  \]  
  \( \text{(III.18)} \)
- More generally, if \( \hat{D} = \bigoplus_{\alpha} a_\alpha \hat{D}^{(\alpha)} \), then
  \[
  \chi_{\hat{D}} = \sum_{\alpha} a_\alpha \chi_{\hat{D}^{(\alpha)}}.
  \]  
  \( \text{(III.19)} \)
  In particular, if the representations \( \hat{D}^{(\alpha)} \) are irreducible, this identity shows that a compound character is a linear combination (with positive integer coefficients) of simple characters.
- The character of a tensor-product representation \( \hat{D} \otimes \hat{D}' \) is the product of the respective characters of \( \hat{D} \) and \( \hat{D}' \):
  \[
  \chi_{\hat{D} \otimes \hat{D}'} = \chi_{\hat{D}} \cdot \chi_{\hat{D}'}.
  \]  
  \( \text{(III.20)} \)
  where the dot stands for the multiplication in \( \mathbb{C} \) of the values taken by the characters.
- The characters of a representation \( \hat{D} \) and its conjugate \( \hat{D}^* \) are conjugate to each other:
  \[
  \chi_{\hat{D}}^* = (\chi_{\hat{D}})^*.
  \]  
  \( \text{(III.21)} \)
- If \( \hat{D} \) is a unitary representation, then
  \[
  \chi_{\hat{D}}(g^{-1}) = \chi_{\hat{D}}(g)^* \quad \text{for all } g \in G.
  \]  
  \( \text{(III.22)} \)

Property III.23. If \( \hat{D} \) is a one-dimensional representation of a group \( G \), then the complex numbers \( \chi_{\hat{D}}(g) \) should obey the group law, i.e. \( \chi_{\hat{D}}(gg') = \chi_{\hat{D}}(g) \chi_{\hat{D}}(g') \) for all \( g, g' \in G \).

### III.3.2 b Orthogonality of characters

Let us consider again the orthogonality relation \( \text{(III.10b)} \). If we multiply the term on the left hand side with \( \delta_{ij} \) and sum over all possible values of \( i \) and \( j \), we are actually computing the trace of \( D^{(\alpha)}(g) \). According to definition \( \text{III.15} \), this trace is precisely the value in \( g \) of the character of \( D^{(\alpha)} \), which we shall denote \( \chi^{(\alpha)}(g) \). Similarly, one can also trace the relation \( \text{(III.10b)} \) over the indices \( k \) and \( l \), which yields

\[
\frac{1}{|G|} \sum_{g \in G} \chi^{(\alpha)}(g) \chi^{(\beta)}(g)^* = \frac{\delta_{\alpha\beta}}{\dim Y^{(\alpha)}} \sum_{i,j,k,l} \delta_{ii} \delta_{jj} \delta_{kk} \delta_{ll}.
\]

The term on the left hand side is precisely the scalar product \( \text{(III.11)} \) of the characters \( \chi^{(\alpha)} \) and \( \chi^{(\beta)} \). In turn, the sum appearing on the right hand side is readily computed and equals \( \dim Y^{(\alpha)} \), canceling the similar factor present in the denominator. All in all, one thus finds

\[
(\chi^{(\beta)}, \chi^{(\alpha)})_G = \frac{1}{|G|} \sum_{g \in G} \chi^{(\alpha)}(g) \chi^{(\beta)}(g)^* = \delta_{\alpha\beta},
\]

which means that the characters of inequivalent irreducible representations of a given group \( G \) are orthogonal — and in fact orthonormal — for the scalar product \( \text{(III.11)} \).

For every element \( g \in G \), let us denote \([g]\) the conjugacy class of \( g \) in \( G \), and \(|[g]|\) the number of elements in this conjugacy class. Writing a sum over all elements of \( G \) as a sum over all conjugacy classes of the sums over all elements in a class, and taking into account the constancy of the
representations of finite groups, the orthogonality relation of characters \[ \text{III.24} \] can be recast in the equivalent form

\[
\frac{1}{|G|} \sum_{[g]} |[g]| \chi^{(\alpha)}([g]) \chi^{(\beta)}([g])^* = \delta_{\alpha\beta}, \tag{III.25}
\]

where we somewhat loosely use the same notations \( \chi^{(\alpha)}, \chi^{(\beta)} \) for functions whose argument is now a conjugacy class \([g]\).

Let \( r \) be the number of conjugacy classes in \( G \). Viewing \( |[g]|^{1/2} \chi^{(\alpha)}([g]) \) and \( |[g]|^{1/2} \chi^{(\beta)}([g]) \) as \( r \)-dimensional vectors, Eq. \[ \text{III.25} \] states that these are orthogonal if \( \alpha \neq \beta \), i.e. for inequivalent irreducible representations. Repeating a similar reasoning as in Sec. \[ \text{III.3.1} \] where we now use the fact that the pairwise orthogonal vectors \( |[g]|^{1/2} \chi^{(\alpha)}([g]) \) belong to an \( r \)-dimensional vector space, we can conclude that there exists at most \( r \) inequivalent irreducible representations of \( G \). In fact, one can even show the following:

**Theorem III.26.** The number of inequivalent irreducible representations of a finite group \( G \) is equal to the number of conjugacy classes in \( G \).

**Examples:**

* Since the dihedral group \( D_3 \) contains 3 conjugacy classes, the three irreducible representations introduced in Sec. \[ \text{II.3.2} \] constitute a whole set of inequivalent irreducible representations of the group.

* In the case of an Abelian group \( G \), there are as many conjugacy classes as elements in the group. According to theorem \[ \text{III.26} \], there are thus \( |G| \) inequivalent irreducible representations of \( G \). To fulfill the inequality \[ \text{III.13} \], these irreps. must all be one-dimensional, which provides a further demonstration of the corollary \[ \text{II.35} \].

### III.3.2 c Character table

To characterize the irreducible representations of a finite group \( G \), one uses a *character table*, which lists the characters of the various (classes of) irreducible representations. More precisely, the columns of the table correspond to the different conjugacy classes of the group, while on each row stands a different irreducible representation:

<table>
<thead>
<tr>
<th></th>
<th>( [g_1] )</th>
<th>( [g_2] )</th>
<th>( \ldots )</th>
<th>( [g_r] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^{(1)} )</td>
<td>( \chi^{(1)}(g_1) )</td>
<td>( \chi^{(1)}(g_2) )</td>
<td>( \ldots )</td>
<td>( \chi^{(1)}(g_r) )</td>
</tr>
<tr>
<td>( \chi^{(2)} )</td>
<td>( \chi^{(2)}(g_1) )</td>
<td>( \chi^{(2)}(g_2) )</td>
<td>( \ldots )</td>
<td>( \chi^{(2)}(g_r) )</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( \chi^{(r)} )</td>
<td>( \chi^{(r)}(g_1) )</td>
<td>( \chi^{(r)}(g_2) )</td>
<td>( \ldots )</td>
<td>( \chi^{(r)}(g_r) )</td>
</tr>
</tbody>
</table>

**Table III.1** – Character table of a group.

For instance, Table \[ \text{III.2} \] shows a character table for the group \( D_3 \), using the irreducible representations \( \xi^{(1)}, \xi^{(e)}, \xi^{(b)} \) introduced in Sec. \[ \text{II.3.2} \]. In the first row, \( [(A B)] \) stands for the conjugacy class of the reflexion, consisting of the three elements \( (A B), (A C), \) and \( (B C) \); in turn, \( [(A B C)] \) is the conjugacy class of the rotations through \( 2\pi/3 \) or \( 4\pi/3 \), i.e. \( [(A B C)] = \{(A B C), (A C B)\} \).

When constructing the table \[ \text{III.2} \] we already knew irreducible representations of \( D_3 \) and thus we only had to compute the respective traces to find the table entries. Even without explicitly knowing irreducible representations, one can often still determine the character table of a finite group \( |G| \), by exhausting the available information:
III.3 Irreducible representations of a finite group

<table>
<thead>
<tr>
<th></th>
<th>[Id]</th>
<th>[(A B)]</th>
<th>[(A B C)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^{(1)}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi^{(e)}$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi^{(s)}$</td>
<td>2</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table III.2 – Character table of the dihedral group $D_3$.

i. The number of rows and columns of the table is given by theorem (III.26), i.e. it is the number of conjugacy classes of the group.

ii. The (squared) dimensions of inequivalent irreducible representations $\hat{\rho}^{(\alpha)}$ of the group should obey Eq. (III.14). (16)

iii. Assuming that the first column of the table corresponds to the conjugacy class of the identity element $e$ of $G$, then the entries in that column are precisely the representation dimensions $\dim \rho^{(\alpha)}$ since $\chi^{(\alpha)}(e) = \text{Tr} \hat{\rho}^{(\alpha)} = \dim \rho^{(\alpha)}$.

iv. There is always at least a one-dimensional irreducible representation, namely the trivial representation. All entries in the corresponding line — which it is convenient to list first — are simply 1.

v. If there are further one-dimensional irreducible representations, their respective characters obey property (III.23).

vi. The orthogonality relation (III.25) should be fulfilled for every pair of representations, i.e. every pair of rows of the table.

Let us rediscover the character table of $D_3$ by applying this set of recipes. We assume that a preliminary study of the group has concluded to the existence of $r = 3$ conjugacy classes [Id], [(A B)], and [(A B C)], with respectively 1, 3, and 2 elements. According to point i., the character table should have 3 rows and 3 columns.

Denoting $d_\alpha$ the representation dimensions, these should according to point ii. be a solution of the equation $(d_1)^2 + (d_2)^2 + (d_3)^2 = |D_3| = 6$: up to a relabeling of the irreducible representations, the only possible solution is $d_1 = 1 = d_2 = 1$, $d_3 = 2$.

Invoking then points iii. and iv., we can already construct the following table, in which four entries are still missing.

<table>
<thead>
<tr>
<th></th>
<th>[Id]</th>
<th>[(A B)]</th>
<th>[(A B C)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^{(1)}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi^{(2)}$</td>
<td>1</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\chi^{(3)}$</td>
<td>2</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

The second line can be filled with the help of point v.: noting that two successive applications of a symmetry gives the identity, e.g. $(A B)^2 = \text{Id}$, property (III.23) gives $\chi^{(2)}((A B))^2 = \chi^{(2)}(\text{Id}) = 1$, i.e. $\chi^{(2)}((A B)) = \pm 1$. In turn, one checks that the composition of a reflexion and a rotation gives a reflexion with respect to an other axis, e.g. $(A B)(A B C) = (B C)$, which gives

\[ \sum_{\alpha} (d_\alpha)^2 = n \] such an equation, where all $d_\alpha$ and $n$ are (positive) integers is called a Diophantine equation.
\[ \chi^2((AB))\chi^2((ABC)) = \chi^2((BC)). \]

From \( \chi^2((AB)) = \chi^2((BC)) \), both reflexions being in the same conjugacy class, and the fact that this character is either 1 or \(-1\), i.e. non-zero, one deduces \( \chi^2((ABC)) = 1 \). This at once leaves only the possibility \( \chi^2((AB)) = -1 \), otherwise the first and second lines of the table would be identical:

<table>
<thead>
<tr>
<th></th>
<th>[Id]</th>
<th>[(AB)]</th>
<th>[(ABC)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^1 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>( \chi^3 )</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The third line can be dealt with entirely by using the orthogonality relation (III.25), as indicated in point vi. More precisely, one must have

\[
||\text{Id}||\chi^3(\text{Id})\chi^3(\text{Id})^* + ||(AB)||\chi^3((AB))\chi^3((AB))^* + ||(AB C)||\chi^3((AB C))\chi^3((AB C))^* = 0
\]

for \( \beta = 1 \) or 2, i.e. using \( ||\text{Id}|| = 1 \), \( ||(AB)|| = 3 \), \( ||(AB C)|| = 2 \) and the already known entries of the first column

\[ 2 + 3\chi^3((AB))\chi^3((AB))^* + 2\chi^3((AB C))\chi^3((AB C))^* = 0. \]

Denoting \( x = \chi^3((AB)) \), \( y = \chi^3((ABC)) \) the two missing characters, one obtains a system of two linear equations with two unknowns:

\[
\begin{cases}
2 + 3x + 2y = 0 \\
2 - 3x + 2y = 0.
\end{cases}
\]

The obvious solution is \( x = 0, y = -1 \), leading eventually to the table II.2.

**III.4 Reduction of a representation**

**III.4.1 Decomposition of a representation into irreducible representations**

Let \( \{ \hat{D}^{(\alpha)} \} \) be a set of inequivalent unitary irreducible representations of a finite group \( G \) and \( \hat{D} \) a finite-dimensional representation of \( G \). Following Maschke’s theorem III.5, \( \hat{D} \) is fully reducible, and the equivalence class of each irreducible representations in its decomposition contains one of the \( \hat{D}^{(\alpha)} \) (theorem III.2). Accordingly, one writes

\[ \hat{D} = \bigoplus_{\text{irreps. } \alpha} a_{\alpha} \hat{D}^{(\alpha)} \quad \text{with } a_{\alpha} \in \mathbb{N}, \quad (III.27) \]

where the integer coefficients \( a_{\alpha} \) signal the possible multiple appearance of the irreducible representation \( \hat{D}^{(\alpha)} \) in the direct sum, as in Eq. (II.21).

**Remark:** The equal sign in Eq. (III.27) in fact stands for an equivalence. Similarly, this equivalence is used to (re)order the irreducible representations according to the (here unspecified) index \( \alpha \).

Let \( \chi^{(\alpha)} \) denote the character of the irreducible representation \( \hat{D}^{(\alpha)} \) and \( \chi_{\hat{D}} \) that of the representation \( \hat{D} \). Applying both sides of Eq. (III.27) to an element \( g \) of \( G \), which yields an equality between operators, and building the trace of the latter, one obtains thanks to property (III.19)
\[ \chi_{\mathcal{D}}(g) = \sum_{\text{irreps. } \alpha} a_{\alpha} \chi^{(\alpha)}(g) \quad \text{for all } g \in \mathcal{G}. \] (III.28)

That is, a compound character is a linear combination with positive integer coefficients of simple characters.

This identity is the key to finding the coefficients \( \{a_{\alpha}\} \) of the decomposition. Invoking the orthogonality relation (III.24) of the irreducible characters, one has

\[ a_{\alpha} = (\chi^{(\alpha)}, \chi_{\mathcal{D}})_G = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi_{\mathcal{D}}(g) \chi^{(\alpha)}(g)^*. \] (III.29)

Proof: denoting \( \beta \) the dummy summation index for irreps. in Eq. (III.28) and using the linearity of the scalar product (III.11) in its second argument, one has

\[ (\chi^{(\alpha)}, \chi_{\mathcal{D}})_G = \sum_{\text{irreps. } \beta} a_{\beta} (\chi^{(\alpha)}, \chi^{(\beta)})_G = \sum_{\text{irreps. } \beta} a_{\beta} \delta_{\alpha\beta} = a_{\alpha}. \]

**Theorem III.30** (Criterion for irreducibility). Let \( \chi_{\mathcal{D}} \) denote the character of a representation \( \hat{\mathcal{D}} \) of a finite group \( \mathcal{G} \). Then \( \hat{\mathcal{D}} \) is irreducible if and only if the scalar product \( (\chi_{\mathcal{D}}, \chi_{\mathcal{D}})_G \) equals 1.

Proof: Let \( \{ \hat{\mathcal{D}}^{(\alpha)} \} \) be a set of inequivalent unitary irreducible representations of \( \mathcal{G} \) and

\[ \hat{\mathcal{D}} = \bigoplus_{\text{irreps. } \alpha} a_{\alpha} \hat{\mathcal{D}}^{(\alpha)} \]

with positive integer coefficients \( a_{\alpha} \) be the decomposition of \( \hat{\mathcal{D}} \) into irreps. Using Eq. (III.28), the sesquilinearity of the scalar product (III.11), and the orthonormality relation (III.24), one has

\[ (\chi_{\mathcal{D}}, \chi_{\mathcal{D}})_G = \sum_{\text{irreps. } \alpha, \beta} a_{\alpha}^{*} a_{\beta} (\chi^{(\alpha)}, \chi^{(\beta)})_G = \sum_{\text{irreps. } \alpha} a_{\alpha}^{*} a_{\alpha} \delta_{\alpha\beta} = \sum_{\text{irreps. } \alpha} |a_{\alpha}|^2. \]

Since every \( a_{\alpha} \) is a positive integer, this also holds for \( |a_{\alpha}|^2 \), and the sum on the right hand side equals 1 if and only if all coefficients \( a_{\alpha} \) vanish but a single one, which has to be equal one — i.e. precisely if \( \hat{\mathcal{D}} \) is an irreducible representation.

### III.4.2 Regular representation

#### III.4.2 a Group algebra

**Algebra**

Let us first define a new algebraic structure, the *algebra*, which we shall need both in the present section and more generally in the study of continuous groups in Chapter V.

**Definition III.31.** An *algebra* is a vector space \( \mathcal{A} \) on which a bilinear operation \( \ast \) mapping \( \mathcal{A} \times \mathcal{A} \) into \( \mathcal{A} \) is defined with the following properties:

- for all \( x, y, z \in \mathcal{A} \), \( (x + y) \ast z = x \ast z + y \ast z \); \hfill (A1)
- for all \( x, y, z \in \mathcal{A} \), \( x \ast (y + z) = x \ast y + x \ast z \); \hfill (A2)
- for all \( x, y \in \mathcal{A} \), \( \lambda \in \mathbb{K} \), \( (\lambda x) \ast y = \lambda (x \ast y) = x \ast (\lambda y) \), \hfill (A3)

where \( \mathbb{K} \) is the base field of the vector space. One often specifies “algebra over the field \( \mathbb{K} \”).

**Remarks:**

* The internal law \( \ast \) is often referred to as the “vector product” or the “multiplication” on the algebra.
Thanks to the bilinearity, it is sufficient to know the products \( e_i \star e_j \) of the vectors of a basis \( \{e_i\} \) of the algebra to fully determine the multiplication:

\[
\text{if } x = \sum_i x_i e_i, \ y = \sum_j y_j e_j \text{ then } x \star y = \sum_{i,j} x_i y_j e_i \star e_j. \tag{III.32}
\]

Denoting \( 0_A \) the zero vector of the algebra \( A \), one deduces from Eq. (A3) that \( 0_A \) is an absorbing element for \( \star \), i.e.

\[
\text{for all } x \in A, \ x \star 0_A = 0_A \star x = 0_A. \tag{III.33}
\]

**Definition III.34.** An algebra \( A \) is called associative if for all \( x, y, z \in A \), \( x \star (y \star z) = (x \star y) \star z \).

**Definition III.35.** An algebra \( A \) is called commutative if for all \( x, y \in A \), \( x \star y = y \star x \).

**Definition III.36.** An algebra is called unital (or unitary) if it has an identity element with respect to the operation \( \star \).

**Examples:**

* The set \( M(n, K) \) of \( n \times n \) matrices with entries in \( K \), with the usual matrix product as multiplication, is a unital associative algebra.

* The set \( \mathbb{R}^3 \) of three-dimensional vectors, with the usual cross product \( \times \) as vector product, is an algebra over \( \mathbb{R} \), which neither associative nor commutative (\( \vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \! \)), nor has an identity element.

* The linear operators on a Hilbert space, with the commutator as multiplication, form an algebra.

**Definition III.37.** The group algebra of a finite\(^{(17)} \) group \( G \) over a field \( K \) is the algebra, denoted \( K G \), consisting of elements of the form

\[
x = \sum_{g \in G} x_g g, \text{ with } x_g \in K, \tag{III.38a}
\]

with the addition and multiplication respectively defined for \( x = \sum_{g \in G} x_g g \) and \( y = \sum_{g \in G} y_g g \) by

\[
x + y = \sum_{g \in G} (x_g + y_g) g \quad \text{and} \quad x \star y = \sum_{h \in G, g, g' \in G, gg' = h} (x_g y_{g'}) h. \tag{III.38b}
\]

**Remarks:**

* Equation (III.38a) is to be understood as meaning that the group elements form a basis of the group algebra. Thus if \( G \) is finite, then \( K G \) is of dimension \( |G| \).

* In turn, the vector product \( \star \) defined in Eq. (III.38b) is induced by the group law: if the latter reads \( gg' = h \) for \( g, g', h \in G \), then in \( K G \) one has \( g \star g' = h \). Accordingly, one quickly drops the symbol \( \star \) when denoting the (vector) product of elements of the group algebra.

**Property III.39.** The group algebra \( K G \) is associative.

**Property III.40.** The identity element \( e \) of the group \( G \) is also an identity element of the group algebra \( K G \).

**Property III.41.** If the group \( G \) is Abelian, then the group algebra \( K G \) is commutative.

These results follow at once from the properties of the group law.

\(^{(17)}\) If the group is infinite, the sums in Eqs. (III.38a)–(III.38b) should run over only a finite number of terms.
### III.4 Reduction of a representation

#### III.4.2 Regular representation and its reduction

**Theorem & Definition III.42.** Let $\mathcal{G}$ be a finite group and $K$ a field. The mapping $\hat{\mathcal{D}}^{(r)} : \mathcal{G} \rightarrow \text{Aut}(K\mathcal{G})$ which to an element $g \in \mathcal{G}$ associates the linear operator $\hat{\mathcal{D}}^{(r)}(g)$ on the group algebra $K\mathcal{G}$ such that for all $h \in \mathcal{G} \subset K\mathcal{G}$

\[
\hat{\mathcal{D}}^{(r)}(g)(h) = gh
\]

is a representation of $\mathcal{G}$ on $K\mathcal{G}$, called the (left) regular representation.

Denoting $g_1, \ldots, g_n$ the elements of $\mathcal{G}$, where $n = |\mathcal{G}|$, the matrix of the operator $\hat{\mathcal{D}}^{(r)}(g)$ in the basis $\{g_i\}_{1 \leq i \leq n}$ of the group algebra is such that

\[
\begin{pmatrix}
\mathcal{D}^{(r)}(g) & g_1 \\
& \vdots \\
& g_n \\
\end{pmatrix}
= 
\begin{pmatrix}
g_{g_1} \\
\vdots \\
g_{g_n}
\end{pmatrix}.
\]  

(III.43)

Since $g$ is one of the basis vectors, the matrix $\mathcal{D}^{(r)}(g)$ contains a single 1 and $n-1$ times 0 on each line and on each column. More precisely, the matrix representing the group identity element $e$ is the $|\mathcal{G}|$-dimensional identity matrix, $\mathcal{D}^{(r)}(e) = I_{|\mathcal{G}|}$, as is always the case; while for $g \neq e$, the mapping $h \rightarrow gh$ is a derangement of the group elements, leaving none of them invariant, so that the matrix $\mathcal{D}^{(r)}(g)$ has only 0 on its diagonal.

Accordingly, denoting $\chi^{(r)}$ the character of the regular representation, the latter obeys

\[
\begin{align*}
\chi^{(r)}(e) &= |\mathcal{G}| & (III.44a) \\
\chi^{(r)}(g) &= 0 \quad \text{for} \quad g \neq e. & (III.44b)
\end{align*}
\]

Let $\{\hat{\mathcal{D}}^{(\alpha)}\}$ be a set of inequivalent unitary irreducible representations of $\mathcal{G}$ on respective vector spaces $\{\mathcal{V}^{(\alpha)}\}$, with characters $\{\chi^{(\alpha)}\}$. The reduction (III.27) of the regular representation reads

\[
\hat{\mathcal{D}}^{(r)} = \bigoplus_{\text{irreps. } \alpha} a^{(r)}_{\alpha} \hat{\mathcal{D}}^{(\alpha)}
\]

with integer coefficients $a^{(r)}_{\alpha}$ given by Eq. (III.29), namely

\[
a^{(r)}_{\alpha} = \left(\chi^{(\alpha)}, \chi^{(r)}\right)_g = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi^{(r)}(g) \chi^{(\alpha)}(g)^*.
\]

Using Eqs. (III.44), the sum on the right hand side of the previous equation actually only involves a single non-zero summand for $g = e$, namely $\chi^{(r)}(e) \chi^{(\alpha)}(e)^* = |\mathcal{G}| \chi^{(\alpha)}(e)^* = |\mathcal{G}| (\dim \mathcal{V}^{(\alpha)})^*$, i.e.

\[
a^{(r)}_{\alpha} = \dim \mathcal{V}^{(\alpha)}.
\]

We thus have proven the following theorem:

**Theorem III.45.** The regular representation $\hat{\mathcal{D}}^{(r)}$ of a finite group $\mathcal{G}$ contains each irreducible representation $\hat{\mathcal{D}}^{(\alpha)}$ of the group\(^{18}\) with a multiplicity equal to the dimension of $\hat{\mathcal{D}}^{(\alpha)}$:

\[
\hat{\mathcal{D}}^{(r)} = \bigoplus_{\text{irreps. } \alpha} (\dim \mathcal{V}^{(\alpha)}) \hat{\mathcal{D}}^{(\alpha)}.
\]  

(III.45)

Taking the trace of the reduction (III.45) yields

\[
\chi^{(r)} = \sum_{\text{irreps. } \alpha} (\dim \mathcal{V}^{(\alpha)}) \chi^{(\alpha)}
\]

(III.46)

for the characters. Applying this equality between functions to the group identity element $e$, using

\(^{18}\)More accurately, it contains an element of each equivalence class of irreducible representations.
the fact that the character of \( e \) is the dimension of the representation space, gives

\[
|\mathcal{G}| = \sum_{\text{irreps. } \alpha} \left( \dim \nu^{(\alpha)} \right)^2,
\]

which proves theorem \textbf{III.14}.

Let us now sketch how one can find subspaces of the group algebra \( \mathbb{K}\mathcal{G} \) which are invariant under the regular representation, thus reducing \( \hat{\mathcal{G}}^{(r)} \). More precisely, we search for “minimal” invariant subspaces \( \mathcal{I}_i \), i.e. such that they admit no proper subspace which is itself invariant. The ideas introduced here will be illustrated in Sec. \textbf{III.5}.

Let \( g \in \mathcal{G} \) and \( y \in \mathcal{I}_i \). Decomposing \( y \) on the basis of \( \mathbb{K}\mathcal{G} \) consisting of the group elements, one finds \( \hat{\mathcal{G}}^{(r)}(g)(y) = g \ast y \). The invariance of \( \mathcal{I}_i \) under \( \hat{\mathcal{G}}^{(r)} \) thus means that for all \( g \in \mathcal{G} \) and all \( y \in \mathcal{I}_i \), \( g \ast y \in \mathcal{I}_i \). Since this holds for every group element \( g \) and given that \( \mathcal{I}_i \) is a vector space, any linear combination of vectors \( g \ast y \) with \( g \in \mathcal{G} \) is again in \( \mathcal{I}_i \), i.e. \( x \ast y \in \mathcal{I}_i \), which one also expresses as \( x \ast \mathcal{I}_i \subset \mathcal{I}_i \) for all \( x \in \mathbb{K}\mathcal{G} \).

Assume that one has found elements \( \{e_i\} \) of the group algebra \( \mathbb{K}\mathcal{G} \) fulfilling the following set of requirements:

i. for all \( i \), \( e_i \ast e_i = e_i \), which defines an idempotent element;

ii. for all \( i \neq j \), \( e_i \ast e_j = 0 \) where 0 denotes the zero vector of \( \mathbb{K}\mathcal{G} \);

iii. each \( e_i \) is primitive, i.e. cannot be expressed as the sum \( e_i' + e_i'' \) of two idempotents obeying requirement ii.;

iv. denoting \( \mathcal{I}_i = \mathbb{K}\mathcal{G} \ast e_i = \{x \ast e_i, x \in \mathbb{K}\mathcal{G}\} \), the various \( \mathcal{I}_i \) are proper subspaces of \( \mathbb{K}\mathcal{G} \) whose direct sum \( \bigoplus \mathcal{I}_i \) equals \( \mathbb{K}\mathcal{G} \). Then one easily checks that the subspaces \( \mathcal{I}_i \) thus defined are invariant under the regular representation and actually provide a reduction in irreducible representations.

The key idea is to recognize that the identity element \( e \) of the group can be written in the form \( e = e_1 + e_2 + \cdots \) and to use \( x \ast e = x \) for all \( x \in \mathbb{K}\mathcal{G} \).

The reduction of the regular representation can thus be brought down to the search for primitive idempotents \( \{e_i\} \) of the group algebra, which will “generate” the invariant subspaces \( \mathcal{I}_i \).

\textbf{Remarks:}

* A single primitive idempotent can generate a subspace \( \mathcal{I}_i = \mathbb{K}\mathcal{G} \ast e_i \) of dimension greater than one. That is, there will in general be less than \( |\mathcal{G}| \) primitive idempotents, so that they do not constitute a basis of \( \mathbb{K}\mathcal{G} \).

* The notation \( e_i \) is traditional for the primitive idempotents defined here, as is the notation \( e_i \) for basis vectors of a vector space! In these lecture notes the distinction is made by denoting the primitive idempotents in italic and the basis vectors in roman font, but the author can guarantee that this is not a universal convention.

* An equivalent formulation to that adopted here is to talk of projectors \( \{p_i\} \) on mutually orthogonal subspaces of the group algebra: a projector is indeed idempotent \( (p_i \circ p_i = p_i) \) and orthogonal projectors obey \( p_i \circ p_j = p_j \circ p_i = 0 \), where 0 now denotes the zero operator, which is clearly reminiscent of requirements i. and ii.

* In practical cases, it might be less time-consuming to search for primitive idempotents “up to a factor”, i.e. such that \( e_i \ast e_i = \lambda_i e_i \) with \( \lambda_i \in \mathbb{K}, \lambda_i \neq 0_{\mathbb{K}} \), since they generate the same invariant subspaces.

\(^{(19)}A\) subset \( \mathcal{I} \) of an algebra \( \mathcal{A} \) obeying \( x \ast \mathcal{I} \subset \mathcal{I} \) for all \( x \in \mathcal{A} \) is called a left ideal of \( \mathcal{A} \), hence the notation \( \mathcal{I}_i \).
III.4.3 Reduction of a tensor-product representation

Let \( \hat{\mathcal{G}} \) and \( \hat{\mathcal{G}}' \) be two representations of a finite group \( \mathcal{G} \) with respective characters \( \chi \) and \( \chi' \). The tensor-product representation \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \) was introduced in §II.3.4b.

**Theorem III.47.** The tensor-product representation \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \) is a representation of the direct-product group \( \mathcal{G} \times \mathcal{G} \). If \( \hat{\mathcal{G}} \) and \( \hat{\mathcal{G}}' \) are irreducible, so is \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \).

Since \( \mathcal{G} \) is (isomorphic to) a subgroup of \( \mathcal{G} \times \mathcal{G} \), the tensor-product representation \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \) is also a representation of \( \mathcal{G} \). Even if \( \hat{\mathcal{G}} \) and \( \hat{\mathcal{G}}' \) are irreducible, yet \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \) can in general be reducible (although theorem III.50 below provides a generic exception). Introducing a set \( \{ \hat{\mathcal{G}}(a) \} \) of inequivalent unitary irreducible representations of \( \mathcal{G} \), one may thus write [cf. Eq. (III.27)]

\[
\hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' = \bigoplus_{\text{irreps. } \alpha} a_{\alpha} \hat{\mathcal{G}}(a) \quad \text{with } a_{\alpha} \in \mathbb{N}. \tag{III.48a}
\]

This decomposition is called the Clebsch–Gordan series of the tensor-product representations. Following Eq. (III.29), the integer coefficients \( a_{\alpha} \) are given by

\[
a_{\alpha} = \langle \chi^{(\alpha)}, \chi_{\hat{\mathcal{G}} \otimes \hat{\mathcal{G}}'} \rangle_{\mathcal{G}} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi_{\hat{\mathcal{G}} \otimes \hat{\mathcal{G}}'}(g) \chi(\alpha)(g)^*,
\]

where \( \chi_{\hat{\mathcal{G}} \otimes \hat{\mathcal{G}}'} \) denotes the character of \( \hat{\mathcal{G}} \otimes \hat{\mathcal{G}}' \). Under consideration of Eq. (III.20) comes

\[
a_{\alpha} = \langle \chi^{(\alpha)}, \chi_{\hat{\mathcal{G}} \otimes \hat{\mathcal{G}}'} \rangle_{\mathcal{G}} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi_{\hat{\mathcal{G}}'}(g) \chi_{\hat{\mathcal{G}}}(g) \chi(\alpha)(g)^*. \tag{III.48b}
\]

**Example III.49.** Consider once again the dihedral group \( D_3 \). Let \( \hat{\mathcal{G}}(s \otimes s) \equiv \hat{\mathcal{G}}(s) \otimes \hat{\mathcal{G}}(s) \) denote the tensor product of the two-dimensional, irreducible standard representation with itself, and \( \chi^{(s \otimes s)} \) the corresponding character. Since \( \chi^{(s \otimes s)} = \chi^{(s)} \chi^{(s)} \) [Eq. (III.20)], one finds the following characters:

<table>
<thead>
<tr>
<th></th>
<th>[Id]</th>
<th>[(A B)]</th>
<th>[(A B C)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^{(s \otimes s)} )</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Comparing with the character table of the group, one finds \( \chi^{(s \otimes s)} = \chi^{(1)} + \chi^{(e)} + \chi^{(s)} \), i.e. \( \hat{\mathcal{G}}^{(s \otimes s)} = \hat{\mathcal{G}}^{(1)} \oplus \hat{\mathcal{G}}^{(e)} \oplus \hat{\mathcal{G}}^{(s)} \). The other possible tensor products of irreducible representations of \( D_3 \), involving either \( \hat{\mathcal{G}}^{(1)} \) or \( \hat{\mathcal{G}}^{(e)} \), are taken care of by the following theorem.

**Theorem III.50.** The tensor product of an irreducible representation of a group \( \mathcal{G} \) with a representation of dimension 1 is irreducible.

**Clebsch–Gordan coefficients**

Let \( \{ \hat{\mathcal{G}}(\gamma) \} \) be a set of inequivalent irreducible representations of a finite group \( \mathcal{G} \). For each \( \gamma \), we denote \( d(\gamma) \) the dimension of \( \hat{\mathcal{G}}(\gamma) \).

Given two irreducible representations \( \hat{\mathcal{G}}^{(\alpha)} \) and \( \hat{\mathcal{G}}^{(\beta)} \) on respective vector spaces \( \mathcal{V}^{(\alpha)} \) and \( \mathcal{V}^{(\beta)} \) with bases \( \{ \psi_i^{(\alpha)} \}_{i=1,\ldots,d(\alpha)} \) and \( \{ \phi_j^{(\beta)} \}_{j=1,\ldots,d(\beta)} \), let

\[
\hat{\mathcal{G}}^{(\alpha)} \otimes \hat{\mathcal{G}}^{(\beta)} = \bigoplus_{\text{irreps. } \gamma} a_{\gamma}^{\alpha\beta} \hat{\mathcal{G}}(\gamma) \quad \text{with } a_{\gamma}^{\alpha\beta} \in \mathbb{N}, \tag{III.51}
\]

be the decomposition of the tensor-product representation in irreducible representations.

The vectors \( \{ \psi_i^{(\alpha)} \otimes \phi_j^{(\beta)} \} \) form a basis of the representation space \( \mathcal{V}^{(\alpha)} \otimes \mathcal{V}^{(\beta)} \) of \( \hat{\mathcal{G}}^{(\alpha)} \otimes \hat{\mathcal{G}}^{(\beta)} \).

---

(m) A. Clebsch, 1833–1872

(n) P. Gordan, 1837–1912
Since the latter is decomposable, one can find an alternative basis \( \{ \Psi^{(\gamma),n_\gamma}_k \} \) that is better “adapted” to the decomposition (III.51) in the following sense. The set of new basis vectors can be partitioned into disjoint subsets, labeled by \( \gamma \) and \( n_\gamma \in \{1, \ldots, a^\gamma \} \) — if the corresponding integer \( a^\gamma \) in the Clebsch–Gordan series is at least equal to 2, otherwise \( n_\gamma \) is irrelevant. For given \( \gamma \) and \( n_\gamma \), there are \( d(\gamma) \) basis vectors \( \Psi^{(\gamma),n_\gamma}_k \) with \( k \in \{1, \ldots, d(\gamma)\} \), which span a subspace of \( \mathcal{V}^{(\alpha)} \otimes \mathcal{V}^{(\beta)} \) invariant under (one of) the representation \( \mathcal{G}(\gamma) \) in the decomposition (III.51).

The basis change between \( \{ \psi^{(\alpha)}_i \otimes \phi^{(\beta)}_j \} \) and \( \{ \Psi^{(\gamma),n_\gamma}_k \} \) is of the form

\[
\Psi^{(\gamma),n_\gamma}_k = \sum_{i,j} C_{\alpha,\beta;i,j}^{\gamma,n_\gamma;k} \psi^{(\alpha)}_i \otimes \phi^{(\beta)}_j,
\]

where the sum over \( i \) resp. \( j \) runs from 1 to \( d(\alpha) \) resp. \( d(\beta) \). The (complex) coefficients \( C_{\alpha,\beta;i,j}^{\gamma,n_\gamma;k} \) are called Clebsch–Gordan coefficients. These are usually normalized to obey the relation

\[
\sum_{i,j} \left| C_{\alpha,\beta;i,j}^{\gamma,n_\gamma;k} \right|^2 = 1
\]

for all \( \gamma, n_\gamma \) and \( k \in \{1, \ldots, d(\gamma)\} \). This choice ensures that the basis \( \{ \Psi^{(\gamma),n_\gamma}_k \} \) is orthonormal for the “natural” scalar product on \( \mathcal{V}^{(\alpha)} \otimes \mathcal{V}^{(\beta)} \) if this is the case of the bases \( \{ \psi^{(\alpha)}_i \} \) and \( \{ \phi^{(\beta)}_j \} \).

**Definition III.54.** A group is called simply reducible if the integer coefficient \( a_\gamma^{\alpha,\beta} \) in the Clebsch–Gordan series (III.51) is either 0 or 1 for all \( \alpha, \beta, \gamma \), that is when the tensor product of any two irreducible representations decomposes into different irreducible representations.

An example is provided by the dihedral group \( D_3 \), as can be seen from example III.49. In the case of a simply reducible group, the label \( n_\gamma \) is unnecessary in Eqs. (III.52)–(III.53).

### III.5 Representations of the symmetric group \( S_n \)

Let us now investigate the representations of the group \( S_n \) of the permutations of \( n \in \mathbb{N}^* \) objects, using the ideas sketched in § III.4.2b. For definiteness, we shall assume that the group algebra, on which the regular representation operates, is a complex vector space, thus to be denoted \( \mathbb{C}S_n \). The multiplication on \( \mathbb{C}S_n \) is (built from) the composition of permutations: for \( \sigma, \tau \in S_n \), \( \tau \circ \sigma = \tau \circ \sigma \), which will be more shortly written \( \tau \sigma \).

Gathering first a few results obtained in previous sections, we know that the group \( S_n \) is of order \( |S_n| = n! \) (property I.19) and that its conjugacy classes are in one-to-one correspondence with the integer partitions of \( n \), which are conveniently depicted with the help of Young diagrams (§ I.4.3b).

#### III.5.1 Irreducible representations of \( S_n \)

According to § III.4.2b, the reduction of the regular representation of \( S_n \) can be done using primitive idempotents, i.e. adequate linear combinations of the group elements \( \sigma \in S_n \) that generate invariant subspaces of \( \mathbb{C}S_n \). In this specific case, the primitive idempotents are called Young elements. Two of them (for \( n \geq 2 \)) are easily found, namely

\[
\text{the symmetrizer: } \mathcal{J} = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma
\]

\[
\text{the antisymmetrizer: } \mathcal{A} = \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) \sigma,
\]

with \( \varepsilon(\sigma) \) the signature of the permutation \( \sigma \). Using the identities \( \tau \mathcal{J} = \mathcal{J} \) and \( \tau \mathcal{A} = \varepsilon(\tau) \mathcal{A} \) for all \( \tau \in S_n \), one indeed checks that \( \mathcal{J} \) and \( \mathcal{A} \) are idempotent and that their product is 0 (the zero

\( ^{20} \) The reduction of \( S_1 \) is straightforward: it has a single irrep., the trivial representation.)
vector of the group algebra). Additionally, one sees that both elements generate one-dimensional invariant subspaces, corresponding to the trivial and the signature (or alternating) representations, respectively. For $S_2$, these are the only two irreducible representations.

For $n \geq 3$, there are further irreducible representations, corresponding to other Young elements. We shall just state a few results regarding those, without proof, and illustrate them on the examples of $S_3$ or $S_4$\footnote{Note that the irreducible representations of $S_3$, which is isomorphic to $D_3$, were actually already introduced in Sec. II.3.2}. For that purpose, let us first introduce two definitions.

### III.5.1 a Young tableaux

**Definition III.56.** A Young tableau of order $n$ is a Young diagram with $n$ boxes in which the numbers $1, \ldots, n$ have been written in any arbitrary order.

**Definition III.57.** A Young tableau is called *standard* if the numbers occur in increasing order in each row and in each column.

For instance, the following are four Young tableaux of order 7 corresponding to the same Young diagram; only the third and fourth tableaux are standard:

\[
\begin{array}{c}
3 & 1 & 8 & 7 \\
4 & 2 \\
6
\end{array}, \quad
\begin{array}{c}
1 & 3 & 6 & 7 \\
4 & 5 \\
2
\end{array}, \quad
\begin{array}{c}
1 & 3 & 6 & 7 \\
2 & 4 \\
5
\end{array}, \quad
\begin{array}{c}
1 & 2 & 3 & 4 \\
5 & 6 \\
7
\end{array}.
\]

As further examples, we list in Fig. III.1 resp. III.2 all standard Young tableaux with 3 resp. 4 boxes, arranging them according to the Young diagrams.

**Figure III.1** – Standard Young tableaux for $S_3$.

\[
\begin{array}{c}
1 & 2 & 3 \\
3
\end{array}, \quad
\begin{array}{c}
1 & 2 & 3 \\
2
\end{array}, \quad
\begin{array}{c}
1 & 3 \\
2
\end{array}.
\]

To the Young diagrams (of any order $n$) with a single row or a single column, there corresponds only one standard Young tableau, respectively. In contrast, when there are at least two rows or columns in the diagram, there are several standard tableaux. To know the number of the latter without enumerating them, let us introduce the *hook length* of a given box in a Young diagram, which is the number of boxes (including itself) either directly to its right or directly below. We illustrate this definition on an example in Fig. III.3 where the number in each box of the right diagram is precisely the hook length of the box.

**Figure III.3** – Example of hook lengths (right) for the Young diagram depicted left.
Using the hook lengths, the number of standard Young tableaux for a given Young diagram of order \( n \) is given by the hook length formula:

\[
d_\alpha = \frac{n!}{\prod \text{(hook lengths)}},
\]

where the product in the denominator runs over all boxes of the diagram, while the subscript \( \alpha \) merely labels that diagram. For the example of Fig. [III.3] one would find \( 7!/(6 \cdot 4 \cdot 2 \cdot 1 \cdot 3 \cdot 1 \cdot 1) = 35 \), different standard Young tableaux, which the reader probably does not want to enumerate. One quickly checks that the formula (III.58) gives the right results for the diagrams with only one row or only one column, as well as for those presented in Figs. [III.1] [III.2].

### III.5.1b Young elements; irreducible representations of \( S_n \)

For each standard Young tableau of order \( n \), corresponding to a Young diagram \( (\alpha) \), one defines an element of the group algebra \( \mathbb{C}S_n \) as follows. Let \( \mathcal{P} \) resp. \( \mathcal{Q} \) denote the subset of \( S_n \) consisting of the permutations that leave every row resp. every line of the standard Young tableau invariant. Denoting

\[
\mathcal{P} \equiv \sum_{\sigma \in \mathcal{P}} \sigma \quad \text{and} \quad \mathcal{Q} \equiv \sum_{\sigma \in \mathcal{Q}} \varepsilon(\sigma) \sigma,
\]

both elements of \( \mathbb{C}S_n \), their product

\[
\mathcal{Y} = \mathcal{Q} \mathcal{P}
\]

is again an element of \( \mathbb{C}S_n \) which is a primitive idempotent up to a factor: \( \mathcal{Y} \) is a Young element. For instance, the standard Young diagram built from the Young diagram with only a single row resp. a single column leads to the symmetrizer (III.55a) resp. to the antisymmetrizer (III.55b).\

**Remark:** \( \mathcal{P} \) resp. \( \mathcal{Q} \) is sometimes called the row resp. column stabilizer of the standard Young tableau under consideration. Both are actually subgroups of \( S_n \).

To each Young diagram \( (\alpha) \) correspond \( d_\alpha \) standard tableaux, with \( d_\alpha \) given by Eq. (III.58), which lead to as many Young elements. In turn, each of these Young elements generates (by left multiplication with the group algebra) a subspace of dimension \( d_\alpha \) that is invariant under the regular representation. More precisely, each Young diagram \( (\alpha) \) leads to \( d_\alpha \) copies of the same irreducible representation, namely of a representation of dimension \( d_\alpha \) — which we shall naturally denote \( \hat{D}_\alpha \).

Going to the standard Young diagrams with 3 boxes as given in Fig. [III.1] this means we should find 4 Young elements \( \mathcal{Y}, \mathcal{Y}, \mathcal{Y}', \mathcal{A} \), respectively. \( \mathcal{Y} \) and \( \mathcal{A} \) generate two one-dimensional invariant spaces with different irreducible representations, which we already know. In turn, \( \mathcal{Y} \) and \( \mathcal{Y}' \) generate two invariant spaces of dimension 2, corresponding to two copies of the same irreducible representation — which is the standard representation of § II.3.2c.

Consider for example the standard Young diagram \( \begin{array}{c}
\mathcal{Y}
\end{array} \). The corresponding row stabilizer is \( \mathcal{P} = \{ \text{Id}, (1 \ 2) \} \) and the column stabilizer \( \mathcal{Q} = \{ \text{Id}, (1 \ 3) \} \), leading to

\[
\mathcal{P} = \text{Id} + (1 \ 2) \quad \text{and} \quad \mathcal{Q} = \text{Id} - (1 \ 3),
\]

leading to the Young element

\[
\mathcal{Y} = \text{Id} - (1 \ 3)[\text{Id} + (1 \ 2)],
\]

that is, after factorizing out and performing the products

\[
\mathcal{Y} = \text{Id} + (1 \ 2) - (1 \ 3) - (1 \ 2 \ 3).
\]

The left multiplication of \( \mathcal{Y} \) with each element of \( S_3 \) gives (somewhat tediously)

\[
\begin{align*}
\text{Id} \mathcal{Y} &= \mathcal{Y} \\
(1 \ 2 \ 3) \mathcal{Y} &= (1 \ 2 \ 3) + (1 \ 3) - (2 \ 3) - (1 \ 3 \ 2) \\
(1 \ 3 \ 2) \mathcal{Y} &= (1 \ 3 \ 2) + (2 \ 3) - (1 \ 2) - \text{Id} = -\mathcal{Y} - (1 \ 2 \ 3) \mathcal{Y} \\
(1 \ 2) \mathcal{Y} &= (1 \ 2) + \text{Id} - (1 \ 3 \ 2) - (2 \ 3) = \mathcal{Y} + (1 \ 2 \ 3) \mathcal{Y} \\
(2 \ 3) \mathcal{Y} &= (2 \ 3) + (1 \ 3 \ 2) - (1 \ 2 \ 3) - (1 \ 3) = -(1 \ 2 \ 3) \mathcal{Y} \\
(1 \ 3) \mathcal{Y} &= (1 \ 3) + (1 \ 2 \ 3) - \text{Id} - (1 \ 2) = -\mathcal{Y}.
\end{align*}
\]
All 6 products can thus be expressed as linear combinations of $S$ and $(1\ 2\ 3)S$, which shows that the subspace generated by multiplying $S$ left with $CS_3$ is of dimension 2.

For every element $\sigma$ of $S_3$, one can also compute the product $\sigma(1\ 2\ 3)S$: using Eqs. (III.60), one can then construct a $2 \times 2$-matrix representing $\sigma$ in the basis consisting of $S$ and $(1\ 2\ 3)S$, i.e. precisely a two-dimensional representation of $S_3$.

Eventually, note that Eqs. (III.60) allow one to compute $S^2 = 3S$, which shows that $S$ is indeed idempotent up to a factor, as well as $S^2 = AS = 0$.

Regarding $S_4$, the standard Young diagrams depicted in Fig. III.2 lead to two representations of dimension 1 — the trivial and alternating representations —, two inequivalent irreducible representations of dimension 3 — which are both contained three times in the regular representation —, and one irreducible representation of dimension 2, contained twice in the regular representation.

### III.5.1 Conjugate Young diagrams

**Definition III.61.** Given a Young diagram of order $n$, the conjugate diagram (or transposed diagram) is that obtained by exchanging the roles of the rows and columns. A self-conjugate diagram is then a Young diagram which coincides with its conjugate diagram.

**Examples:**

* Considering the Young diagrams of order 4, the diagrams

$$
\begin{array}{cc}
\begin{array}{cc}
\bullet & \bullet \\
\bullet & \bullet
\end{array}
\end{array}
$$

and

$$
\begin{array}{cc}
\begin{array}{cc}
\bullet & \bullet \\
\bullet & \bullet
\end{array}
\end{array}
$$

resp.

$$
\begin{array}{cc}
\begin{array}{cc}
\bullet & \bullet \\
\bullet & \bullet
\end{array}
\end{array}
$$

and

$$
\begin{array}{cc}
\begin{array}{cc}
\bullet & \bullet \\
\bullet & \bullet
\end{array}
\end{array}
$$

form two pairs of conjugate diagrams, while the diagram

$$
\begin{array}{cc}
\begin{array}{cc}
\bullet & \bullet \\
\bullet & \bullet
\end{array}
\end{array}
$$

is self-conjugate.

* More generally, for any $n$ the diagrams consisting of a single row and that consisting of a single column, corresponding respectively to the Young symmetrizer and antisymmetrizer, are conjugate to each other.

Coming to the irreducible representations associated to the standard Young tableaux corresponding to conjugate diagrams — which according to the hook length formula will have the same dimension —, there exist a handful of results, two of which we now quote.

**Theorem III.62.** Let $\hat{D}^{(\alpha)}$ be the irreducible representation of $S_n$ associated with a given Young diagram of order $n \geq 2$. The irreducible representation associated with the conjugate Young diagram is obtained by multiplying $\hat{D}^{(\alpha)}$ with the alternating representation.

The fact that the Young symmetrizer generates the trivial representation and the Young antisymmetrizer the alternating representation is a special case of that theorem.

**Corollary III.63.** In a self-conjugate irreducible representation, the characters of the odd permutations are zero.

### III.5.2 Product of irreducible representations of the symmetric group

Consider two irreducible representations $\hat{D}^{(\alpha)}_{(n)}$ of $S_n$ and $\hat{D}^{(\beta)}_{(p)}$ of $S_p$, where $n$ and $p$ are not necessarily equal. The corresponding Young diagrams will be denoted $(\alpha)$ and $(\beta)$. One can define an “outer product” of these representations, or equivalently of the Young diagrams, which leads to a representation of $S_{n+p}$ that will in general be reducible:

$$
\hat{D}^{(\alpha)}_{(n)} \boxtimes \hat{D}^{(\beta)}_{(p)} = \bigoplus_{\text{irreps. } \gamma} c_{\alpha\beta}^{\gamma} \hat{D}^{(\gamma)}_{(n+p)},
$$

(III.64)
where the $\hat{\mathcal{D}}^{(\gamma)}_n$ are now irreducible representation of $S_{n+p}$ while the integer coefficients $c^{\alpha\beta}_\gamma$ are their respective multiplicities. In this section, we introduce the recipe for this outer product, without providing any proof.\(^{(22)}\)

**Remark III.65.** A somewhat “cleaner” setup to motivate the outer product and Eq. (III.64) is to choose $n$ elements $i_1, \ldots, i_n$ among $\{1, \ldots, n+p\}$, for which choice there are $\binom{n+p}{n}$ possibilities, denoting $i_{n+1}, \ldots, i_{n+p}$ the remaining $p$ elements. The permutations that only affect the subset $\{i_1, \ldots, i_n\}$, leaving $\{i_{n+1}, \ldots, i_{n+p}\}$ invariant, form a subgroup $H_1$ of $S_{n+p}$ which is isomorphic to $S_n$. Conversely, the permutations that only affect $\{i_{n+1}, \ldots, i_{n+p}\}$ while leaving $\{i_1, \ldots, i_n\}$ unchanged form a subgroup $H_2$ isomorphic to $S_p$. The outer product (III.64) is then the product of irreducible representations of different subgroups of $S_{n+p}$.

One easily checks that every permutation of $S_{n+p}$ can be written as the (commutative) composition of an element of $H_1$ and an element of $H_2$, and that the intersection $H_1 \cap H_2$ reduces to the singleton consisting of the identity. $H_1$ and $H_2$ are thus so-called complementary subgroups of $S_{n+p}$. Since $H_1$ and $H_2$ are in general not normal, $S_{n+p}$ is not their direct product.

Let us illustrate the idea, as well as the recipe to find the decomposition on the right hand side of Eq. (III.64), on an example:

\[
\begin{array}{c}
\begin{array}{ccc}
\hline
& & \\
\hline
\end{array} \\
\begin{array}{ccc}
\hline
\end{array} \\
\end{array} \otimes \\
\begin{array}{c}
\begin{array}{ccc}
\hline
\hline
\end{array} \\
\begin{array}{ccc}
\hline
\end{array} \\
\end{array}
\]

Here the Young diagram $(\alpha)$ resp. $(\beta)$ corresponds to an irreducible representation of $S_4$ resp. $S_3$. To perform their outer product, one proceeds as follows.

- Take one of the diagrams — usually the smallest one, thus here $(\beta)$ — and place an $a$ in each box of its first row, a $b$ in each box of the second row and so on:

\[
\begin{array}{c}
\begin{array}{c}
 a \\
 a \\
 b
\end{array}
\end{array}
\]

- Add the boxes $[a]$ to the diagram $(\alpha)$, to the right (or at the bottom of the first column) and in such a way that (i.) two of them do not appear in the same column and (ii.) the resulting diagrams are valid ones, i.e. the number of boxes decreases from top to bottom:

\[
\begin{array}{c}
\begin{array}{ccc}
 a & a \\
 a \\
 a \\
\end{array} , \begin{array}{ccc}
 a & a \\
 a \\
 a \\
\end{array} , \begin{array}{ccc}
 a \\
 a & a \\
 a \\
\end{array} , \begin{array}{ccc}
 a \\
 a & a \\
 a \\
\end{array} , \begin{array}{ccc}
 a & a \\
 a & a \\
 a \\
\end{array} , \\
\end{array}
\]

- Add the boxes $[b]$ to each of the diagrams obtained at the previous step, respecting the rules (i.) and (ii.) as well as an extra restriction, namely that (iii.) when reading the new diagram from top to bottom and from right to left, the number of boxes $[b]$ always remains smaller than or equal to the number of boxes $[a]$:

\[
\begin{array}{c}
\begin{array}{ccc}
 a & a & a \\
 b & a \\
 b \\
\end{array} , \begin{array}{ccc}
 a & a & a \\
 b & a \\
 b \\
\end{array} , \begin{array}{ccc}
 a & a & a \\
 b & a \\
 b \\
\end{array} , \\
\end{array}
\]

If diagram $(\beta)$ contains further rows, add the corresponding boxes $[c], [d], \ldots$ using the same rules (i.), (ii.), and (iii.).

\(^{(22)}\)The outer product of Young diagrams will prove useful again when we consider the representations of the general linear groups.
• Eventually, remove the letters from the boxes: some diagrams may appear several times, which means that the corresponding irreducible representation is contained with a multiplicity \( c_\alpha^\beta \geq 2 \) in the direct sum on the right hand side of Eq. (III.64).

\[
\begin{array}{cccc}
\emptyset & \otimes & \boxtimes & = \\
\emptyset & \oplus & \boxtimes & \oplus & \boxtimes & \oplus & 2 \\
\end{array}
\]

Remarks:

* One can compute the dimension of each of the irreducible representations involved in the outer product. For the previous example, one finds for the representations of \( S_7 \) on the right hand side \( 14, 15, 14, (2 \times 35), 20, 21, 21, 35 \) — where one recognizes that conjugate representations have the same dimension. Their sum, under consideration of the twofold appearance of one of them, thus equals 210. On the other hand, the irreducible representations of \( S_4 \) and \( S_3 \) on the left hand side are of dimension 3 and 2, respectively: multiplying these dimensions, and multiplying the result (6) by \( \binom{7}{4} = \binom{7}{3} = 35 \) — corresponding to the number of ways to choose 4 (or 3) elements among 7, see remark III.65 — also yields 210.

* The outer product of two Young diagrams (or of the corresponding irreducible representations) is commutative, i.e. one can exchange the roles of \( \alpha \) and \( \beta \) and still obtain the same result.
CHAPTER IV

Physical applications

IV.1 Electric dipole of a molecule
coming soon

IV.2 Molecular vibrations
coming soon
IV.3 Crystal field splitting

The energy levels of electrons orbiting around a given atomic nucleus depend on the environment of the atom. If the latter is “isolated”, as e.g. in a gas, one can assume that the potential felt by an electron has spherical symmetry, i.e. is preserved by any three-dimensional rotation about an axis going through the nucleus. As the reader learned in quantum mechanics, this symmetry leads to degeneracy in the energy spectrum: all levels with a given azimuthal quantum number $\ell$ yet with different magnetic quantum numbers $m_\ell \in \{-\ell, \ldots, \ell - 1, \ell\}$ have the same energy. However, if the atom is part of a crystal, the spherical symmetry is broken, i.e. the symmetric group is smaller, which leads to (partial) lifting of some of the degeneracy. In this Section, we illustrate this loss of degeneracy, known as “crystal field splitting”\[^{(23)}\] on an example.

Consider a cubic lattice of identical ions forming a crystal, as shown in Fig. IV.1. Focusing on the nearest neighbors, a given “tagged” ion in this crystal is surrounded by 6 other ions at the same distance $a$ — the lattice spacing — (Fig. IV.2, left), which constitute the vertices of a regular octahedron with the tagged ion at its center (Fig. IV.2, right).

\[^{(23)}\] Or “ligand field splitting” in chemistry.
Accordingly, the rotations that leave the geometry of the system invariant are those that preserve a regular octahedron — or equivalently a cube —, which we now enumerate. These rotations constitute a group $\mathcal{O}$, whose conjugacy classes we also give without proof.

- The identity transformation $\text{Id}$, which as always is in a conjugacy class on its own, which will hereafter will be denoted $[\text{Id}]$.

- The octahedron possesses four axes with threefold symmetry, namely the straight lines going through the centers of opposite faces (see Fig. IV.3). For each of these axes, there are two proper rotations, namely through $120^\circ$ and $240^\circ$. This gives all in all 8 such rotations, which constitute a conjugacy class $C_3$.

- The three diagonals of the octahedron, i.e. the straight lines joining vertices that do not belong to a same face (Fig. IV.4), are axes with fourfold symmetry. All in all, there are thus 9 proper

\begin{center}
\textbf{Figure IV.3} – Example of threefold-symmetry axis.
\end{center}

\begin{center}
\textbf{Figure IV.4} – The three axes with fourfold symmetry of a regular octahedron.
\end{center}
rotations, through 90, 180 or 270°, around these axes. The 3 rotations through 180° constitute a conjugacy class $3C_2$, while the 6 other rotations are in a class $C_4$.

- Eventually, the 12 edges of the octahedron form 3 squares. Consider the straight line joining the middle points of opposite sides of such a square, as shown in Fig. IV.5 a rotation through 180° about such an axis leaves the octahedron invariant. There are 6 such axes with twofold symmetry, yielding 6 new rotations through 180°, which constitute a conjugacy class $6C_2$.

![Figure IV.5](#)

All in all, $O$ thus consists of 24 rotations (including the identity), which are partitioned into 5 conjugacy classes.

**Remarks:**

* Under each of the rotations of $O$, two opposite faces of the octahedron are “locked” together, so that only four pairs of faces are permuted. Accordingly, the 24 rotations of $O$ are in one-to-one correspondence with the 24 permutations of $S_4$.

* The symmetry group of a regular octahedron actually also includes reflections and their product with the rotations considered here. For simplicity we leave those transformations aside in the following.

Since the group $O$ has five conjugacy classes, it admits 5 (inequivalent) irreducible representations. The corresponding character table is given (without proof) in Table IV.1.

<table>
<thead>
<tr>
<th></th>
<th>$\text{Id}$</th>
<th>$C_3$</th>
<th>$3C_2$</th>
<th>$6C_2$</th>
<th>$C_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi^{(1)}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\varphi^{(1')}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>$\varphi^{(2)}$</td>
<td>2</td>
<td>−1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varphi^{(3)}$</td>
<td>3</td>
<td>0</td>
<td>−1</td>
<td>−1</td>
<td>1</td>
</tr>
<tr>
<td>$\varphi^{(3')}$</td>
<td>3</td>
<td>0</td>
<td>−1</td>
<td>1</td>
<td>−1</td>
</tr>
</tbody>
</table>

**Table IV.1** – Character table of the group $O$ of octahedron-preserving rotations.
The group $O$ of octahedron-preserving rotations is a subgroup of the group of three-dimensional rotations. Accordingly, every representation of the latter will also be a representation of the former. Anticipating on the results of a later chapter, the group of three-dimensional rotations admits for every $\ell \in \mathbb{N}$ an irreducible $(2\ell + 1)$-dimensional representation $D_\ell$ with character

$$\chi_\ell(\theta) = \sum_{m_\ell = -\ell}^{\ell} e^{im_\ell \theta} = \begin{cases} 2\ell + 1 & \text{if } \theta = 0 \\ \frac{\sin[(\ell + \frac{1}{2})\theta]}{\sin \frac{\theta}{2}} & \text{for } \theta \neq 0 \end{cases}$$ (IV.1)

for a rotation through the angle $\theta$ [see Eq. (VI.48)]. From what was just said, $D_\ell$ is also a representation of $O$, yet not necessarily irreducible. We shall hereafter find the reduction of $D_\ell$ in irreducible representations of $O$.

In the context of the physical problem of an electron in a central potential, $D_\ell$ is the representation that operates on the energy eigenfunctions with azimuthal quantum number $\ell$, which all correspond to the same energy. If $D_\ell$ becomes reducible when the symmetry group is broken to $O$, this means that the degeneracy of the energy eigenstates is possibly lifted.

Now, using Eq. (IV.1) one can compute the character $\chi_\ell$ for each conjugacy class of $O$. Thus, since the identity transformation is a rotation through $0^\circ$ one finds

$$\chi_\ell([\text{Id}]) = 2\ell + 1,$$ (IV.2)

which simply means that $D_\ell$ is $(2\ell + 1)$-dimensional. The conjugacy classes $3C_2$ and $6C_2$ both correspond to rotations through $180^\circ$ — which are all in the same conjugacy class in the group of three-dimensional rotations. Using $\chi_\ell(\pi)$, one thus finds

$$\chi_\ell(3C_2) = \chi_\ell(6C_2) = (-1)^\ell.$$ (IV.3)

The character (IV.1) is an even function of $\theta$, which in particular gives $\chi_\ell(2\pi/3) = \chi_\ell(-2\pi/3)$ and $\chi_\ell(\pi/2) = \chi_\ell(-\pi/2)$. One thus obtains on the one hand

$$\chi_\ell(C_3) = \begin{cases} 1 & \text{for } \ell = 0 \text{ mod } 3 \\ 0 & \text{for } \ell = 1 \text{ mod } 3 \\ -1 & \text{for } \ell = 2 \text{ mod } 3, \end{cases}$$ (IV.4)

and on the other hand

$$\chi_\ell(C_4) = \begin{cases} 1 & \text{for } \ell = 0 \text{ or } 1 \text{ mod } 4 \\ -1 & \text{for } \ell = 2 \text{ or } 3 \text{ mod } 4. \end{cases}$$ (IV.5)

The characters of the representations $D_\ell$ with $\ell = 0, 1, 2, 3, 4$ for the five conjugacy classes of $O$ are summarized in table IV.2.

<table>
<thead>
<tr>
<th></th>
<th>[Id]</th>
<th>$C_3$</th>
<th>$3C_2$</th>
<th>$6C_2$</th>
<th>$C_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_\ell=0$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$D_\ell=1$</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$D_\ell=2$</td>
<td>5</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$D_\ell=3$</td>
<td>7</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$D_\ell=4$</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table IV.2
 Comparing this table with the character table IV.1 of $\mathcal{O}$, one sees
\[ \chi_{\ell=0} = \chi^{(1)} \quad \text{i.e.} \quad \mathcal{D}_{\ell=0} = \mathcal{D}^{(1)} \]
and
\[ \chi_{\ell=1} = \chi^{(3)} \quad \text{i.e.} \quad \mathcal{D}_{\ell=1} = \mathcal{D}^{(3)}. \]
The representations $\mathcal{D}_{\ell=0}$ and $\mathcal{D}_{\ell=1}$ of the group of three-dimensional rotations are thus irreducible representations of $\mathcal{O}$. The corresponding physical states of an electron in a central potential — in the usual denomination of atomic physics, s-states with $l = 0$ and p-states with $\ell = 1$ — are thus still in irreducible representations of $\mathcal{O}$, so that their respective degeneracies persist.

Using now the results of Sec. III.4.1, as e.g. Eq. (III.29) for the integer coefficients of the reduction of a representation in irreducible representations, one easily finds
\[ \chi_{\ell=2} = \chi^{(2)} + \chi^{(3')} \quad \text{i.e.} \quad \mathcal{D}_{\ell=2} = \mathcal{D}^{(2)} \oplus \mathcal{D}^{(3')}, \]
which shows that the eigenstates with azimuthal quantum number $\ell = 2$, the d-states, will belong to two different irreducible representations of the symmetry group $\mathcal{O}$. Accordingly, the 5 orthogonal energy eigenstates with $\ell = 2$ will split into a doublet (corresponding to the two-dimensional irrep. $\mathcal{D}^{(2)}$) and a triplet (corresponding to $\mathcal{D}^{(3')}$), with possibly different energies.

Similarly, one finds
\[ \chi_{\ell=3} = \chi^{(1')} + \chi^{(3)} + \chi^{(3')} \quad \text{i.e.} \quad \mathcal{D}_{\ell=3} = \mathcal{D}^{(1')} \oplus \mathcal{D}^{(3)} \oplus \mathcal{D}^{(3')}, \]
and
\[ \chi_{\ell=4} = \chi^{(1)} + \chi^{(2)} + \chi^{(3)} + \chi^{(3')} \quad \text{i.e.} \quad \mathcal{D}_{\ell=4} = \mathcal{D}^{(1)} \oplus \mathcal{D}^{(2)} \oplus \mathcal{D}^{(3)} \oplus \mathcal{D}^{(3')}, \]
which illustrates the reduction of higher-dimensional representations $\mathcal{D}_{\ell}$, which will also lead to a loss of degeneracy.
In physics, one encounters not only finite or discrete symmetry groups, but also “continuous symmetries”, corresponding to transformations described by one or several real parameters that take their values in an interval of $\mathbb{R}$ or in a continuous subset of some $\mathbb{R}^s$. Obvious examples are translations or rotations in two- or three-dimensional space, or the multiplication of the state vector $|\psi\rangle$ of a quantum-mechanical system by a phase factor $e^{i\delta}$ where $\delta \in \mathbb{R}$ is arbitrary.

All in all, we are thus led to consider groups consisting of elements $g(\theta_1, \ldots, \theta_s)$ labeled by a number $s \in \mathbb{N}$ of real parameters $\theta_1, \ldots, \theta_s$, such that every parameter is relevant — i.e. one cannot make $s$ smaller — and that each $\theta_a$ takes values in an interval of $\mathbb{R}$. Such a group will be referred to as an $s$-parameter continuous group, and its parameters will sometimes be collectively denoted either $\{\theta_a\}$ or $\theta$.

Remark: When counting the parameters of a continuous group, only those parameters which can take continuous values are to be taken into account.

In addition, we shall require that the correspondence $\theta \mapsto g(\theta)$ be continuous — we shall in fact even require analyticity —, which necessitates the introduction of topological notions on the set underlying the group (Sec. V.1 and Appendix A).

V.1 Lie groups

The continuous groups of interest in physics are actually more than merely continuous: their dependence on the parameters is actually (real) analytic, constituting the so-called Lie\(^{(o)}\) groups, which we now define (Sec. V.1.1) and of which we give the most currently encountered examples (Sec. V.1.2).

V.1.1 Definition and first results

Definition V.1. A real resp. complex Lie group $\mathcal{G}$ is a real resp. complex differentiable manifold on which two analytic functions $\varphi : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ and $f : \mathcal{G} \to \mathcal{G}$ are defined, which provide $\mathcal{G}$ with a group structure: the group law is such that $g_1 \cdot g_2 = \varphi(g_1, g_2)$ and the inverse element given by $g^{-1} = f(g)$.

Stated differently, a Lie group is a group, as defined by axioms [G1]–[G4], whose operations — multiplication and inversion — are analytic (or at least smooth).

Remarks:

* The dimension of $\mathcal{G}$ as a manifold\(^{(24)}\) is the number of parameters needed to label a group element.

\(^{(24)}\)The definition of a manifold and of closely related mathematical notions is recalled in Appendix A.2

\(^{(o)}\)S. Lie, 1842–1899
That the functions \( \varphi \) and \( f \) provide a group structure constrains their possible forms. For instance, the associativity property \( G2 \) becomes

\[
\varphi(\varphi(g, g'), g'') = \varphi(g, \varphi(g', g'')) \quad \forall g, g', g'' \in G,
\]

while the existence of an identity element \( e \) (axiom \( G3 \)) reads

\[
\varphi(e, g) = \varphi(g, e) = g \quad \forall g \in G.
\]

In turn, the characteristic property of the inverse requires

\[
\varphi(f(g), g) = \varphi(g, f(g)) = e \quad \forall g \in G.
\]

One can show that these requirements, considered for elements \( g \in G \) in the neighborhood of the identity element and using a chart on that neighborhood, can be used to deduce the existence of the structure constants introduced below.

Instead of requiring the analyticity of \( \varphi \) and \( f \), one can use weaker assumptions which lead to the same result, especially in the finite-dimensional case we are mostly interested in.

In the following we shall only consider Lie groups consisting of matrices — the group law being the usual matrix multiplication. These groups obviously depend on only a finite number \( s \) of real parameters, and the identity element of the group will always be an identity matrix (whose dimension naturally depends on the specific group under consideration), which will generically be denoted \( 1 \). More precisely, one has the following definition:

**Definition V.2.** A matrix Lie group is a closed subgroup of the group \( \text{GL}(n, \mathbb{C}) \) of regular \( n \times n \) matrices with complex entries.

**V.1.2 Examples**

We now list a few families of groups which are often encountered in physics, in the role of symmetry groups either for “physical problems” or for toy models. In the following paragraphs (except in § V.1.2d), \( n \) denotes a natural number.

**V.1.2a General linear groups**

**Theorem & Definition V.3.** The \( n \times n \) invertible matrices with real resp. complex entries form a group, the general linear group \( \text{GL}(n, \mathbb{R}) \) resp. \( \text{GL}(n, \mathbb{C}) \).

\( \text{GL}(n, \mathbb{R}) \) is a Lie group with \( n^2 \) continuous parameters — the condition that the matrix determinant be non-zero does not constitute a strong enough condition to decrease the number of parameters.\(^{25} \) On the other hand, since the image of \( \text{GL}(n, \mathbb{R}) \) by the determinant, i.e. a continuous mapping, is the disconnected set \( \mathbb{R}^* \) proves that \( \text{GL}(n, \mathbb{R}) \) cannot be connected. More precisely, \( \text{GL}(n, \mathbb{R}) \) consists of two connected components (two sheets) corresponding to the two possible signs of the determinant. Eventually, \( \text{GL}(n, \mathbb{R}) \) is noncompact.

**Remark:** The connected component of the identity matrix \( \mathbf{1}_n \), i.e. the \( n \times n \) invertible matrices with positive determinant, form a (connected!) subgroup — this is a generic property —, denoted \( \text{GL}^+(n, \mathbb{R}) \), which is also of dimension \( n^2 \).

\( \text{GL}(n, \mathbb{C}) \) is a complex Lie group with \( 2n^2 \) real parameters, which unlike \( \text{GL}(n, \mathbb{R}) \) is connected, yet also noncompact.

\(^{25}\) A trick to quickly check that a general formula for the number of parameters is not wrong (if not correct) is to check the case \( n = 1 \). Thus, \( \text{GL}(1, \mathbb{R}) \) is \( \mathbb{R}^* \), i.e. a one-parameter group; \( \text{GL}(1, \mathbb{C}) \) is \( \mathbb{C} \setminus \{0\} \), i.e. a two-parameter group; \( \text{SL}(1, \mathbb{R}), \text{SL}(1, \mathbb{C}), \text{SO}(1), \text{SU}(1) \) are reduced to the singleton \( \{1\} \), i.e. are zero-parameter groups; \( \text{O}(1) \) contains only the elements \( 1 \) and \( -1 \), i.e. is also a zero-parameter group; \( \text{U}(1) \) is (isomorphic to) the unit circle in the complex plane, i.e. it is a one-parameter group.
V.1.2 b Special linear groups

**Theorem & Definition V.4.** The $n \times n$ invertible real resp. complex matrices with determinant one constitute the special linear group $SL(n, \mathbb{R})$ resp. $SL(n, \mathbb{C})$.

The condition $\det M = 1$ is strong enough to “cost” one real or complex parameter: a matrix of $GL(n, \mathbb{R})$ resp. $GL(n, \mathbb{C})$ can always be written as the product of its real- or complex-valued determinant with a matrix of $SL(n, \mathbb{R})$ resp. $SL(n, \mathbb{C})$. Accordingly $SL(n, \mathbb{R})$ is of dimension $n^2 - 1$ and $SL(n, \mathbb{C})$ is of dimension $2(n^2 - 1)$.(25) Both groups are connected and noncompact.

**Remark:** Geometrically, $SL(n, \mathbb{R})$ corresponds to the geometrical transformations that preserve the $(n$-dimensional) volume and the orientation.

V.1.2 c Orthogonal groups

**Theorem & Definition V.5.** The $n \times n$ real matrices $O$ such that $O^T O = O O^T = \mathbb{I}_n$ form a group, the orthogonal group, denoted $O_n$.

**Property V.6.** The orthogonal matrices represent the isometries of the $n$-dimensional Euclidean space that leave the origin invariant.

Stated differently, the orthogonal matrices $O$ represent the transformations $X \mapsto X' = OX$ that preserve the Euclidean scalar product

$$X^T Y = \sum_{i=1}^{n} x_i y_i,$$

where $X$ resp. $Y$ is a vector with $n$ components $(x_1, \ldots, x_n)$ resp. $(y_1, \ldots, y_n)$.

**Property V.8.** The Lie group $O(n)$ is of dimension $\frac{n(n-1)}{2}$.

Proof: A matrix $O \in O(n)$ transforms an orthonormal basis for the scalar product (V.7) into another orthonormal basis, so that its columns (or equivalently its lines) are of norm 1 and pairwise orthogonal. Accordingly, one can construct $O$ by choosing $n - 1$ independent real parameters for the first column — the $n$ entries have to obey the normalization condition, which removes one degree of freedom —, $n - 2$ parameters for the second column — the normalization fixes one parameter, the orthogonality with the first column a second one —, and more generally $n - k$ independent parameters for the $k$-th column. All in all there are thus

$$\sum_{k=1}^{n} (n - k) = n^2 - \frac{n(n + 1)}{2} = \frac{n(n - 1)}{2}$$

independent real parameters. \(\square\)

The defining property $O^T O = O O^T = \mathbb{I}_n$ leads at once to $|\det O| = 1$ and to the normalization of each column to 1 for the Euclidean scalar product (V.7). Accordingly, the entries of an orthogonal matrix are bounded, so that the group $O(n)$ is compact. On the other hand, since both $\det O = +1$ and $\det O = -1$ are possible, $O(n)$ is disconnected, consisting of two connected components.

The connected component of the identity matrix is a Lie subgroup of $O(n)$:

**Theorem & Definition V.9.** The $n \times n$ real matrices $O$ such that $O^T O = O O^T = \mathbb{I}_n$ and $\det O = 1$ form a group, the special orthogonal group $SO(n)$.

**Property V.10.** The special orthogonal matrices represent the isometries of the $n$-dimensional Euclidean space that leave the origin invariant and preserve the orientation, i.e. precisely the rotations in $n$-dimensional space.

$SO(n)$ is of dimension $\frac{n(n-1)}{2}$, connected(!), and compact.
V.1.2d Indefinite orthogonal groups

Let \( p \) and \( q \) be two natural numbers and \( \eta \) be the diagonal \((p+q) \times (p+q)\) matrix defined as

\[
\eta = \text{diag}(1, 1, \ldots, 1, -1, \ldots, -1).
\]  
(V.11)

**Theorem & Definition V.12.** The \((p+q) \times (p+q)\) real matrices \( \Lambda \) such that \( \Lambda^T\eta\Lambda = \eta \) form a group, the indefinite orthogonal group \( O(p, q) \).

One at once sees that the matrices \( \Lambda \) obey \(|\det \Lambda| = 1\), i.e. are invertible, and the defining property gives \( \Lambda^{-1} = \eta^{-1}\Lambda^T\eta \), where actually \( \eta^{-1} = \eta \).

Defining a (non-positive definite) scalar product by

\[
X^T\eta Y = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{p+q} x_i y_i,
\]  
(V.13)

where \( X \) resp. \( Y \) is a vector with \( p + q \) components \((x_1, \ldots, x_{p+q})\) resp. \((y_1, \ldots, y_{p+q})\), a matrix \( \Lambda \) of \( O(p, q) \) represents a transformation that preserve that scalar product.

**Remarks:**
- Replacing \( \eta \) by \(-\eta\) (with \( p \) minus signs and \( q \) plus signs in the scalar product) leaves \( O(p, q) \) unchanged.

- The special case \( p = 3, q = 1 \) (or equivalently \( p = 1, q = 3 \), see previous remark) gives the Lorentz\(^{(p)}\) group of relativistic physics, which will be studied in further detail in Chap. VIII.

Denoting \( n = p + q \), \( O(p, q) \) is a noncompact group with \( n(n-1)/2 \) parameters, and with four connected components, two of which correspond to matrices with determinant \( \det \Lambda = +1 \), the other two to matrices with determinant \(-1\).

**Theorem & Definition V.14.** The \((p+q) \times (p+q)\) real matrices \( \Lambda \) such that \( \Lambda^T\eta\Lambda = \eta \) and \( \det \Lambda = 1 \) form a group, the indefinite special orthogonal group \( \text{SO}(p, q) \).

\( \text{SO}(p, q) \) is again noncompact and still consists of two connected components. The connected component of the identity matrix is a Lie group, denoted \( \text{SO}^+(p, q) \).

**Remarks:**
- The group \( \text{SO}(3, 1) \) is that of proper Lorentz transformations.

- One can also find a Lie subgroup \( \text{O}^+(p, q) \) of \( O(p, q) \), which consists of two of the four sheets. In the case of the Lorentz group \( O(3,1) \), this subgroup \( \text{O}^+(3, 1) \) will be that of orthochronous Lorentz transformations.

V.1.2e Unitary groups

**Theorem & Definition V.15.** The \( n \times n \) complex matrices \( U \) such that \( U^*U = UU^* = I_n \) form a group, the unitary group, denoted \( U(n) \).

**Property V.16.** The unitary matrices represent the isometries of the \( n \)-dimensional Hermitian space that leave the origin invariant.

Stated differently, the unitary matrices \( U \) represent the transformations \( X \mapsto X' = UX \) that preserve the Hermitian product

\[
X^\dagger Y = \sum_{i=1}^{n} x_i^* y_i,
\]  
(V.17)

where \( X \) resp. \( Y \) is a vector with \( n \) complex components \((x_1, \ldots, x_n)\) resp. \((y_1, \ldots, y_n)\).

\(^{(p)}\)H. Lorentz, 1853–1928
Property V.18. The Lie group $U(n)$ is of dimension $n^2$.

Proof: A matrix $U \in U(n)$ transforms an orthonormal basis for the scalar product (V.17) into another orthonormal basis, so that its columns (or equivalently its lines) are of norm 1 and pairwise orthogonal. Accordingly, one can construct $U$ by choosing $2n - 1$ independent real parameters for the first column — the $n$ complex entries have to obey the normalization condition, which removes one degree of freedom —, $2n - 3$ real parameters for the second column, and more generally $2n - (2k - 1)$ independent parameters for the $k$-th column. All in all there are thus
\[
\sum_{k=1}^{n} [2n - (2k - 1)] = 2n^2 - 2\frac{n(n + 1)}{2} + n = n^2
\]

independent real parameters. \qed

The defining property $U^\dagger U = UU^\dagger = \mathbb{I}_n$ leads at once to $|\det U| = 1$ and to the normalization of each column of $U$ to 1 for the Hermitian product (V.17). The entries of a unitary matrix are thus bounded, so that the group $U(n)$ is compact. Unlike the orthogonal group $O(n)$, it is also connected. It contains an important subgroup:

**Theorem & Definition V.19.** The $n \times n$ complex matrices $U$ such that $U^\dagger U = UU^\dagger = \mathbb{I}_n$ and $\det U = 1$ form a group, the special unitary group $SU(n)$.

Requiring that the determinant be equal to 1 suppresses a real parameter with respect to a general matrix of $U(n)$ — instead of $\det U = e^{i\delta}$ with an arbitrary $\delta \in [0, 2\pi]$, the latter is now fixed to 0. Accordingly, $SU(n)$ is of dimension $n^2 - 1$. It is still connected and compact.

**V.1.2f Symplectic groups**

Let $\mathbb{I}_n$ denote the $n \times n$ identity matrix with $n \in \mathbb{N}^*$. One defines a $2n \times 2n$ antisymmetric matrix $J_{2n}$ as
\[
J_{2n} \equiv \begin{pmatrix} 0 & \mathbb{I}_n \\ -\mathbb{I}_n & 0 \end{pmatrix}.
\] (V.20)

**Theorem & Definition V.21.** The $2n \times 2n$ real matrices $M$ such that $M^T J_{2n} M = J_{2n}$ form a group, the symplectic group, denoted $Sp(2n, \mathbb{R})$.

**Remark V.22.** The defining property gives at once $\det M = \pm 1$, which ensures the regularity of $M$ and thus the existence of an inverse, which is indeed given by $M^{-1} = J_{2n}^T M^T J_{2n}$.

**Property V.23.** The symplectic matrices $M$ represent the group of transformations $X \mapsto X' = MX$ that preserve the bilinear product
\[
X^T J_{2n} Y = \sum_{i=1}^{n} (x_{i+y_{i+n}} - x_{i+n}y_i),
\] (V.23)

where $X$ resp. $Y$ is a vector with $2n$ real components $(x_1, \ldots, x_{2n})$ resp. $(y_1, \ldots, y_{2n})$.

**Property V.24.** The Lie group $Sp(2n, \mathbb{R})$ is of dimension $n(2n + 1)$.

A possible proof runs similar to that of property V.8. Note, however, that $Sp(2n, \mathbb{R})$ has the same dimension as $O(2n + 1)$.

Unlike the orthogonal and unitary groups, the symplectic group $Sp(2n, \mathbb{R})$ is noncompact. It is connected, and one can show that the determinant of a symplectic matrix actually always equals 1 \textsuperscript{(26)}.

**Remark:** The symplectic group $Sp(2, \mathbb{R})$ coincides with the special linear group $SL(2, \mathbb{R})$.

\textsuperscript{(26)} An elementary proof can be found in Ref. [1].
V.2 Lie algebras

V.2.1 General definitions

V.2.1 a Lie algebras

We recall that algebras on a field were defined in definition III.31, with in particular the three properties \( A_1 \)–\( A_3 \) of the vector multiplication.

Definition V.25. A \( \textit{Lie algebra} \ \mathcal{A} \) is an algebra such that its vector multiplication fulfills two extra requirements:

- \textbf{alternativity:} for all \( x \in \mathcal{A} \), \( x \star x = 0_{\mathcal{A}} \); \( (\mathcal{L}A1) \)

- \textbf{Jacobi identity:} for all \( x, y, z \in \mathcal{A} \), \( x \star (y \star z) + y \star (z \star x) + z \star (x \star y) = 0_{\mathcal{A}} \), \( (\mathcal{L}A2) \)

where \( 0_{\mathcal{A}} \) is the zero vector of the algebra.

Property V.26. Being bilinear and alternating, the vector multiplication is also anticommutative, that is for all \( x, y \in \mathcal{A} \), \( x \star y = -y \star x \), \( (\mathcal{L}A1') \)

which is often used in lieu of Eq. \( (\mathcal{L}A1) \) in the definition. \( (27) \)

This follows at once from expanding the identity \((x + y) \star (x + y) = 0_{\mathcal{A}}\).

Examples:

- The vector space \( \mathbb{R}^3 \) with the cross product \( \times \) is a Lie algebra.

- The set \( \mathcal{M}(n, \mathbb{R}) \) resp. \( \mathcal{M}(n, \mathbb{C}) \) of \( n \times n \) matrices with real resp. complex entries with the commutator \( AB - BA \) as vector multiplication is a Lie algebra of dimension \( n^2 \).

The 3 \( \times \) 3 strictly upper triangular real matrices constitute a Lie subalgebra of \( \mathcal{M}(3, \mathbb{R}) \) called the \textit{Heisenberg algebra}, and which we shall denote \( \mathfrak{h}_3 \). A basis of this Lie algebra is obviously

\[
M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

One quickly checks \( M_1 \star M_2 \equiv M_1M_2 - M_2M_1 = M_3 \), \( M_1 \star M_3 \equiv M_1M_3 - M_3M_1 = 0 \), and \( M_2 \star M_3 \equiv M_2M_3 - M_3M_2 = 0 \), and thus for all \( A, B, C \in \mathfrak{h}_3 \), \( A \star (B \star C) = (A \star B) \star C = 0 \), where 0 stands for the zero 3 \( \times \) 3 matrix. \( \mathfrak{h}_3 \) is thus an example of associative Lie algebra, but one sees that requiring associativity, alternativity and the Jacobi identity leads to the strong constraint \( A \star (B \star C) = 0 \) for all \( A, B, C \in \mathfrak{h}_3 \).

Notation V.27. The vector multiplication \( x \star y \) of a Lie algebra is traditionally denoted \([x, y]\), called the \textit{Lie bracket}, so that the Jacobi identity \( (\mathcal{L}A2) \) reads

\[
[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0_{\mathcal{A}} \quad \text{for all } x, y, z \in \mathcal{A}.
\]

(\( \mathcal{L}A2' \))

In the following, we shall only consider matrix Lie algebras on which the vector multiplication is the commutator \( AB - BA \), so that the notation \([A, B]\) should hopefully not confuse the reader used to employ it precisely for that purpose.

\( (27) \) There is indeed equivalence between Eq. \( (\mathcal{L}A1) \) (with bilinearity) and Eq. \( (\mathcal{L}A1) \) in case the base field of the algebra is \( \mathbb{R} \) or \( \mathbb{C} \), yet not for a field of characteristic 2.

\( \text{C. G. Jacobi, 1804–1851} \) \( \text{W. Heisenberg, 1901–1976} \)
V.2.1b Structure constants

Definition V.28. Let \( \{ X^a \}_{1 \leq a \leq s} \) be a basis of a (matrix) Lie algebra \( \mathcal{A} \) of dimension \( s \). The product of two basis vectors — often called generators — \( X^a, X^b \) is a linear combination of the basis vectors

\[
[X^a, X^b] = \sum_{c=1}^{s} c^{ab}_c X^c, \quad (V.28)
\]

which defines the structure constants \( c^{ab}_c \) of the Lie algebra.

More generally, one may define the structure constants of an algebra which is not a Lie algebra, yet we do not need those in these lecture notes.

Notation V.29. The previous definition uses the mathematicians’ convention. Physicists rather denote the generators \( \{ T^a \}_{1 \leq a \leq s} \) and define “their” structure constants through the relation

\[
[T^a, T^b] = i \sum_{c=1}^{s} f^{ab}_c T^c. \quad (V.29)
\]

Setting \( T^a = iX^a \) for all \( a \in \{1, \ldots, s\} \), one sees that Eqs. (V.28) and (V.29) define the same structure constants \( f^{ab}_c = c^{ab}_c \).

The coexistence of both conventions naturally means that both have advantages. Anticipating on Sec. V.2.2, we shall see that the mathematicians’ choice gives “more natural” expressions for the form of the elements of a Lie group in the neighborhood of the identity matrix, see Eq. (V.32), while the physicists’ convention lead to Hermitian generators \( T^a \) for the usual matrix groups (see below), instead of anti-hermitian ones.

Remarks:

* Multiplying each generator \( X^a \) or \( T^a \) by a constant number \( \lambda \) in Eqs. (V.28) and (V.29) obviously leads to multiplying the structure constants by the same factor \( \lambda \). However, once a normalization condition has been chosen for the generators, then one can show that the structure constants are in fact independent of the basis. Thus, they really encode an intrinsic property of the Lie algebra, irrespective of a basis choice.

* For compact Lie groups, the position of the indices labeling the structure constants has no meaning. Thus the reader will often find in the (physics) literature \( f^{ab}_c, f^{ab}_c, f^{abc} \ldots \) used somewhat interchangeably. In the case of a non-compact group, some more care is needed.

Property V.30. The structure constants \( f^{ab}_c \) satisfy generic relations, which are inherited from the anticommutativity (LA1) and the Jacobi identity (LA2), namely

\[
f^{ab}_c = - f^{ba}_c \quad \forall a, b, c, \quad (V.30a)
\]

i.e. the structure constants are antisymmetric in their two upper indices, and

\[
\sum_{c=1}^{s} (f^{ab}_c f^{cd}_e + f^{bd}_c f^{ca}_e + f^{da}_c f^{eb}_e) \quad \forall a, b, d, e. \quad (V.30b)
\]

If you wish to memorize this equation, note its structure: \( e \) is fixed, \( c \) is summed over, and \( a, b, d \) are cyclically permuted.

We shall give examples of structure constants, on which the reader can check the above properties, in the following sections.
V.2.2 Lie algebra associated to a Lie group

V.2.2a Tangent space at the identity in a Lie group

Let \( G \) be a matrix Lie group of dimension \( s \). Restricting oneself to a chart that encompasses the identity matrix \( \mathbf{1} \), one chooses a system of coordinates \( \{\theta_a\}_{1 \leq a \leq s} \) on that chart such that \( \mathbf{1} \) is the group element \( g(\theta_1, \ldots, \theta_s) \) corresponding to the origin of the coordinates, i.e.

\[
  g(0, \ldots, 0) = \mathbf{1}. \tag{V.31}
\]

Invoking the differentiability of the Lie group and the smoothness of the charts, one may write a group element in the neighborhood of \( \mathbf{1} \) — i.e. for small enough \( \theta_1, \ldots, \theta_s \) — in the form

\[
g(\theta_1, \ldots, \theta_s) = g(0, \ldots, 0) + \sum_a \theta_a X^a + \mathcal{O}(\theta_a^2) = \mathbf{1} + \sum_a \theta_a X^a + \mathcal{O}(\theta_a^2), \tag{V.32a}
\]

or equivalently, inserting an imaginary unit \( i \) in front of the small parameters

\[
g(\theta_1, \ldots, \theta_s) = g(0, \ldots, 0) - i \sum_a \theta_a T^a + \mathcal{O}(\theta_a^2) = \mathbf{1} - i \sum_a \theta_a T^a + \mathcal{O}(\theta_a^2). \tag{V.32b}
\]

Obviously, one has for each \( a \in \{1, \ldots, s\} \)

\[
  X^a = \lim_{\theta_a \to 0} \frac{g(0, \ldots, 0, \theta_a, 0, \ldots, 0) - \mathbf{1}}{\theta_a}, \quad T^a = i \lim_{\theta_a \to 0} \frac{g(0, \ldots, 0, \theta_a, 0, \ldots, 0) - \mathbf{1}}{\theta_a}. \tag{V.33}
\]

This shows that \( X^a \) resp. \( T^a \) is a vector of the tangent space \( T_1 G \) at the identity of the Lie group.

**Example V.34.** Consider a so-called one-parameter subgroup of \( G \), i.e. a continuous set of elements \( g(t) \) of the Lie group labeled by a single real parameter \( t \) taking its values in a neighborhood of \( 0 \), with the following properties:

\[
g(0) = \mathbf{1} \tag{V.35a}
\]

\[
\forall t_1, t_2, \quad g(t_1 + t_2) = g(t_1)g(t_2), \tag{V.35b}
\]

where the latter holds at least in a neighborhood of \( t = 0 \). Stated differently, \( g(t) \) is a “continuous curve” drawn on the Lie group \( G \) and going through the identity matrix, and which is at least locally a group homomorphism from \( (\mathbb{R}, +) \) into \( G \).

More precisely, we require that \( g(t) \) be a differentiable function of \( t \). For a small enough \( \delta t \) one may then write

\[
g(\delta t) = g(0) + X\delta t + \mathcal{O}((\delta t)^2) = \mathbf{1} + X\delta t + \mathcal{O}((\delta t)^2) \quad \text{with} \quad X = \frac{dg(t)}{dt} \bigg|_{t=0} = g'(0).
\]

\( X \) is a tangent vector at the identity to the Lie group \( G \), \( X \in T_1 G \).

For any \( t \), one may also write \( g(t + \delta t) = g(t) + g'(t)\delta t + \mathcal{O}((\delta t)^2) \), while the one-parameter group property \((V.35b)\) yields

\[
g(t + \delta t) = g(t)g(\delta t) = g(t)\left[\mathbf{1} + X\delta t + \mathcal{O}((\delta t)^2)\right] = g(t) + g(t)X\delta t + \mathcal{O}((\delta t)^2),
\]

where we used the expression of \( g(\delta t) \) given above. Identifying both expressions of \( g(t + \delta t) \), one thus finds after dividing by \( \delta t \) and taking the limit \( \delta t \to 0 \)

\[
g'(t) = g(t)X.
\]

One recognizes a linear differential equation of first order, whose solution must also obey the initial condition \((V.35a)\), yielding

\[
g(t) = e^{tX}.
\]

That is, the one-parameter subgroup \( g(t) \) follows from exponentiating a tangent vector — which, \( G \) being a matrix Lie group, is a matrix — at the identity. Accordingly the definition and properties of the exponential of a matrix will be discussed in §V.2.2c hereafter.
Similarly, we shall see that the relations \([V.32]\) are actually approximations to first order in the parameters \(\{\theta_a\}\) to the more general formulas
\[
\begin{align*}
g(\theta_1, \ldots, \theta_s) &= \exp \left( \sum_a \theta_a X^a \right) \\
g(\theta_1, \ldots, \theta_s) &= \exp \left( -i \sum_a \theta_a T^a \right),
\end{align*}
\]
which we shall also encounter on specific examples.

**Remark:** The example \([V.34]\) shows why the tangent space \(T_1 G\) at the identity of the group seems to play a special role in the study of Lie groups — which will be further discussed in the following paragraph. Thanks to the group law, one can map the behavior of the Lie group in the vicinity of any of its points \(g\) to that in the neighborhood of the identity — this is the step encoded in the equation \(g(t + \delta t) = g(t)g(\delta t) = g(t)[I + \cdots]\) in the example —, so that the tangent spaces \(T_g G\) at \(g\) is isomorphic to that at the identity. In other words, the behavior in a neighborhood of the identity matrix, described by \(T_1 G\), is actually representative of the behavior in a neighborhood of any element of the Lie group.

**V.2.2 b Lie algebra of a Lie group**

By studying a Lie group \(G\) in the neighborhood of the identity matrix, one identifies a Lie algebra.

**Theorem V.37.** The tangent space at the identity of a matrix Lie group \(G\) is a Lie algebra, referred to as the Lie algebra of the group and traditionally denoted \(\mathfrak{g}\), with the matrix commutator as Lie bracket.

Proof: Let us first show that for any tangent vector \(Y \in T_1 G\) and any element \(g \in G\) of the Lie group, the (matrix) product \(gYg^{-1}\) defines a one-parameter subgroup of \(G\), i.e. is also a tangent vector. For that purpose, the property \((V.46)\) of the exponential map yields for every \(t\) such that \(e^{tY} \in G\) the identity
\[
e^{t(gYg^{-1})} = g e^{tY} g^{-1}.
\]
As product of elements of the Lie group, \(e^{t(gYg^{-1})}\) is itself in \(G\), which gives the desired result \(gYg^{-1} \in T_1 G\).

Taking now \(g = e^{\lambda X}\), where \(\lambda \in \mathbb{R}\) and \(X \in T_1 G\) is another tangent vector, one has \(g^{-1} = e^{-\lambda X}\) and thus
\[
e^{\lambda X} Y e^{-\lambda X} \in T_1 G.
\]
The latter result hold for \(\lambda = 0\) and for any \(\lambda\) in a small enough neighborhood of 0: using the linearity of a vector space, the difference \(e^{\lambda X} Y e^{-\lambda X} - Y\) is again in the tangent space, as well as its quotient by \(\lambda\). This still holds in the limit \(\lambda \to 0\), i.e. the derivative of \(e^{\lambda X} Y e^{-\lambda X}\) with respect to \(\lambda\), and compute at \(\lambda = 0\), is still in the tangent space:
\[
\frac{d}{d\lambda} \left( e^{\lambda X} Y e^{-\lambda X} \right) \bigg|_{\lambda = 0} = \left( X e^{\lambda X} Y e^{-\lambda X} - e^{\lambda X} Y e^{-\lambda X} X \right) \bigg|_{\lambda = 0} = X Y - Y X \in T_1 G.
\]
Since the commutator of two matrices obeys the properties of a Lie bracket, the tangent space \(T_1 G\) is a Lie algebra.

One should note at once that while the Lie algebra associated to a given Lie group is unique, yet different Lie groups can have the same Lie algebra — the most obvious example being that of a disconnected group and its subgroup consisting of the connected component of the identity, like \(O(n)\) and \(SO(n)\).

**Corollary V.38.** The Lie algebra \(\mathfrak{g}\) associated to a matrix Lie group \(G\) is the set of complex matrices \(X\) such that \(e^{tX} \in G\) for all \(t \in \mathbb{R}\).

We shall use this result in Sec. **V.2.3** below to write down the Lie algebras of the classical Lie groups of Sec. **V.1.2**

\(^{28}\)The Lie algebra of a group \(G\) is also sometimes denoted \(\text{Lie}(G)\).
Let $X$, $Y$ be two tangent vectors at the identity of a Lie group $G$, i.e. $X, Y \in \mathfrak{g}$. At least for $t, u$ close enough to 0, $e^{tX}$ and $e^{uY}$ belong to $G$. Let us compute their commutator using the approximation $e^{\alpha A} = 1 + \alpha A + \frac{\alpha^2}{2} A^2 + \mathcal{O}(\alpha^3)$. To quadratic order, one finds

$$e^{tX}e^{uY}e^{−tX}e^{−uY} = 1 + tu(XY − YX) + \mathcal{O}(t^3, u^3). \quad (V.39)$$

One thus finds that the Lie bracket $[X, Y] = XY − YX \in \mathfrak{g}$ measures the lack of commutativity of the elements of the one-parameter subgroups $e^{tX}$ and $e^{uY}$.

### V.2.2 Exponential map

**Theorem & Definition V.40.** Let $A$ be an $n \times n$ matrix with real or complex entries. The exponential of $A$ is the matrix given by the power series

$$\exp(A) \equiv e^A \equiv \sum_{k=0}^{\infty} \frac{A^k}{k!}, \quad (V.40)$$

with the convention that $A^0$ is the identity $n \times n$ matrix.

One indeed checks that the series converges for any $A$.

Let us list without proof a few properties of the exponential map. Throughout, $A$ or $B$ are $n \times n$ matrices with real or complex entries.

- $e^{0n} = 1_n$ with $0_n$ resp. $1_n$ the zero resp. identity $n \times n$ matrix. \hfill (V.41)
- $\det e^A = e^{\text{Tr} A}$ \hfill (V.42)
- If $A$ and $B$ are two commuting matrices, i.e. $AB = BA$, then $e^{A+B} = e^A e^B$. \hfill (V.43)

Note that the converse is not true: the matrices

$$A = \begin{pmatrix} i\pi & 0 \\ 0 & -i\pi \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & -2i\pi \end{pmatrix}$$

yield $e^A = -1_2$, $e^B = 1_2$ and $e^{A+B} = -1_2 = e^A e^B$, although they do not commute.

- $(e^A)^{-1} = e^{-A}$. \hfill (V.44)
- For all $t, u \in \mathbb{C}$, $e^{(t+u)A} = e^{tA} e^{uA}$. \hfill (V.45)
- If $B$ is invertible, then $Be^AB^{-1} = e^{BAB^{-1}}$. \hfill (V.46)
- $\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}A$, and in particular $\left. \frac{d}{dt} e^{tA} \right|_{t=0} = A$ \hfill (V.47)

According to corollary V.38, the exponential map provides a mapping between a Lie algebra $\mathfrak{g}$ and a Lie group $\mathcal{G}$. More precisely, one can show the following results:

**Property V.48.** Given a matrix Lie group $\mathcal{G}$ and its Lie algebra $\mathfrak{g}$, there exist neighborhoods $U_0$ of the zero matrix in $\mathfrak{g}$ and $U_1$ of the identity matrix in $\mathcal{G}$ such that the exponential map $X \mapsto e^X$ is a bijection from $U_0$ onto $U_1$.

**Remark:** Obviously, the mention of a neighborhood of the zero matrix in $\mathfrak{g}$ presupposes the introduction of a topological structure on the Lie algebra. Since the latter is a finite-dimensional space of matrices, this represents no difficulty.

**Property V.49.** A sufficient condition to ensure that the exponential map $X \in \mathfrak{g} \mapsto e^X \in \mathcal{G}$ is surjective is if the Lie group $\mathcal{G}$ is connected and compact.

---

(29) The commutator of two elements $g, h$ of a group $\mathcal{G}$ is the element $ghg^{-1}h^{-1}$. 

The condition is not necessary: for example, every matrix of the general linear group \( \text{GL}(n, \mathbb{C}) \), which is non-compact, can be written as the exponential of a matrix of \( \mathfrak{gl}(n, \mathbb{C}) \), i.e. of an \( n \times n \) matrix.

**Property V.50.** A necessary condition for the exponential map to be injective is that the Lie group \( G \) is simply connected.

The condition is not sufficient: the exponential maps both \( 2 \times 2 \) traceless antihermitian matrices

\[
\begin{pmatrix}
\pm i \pi & 0 \\
0 & \mp i \pi \\
\end{pmatrix}
\in \mathfrak{su}(2),
\]

to the special unitary matrix \(-I_2 \in \text{SU}(2)\), i.e. the exponential map is not injective on \( \mathfrak{su}(2) \), although \( \text{SU}(2) \) is simply connected.

An example of non-simply connected Lie group is the abelian group \( \text{U}(1) \), whose Lie algebra \( \mathfrak{u}(1) \) is \( i\mathbb{R} \), i.e. the set of imaginary complex numbers — which is obviously isomorphic to the space of \( 1 \times 1 \) antihermitian matrices. Indeed, every element of \( \mathfrak{u}(1) \) of the form \( 2k\pi i \) with \( k \in \mathbb{Z} \) is mapped by the exponential function to the same element 1 of \( \text{U}(1) \).

According to property \[ \mathbf{48} \] if two matrices \( X, Y \in \mathfrak{g} \) are “close enough” to the zero matrix, then their respective exponentials are elements of a Lie group \( G \) of which \( \mathfrak{g} \) is the Lie algebra, both \( e^X \) and \( e^Y \) being in the vicinity of the identity matrix \( I \). Accordingly, the product \( e^X e^Y \) is also an element of \( G \), which should also lie close to \( I \), and thus be writable in the form \( e^Z \) with \( Z \in \mathfrak{g} \). A natural question then arises, as to whether, and if yes how, \( Z \) is related to \( X \) and \( Y \).

If the latter are “small enough” (for the infinite series involved in the formula to converge), \( Z \) indeed exists and is given by the \( \text{Baker}^{(9)} \), \( \text{Campbell}^{(10)} \), \( \text{Hausdorff}^{(11)} \) formula

\[
e^X e^Y = e^{Z(X,Y)} \quad \text{with} \quad Z(X,Y) = X + Y + \frac{1}{2} [X,Y] + \frac{1}{12} \left( [X,[X,Y]] + [Y,[Y,X]] \right) + \cdots, \tag{V.51}
\]

where all successive terms only depend on iterated commutators of \( X \) and \( Y \).

One can write down several closed forms for the series on the right hand side, especially involving the adjoint mapping introduced below. Instead, let us give special cases of the formula, in which the series is truncated at a low order:

- If \( X \) and \( Y \) commute, one recovers property \[ \mathbf{43} \], namely \( e^X e^Y = e^{X+Y} \).
- If the commutator \([X,Y]\) commutes with both \( X \) and \( Y \), then

\[
e^X e^Y = e^{X+Y+\frac{1}{2}[X,Y]} = e^{X+Y} e^{\frac{1}{2}[X,Y]}. \tag{V.52}
\]

Eventually, let us mention a related formula:

\[
e^X Y e^{-X} = Y + [X,Y] + \frac{1}{2!} [X,[X,Y]] + \frac{1}{3!} [X,[X,[X,Y]]] + \cdots, \tag{V.53}
\]

which is also useful in physical applications.

**V.2.3 Lie algebras associated to the classical Lie groups**

\[ \textbf{V.2.3a} \quad \mathfrak{gl}(n, \mathbb{R}), \mathfrak{gl}(n, \mathbb{C}) \]
\( n \times n \) matrices with real resp. complex entries

\[ \textbf{V.2.3b} \quad \mathfrak{sl}(n, \mathbb{R}), \mathfrak{sl}(n, \mathbb{C}) \]
traceless \( n \times n \) matrices with real resp. complex entries

\[ \textbf{V.2.3c} \quad \mathfrak{o}(n) = \mathfrak{so}(n) \]
traceless antisymmetric \( n \times n \) matrices

\[ \textbf{V.2.3d} \quad \mathfrak{u}(n) = \mathfrak{su}(n) \]
traceless antihermitian \( n \times n \) matrices

Chapter VI
The groups SO(3) and SU(2)
and their representations

Two continuous groups of transformations that play an important role in physics are the special orthogonal group of order 3, SO(3), and the special unitary group of order 2, SU(2), which are in fact related to each other, and to which the present chapter is devoted. We first recall in Secs. VI.1 and VI.2 their most useful properties, which the reader probably knows from previous lectures, and introduce their respective Lie algebras. We then discuss their representations (Sec. VI.3).

VI.1 The group SO(3)
In this section, we begin with a reminder on the rotations in three-dimensional Euclidean space (Sec. VI.1.1) — which as the reader knows models the spatial arena in which the physical objects of non-relativistic physics evolve.

VI.1.1 Rotations in three-dimensional space
Consider first the rotations in three-dimensional Euclidean space that leave a given point O invariant. For the usual composition, they form a continuous group, which is non-Abelian since rotations about different axes do not commute. By choosing a Cartesian coordinate system with O at its origin, one maps these rotations one-to-one to those on \( \mathbb{R}^3 \).

To describe such a rotation, one mostly uses either a set of three angles, or the direction of the rotation axis and the rotation angle, as we now recall.

VI.1.1a Euler angles
A given rotation \( \mathcal{R} \) is fully characterized by its action on three pairwise orthogonal directions in three-dimensional Euclidean space — or equivalently, on an orthonormal basis of \( \mathbb{R}^3 \). Thus, \( \mathcal{R} \) transforms axes \( Ox, Oy, Oz \) into new axes \( Ou, Ov, Ow \), as illustrated in Fig. VI.1. Note that it is in fact sufficient to know the transformations from \( Ox \) to \( Ou \) and from \( Oz \) to \( Ow \), since the remaining axes \( Oy \) and \( Ov \) are entirely determined by the first two directions of a given triplet.

As was already mentioned before, in these lecture notes we consider “active” transformations, i.e. the coordinate system is fixed while the physical system is being transformed — here, rotated. Accordingly, one should not view the transformation from the axes \( Ox, Oy, Oz \) to the directions \( Ou, Ov, Ow \) as a change of coordinates, but as the rotation of “physical” axes — which possibly coincided with the coordinate axes in the initial state, in case the Cartesian coordinates are denoted \( x, y, z \).

The action of \( \mathcal{R} \) can be decomposed into the composition of three rotations about axes that are simply related to \( Oy, Oz, Ou, \) and \( Ow \):

• First, one performs a rotation about \( Oz \) through an angle \( 0 \leq \alpha < 2\pi \), which is actually determined by the second step, and which brings the axis \( Oy \) along a new direction \( On \), the so-called line of nodes. This first rotation will be denoted \( \mathcal{R}_z(\alpha) \).
The groups SO(3) and SU(2) and their representations

Figure VI.1 – Euler angles

- Second, one performs a rotation \( R_n(\beta) \) about \( On \) through an angle \( 0 \leq \beta \leq \pi \) such that it brings the “old” axis \( Oz \) along the “new” direction \( Ow \).
- Eventually, one performs a rotation \( R_w(\gamma) \) about \( Ow \) through an angle \( 0 \leq \gamma < 2\pi \) which brings the line of nodes \( On \) along the direction \( Ov \).

All in all, one may label \( R \) by the three Euler\(^{(v)}\) angles \( \alpha, \beta, \gamma \) and write

\[
R(\alpha, \beta, \gamma) = R_w(\gamma)R_n(\beta)R_z(\alpha).
\]  

(VI.1a)

One can actually express \( R_w(\gamma) \) in terms of \( R_z(\gamma) \) and \( R_n(\beta) \) in terms of \( R_y(\beta) \), and thereby derive the alternative decomposition

\[
R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma),
\]  

(VI.1b)

which only involves rotations about the “old” axes.

Remarks:

* The convention for the Euler angles is far from being universal. Here we introduced the choice which is the most usual in classical mechanics or quantum mechanics, and which leads to the “z-y-z-convention” of Eq. (VI.1b), which we shall encounter again in § VI.3.2b.

* The three Euler angles take their values in a bounded region of the parameter space \( \mathbb{R}^3 \), which shows that the group of rotations of \( \mathbb{R}^3 \) is a three-parameter compact group.

VI.1.1b Parameterization through a unit vector and an angle

The other usual description of rotations consists in using a unit vector \( \vec{n} \) along the rotation axis together with the angle \( \psi \) of the rotation, where the latter can be taken in the interval \([0, 2\pi]\) or equivalently \([-\pi, \pi]\). Accordingly, a given rotation is then denoted \( R_{\vec{n}}(\psi) \).

Remark VI.2. Obviously, a given rotation can be described by two equivalent sets consisting of a unit vector and an angle:

\[
R_{\vec{n}}(\psi) = R_{-\vec{n}}(-\psi),
\]  

(VI.2)

as can also be checked on Eq. (VI.3) below. One may reduce the interval in which \( \psi \) takes its values

\(^{(v)}\)L. Euler, 1707–1783
to $0 \leq \psi \leq \pi$ to decrease this redundancy, yet it cannot be totally removed since there still remains $\mathcal{R}_n(\pi) = \mathcal{R}_{-n}(\pi)$ for any unit vector $\vec{n}$.

Now, since a unit vector of $\mathbb{R}^3$ is parameterized through 2 real parameters — for instance, identifying $\vec{n}$ with a point on the unit sphere $S^2$, by the polar angle $\theta \in [0, \pi]$ and azimuthal angle $\varphi \in [0, 2\pi]$ of the spherical-coordinates system —, we recover the fact that the rotations form a three-parameter group.

The effect on a vector $\vec{x} \in \mathbb{R}^3$ of the rotation through the angle $\psi$ about the direction along the unit vector $\vec{n}$ is given by Rodrigues’ formula:

$$\mathcal{R}_n(\psi) \vec{x} = (\cos \psi) \vec{x} + (1 - \cos \psi)(\vec{n} \cdot \vec{x})\vec{n} + (\sin \psi)\vec{n} \times \vec{x}.$$  \hspace{1cm} (VI.3)

The formula follows from writing $\vec{x} = \vec{x} + \vec{x}$, where $\vec{x}$ is the projection of $\vec{x}$ parallel to $\vec{n}$, and realizing that $\mathcal{R}_n(\psi) \vec{x}$ leaves $\vec{x}$ invariant while transforming $\vec{x}$ into $(\cos \psi) \vec{x} + (\sin \psi)\vec{n} \times \vec{x}$.

In the case of an infinitesimally small angle $|\delta \psi| \ll 1$, a Taylor expansion of the Rodrigues formula to first order gives

$$\mathcal{R}_n(\delta \psi) \vec{x} = \vec{x} + \delta \psi \vec{n} \times \vec{x} + \mathcal{O}(\delta \psi^2).$$  \hspace{1cm} (VI.4)

The reader is invited to check the validity of this relation by drawing $\vec{x}$ and $\mathcal{R}_n(\delta \psi) \vec{x}$ in the plane orthogonal to the direction of $\vec{n}$.

The description by a unit vector and an angle is especially useful for expressing the rotation conjugate to another rotation $\mathcal{R}_n(\psi)$. For any rotation $\mathcal{R}$, one checks the relation

$$\mathcal{R} \mathcal{R}_n(\psi) \mathcal{R}^{-1} = \mathcal{R}_n(\psi) \quad \text{with} \quad \vec{n}' = \mathcal{R}\vec{n}.$$  \hspace{1cm} (VI.5)

Accordingly, two conjugate rotations have the same rotation angle. Conversely, two rotations through the same angle $\psi$ are conjugate, the conjugating element being the rotation that maps the unit vector of the first one to the unit vector of the second one. We thus recover the result already found in §I.4.3c, according to which a conjugacy class of the group of rotations in three-dimensional space consists of all rotations with a given angle.

### VI.1.2 The group SO(3) and its Lie algebra $\mathfrak{so}(3)$

Once a Cartesian coordinate system has been fixed, the rotations in three-dimensional space preserving a given point are represented by $3 \times 3$ real matrices $O$ such that $O^T O = O O^T = I_3$ and with determinant 1, i.e. by the matrices of the group $\text{SO}(3)$. Hereafter, we shall denote these matrices with the same symbol $\mathcal{R}$ as the corresponding rotations.

**Remark:** According to the general formula given in §V.1.2c, $\text{SO}(3)$ is a Lie group of dimension 3, in agreement with what was found above for the group of rotations in three-dimensional Euclidean space.

### VI.1.2a Generators of $\mathfrak{so}(3)$

As stated in §V.2.3c, the Lie algebra $\mathfrak{so}(3)$ consists of the antisymmetric real $3 \times 3$ matrices. That is, a rotation through an infinitesimally small angle $|\delta \psi| \ll 1$ about the direction along the unit vector $\vec{n}$ can be written as

$$\mathcal{R}_n(\delta \psi) = I_3 - i \delta \psi J_{\vec{n}} + \mathcal{O}(\delta \psi^2)$$  \hspace{1cm} (VI.6)

where the matrix $-iJ_{\vec{n}}$ is real and antisymmetric — and where the physicists’ convention (nota-
The groups SO(3) and SU(2) and their representations

V.29 was used. In turn, the rotation about \( \vec{n} \) through a finite angle \( \psi \) reads

\[
\mathcal{R}_\vec{n}(\psi) = e^{-i \psi J_\vec{n}},
\]

as follows from the identity \( \lim_{p \to \infty} \left( 1 + \frac{x}{p} \right)^p = e^x \).

Considering the infinitesimal rotations about the three axes of the coordinate system, i.e. about the unit vectors \( \vec{e}_1, \vec{e}_2, \vec{e}_3 \), one finds a basis of \( \mathfrak{so}(3) \), namely

\[
-i J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad
-i J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad
-i J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

(VI.8)

Considering for instance the well-known matrix of a rotation around the third coordinate axis, \( \mathcal{R}_3(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \), and invoking Eq. (V.33), one indeed finds

\[
\lim_{\psi \to 0} \mathcal{R}_3(\psi) - I_3 = -i J_3.
\]

The corresponding matrices \( J_k \) — the generators, in the physicists’ convention — are given by

\[
J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad
J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad
J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\]

(VI.9a)

i.e. the \( ij \) entry of the matrix \( J_k \) reads

\[
(J_k)_{ij} = -i \epsilon_{ijk}
\]

(VI.9b)

where \( \epsilon_{ijk} \) is the usual totally antisymmetric Levi-Civita symbol with \( \epsilon_{123} = 1 \).

Remark VI.10. The generators (VI.9) obey the relation

\[
\text{Tr}(J_i J_j) = 2 \delta_{ij} \quad \text{for all } i,j \in \{1, 2, 3\},
\]

(VI.10)

which can be viewed as a normalization condition on the generators.

VI.1.2b Commutation relations of \( \mathfrak{so}(3) \)

Calculating the various commutators of the generators (VI.9), one finds

\[
[J_1, J_2] = i J_3, \quad [J_2, J_3] = i J_1, \quad [J_3, J_1] = i J_2.
\]

(VI.11a)

These three equations, as well as the trivial commutator of every matrix \( J_i \) with itself, can be gathered together in the form of the single equation

\[
[J_i, J_j] = i \sum_{k=1}^3 \epsilon_{ijk} J_k \quad \text{for all } i,j \in \{1, 2, 3\}.
\]

(VI.11b)

One sees that the commutator of two different generators is non-zero, which is equivalent to the fact that rotations about different axes do not commute.

Comparing the Lie bracket (VI.11) with Eq. (V.29) shows that the structure constants of the Lie algebra \( \mathfrak{so}(3) \) are \( f^{ij}_k = \epsilon_{ijk} \) for generators obeying the normalization condition (VI.10). In the

(30) Yet with the notation \( J_k \) instead of \( T^a \).

(x) T. Levi-Civita, 1873–1941
present case, the Jacobi identity (V.30b) becomes
\[
\sum_{k=1}^{3} \left( \epsilon_{ijk} \epsilon_{klm} + \epsilon_{jik} \epsilon_{kjm} + \epsilon_{lik} \epsilon_{kjm} \right) = 0,
\]
(VI.12)
as can be checked using the possibly more familiar identity
\[
\sum_{k=1}^{3} \epsilon_{kij} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}
\]
and the property \( \epsilon_{kij} = \epsilon_{ijk} \).

**VI.1.2 Generator of the rotations about an arbitrary axis**

Comparing now the infinitesimal version (VI.4) of Rodrigues’ formula and Eq. (VI.6), one sees that the generator \( J_\vec{n} \) is such that 
\[-iJ_\vec{n} \vec{x} = \vec{n} \times \vec{x} \]
for all \( \vec{x} \in \mathbb{R}^3 \). That is, \(-iJ_\vec{n}\) is a \( 3 \times 3 \) matrix whose \( ij \) element obeys
\[-i(J_\vec{n})_{ij} = \sum_{k=1}^{3} \epsilon_{ikj} n_k = -\sum_{k=1}^{3} \epsilon_{ijk} n_k,
\]
where the second identity uses the antisymmetry of the Levi-Civita symbol while the \( n_k \) are the components of the unit vector \( \vec{n} \). Invoking Eq. (VI.9b), one may now write
\[-i(J_\vec{n})_{ij} = \sum_{k=1}^{3} [-i(J_k)_{ij}] n_k
\]
i.e.
\[J_\vec{n} = \sum_{k=1}^{3} n_k J_k \equiv \vec{n} \cdot \vec{J}.
\]
(VI.13)

In the rightmost term of this equation, the dot product is not a scalar product, but merely a convenient notation, since \( \vec{n} \) is a “normal” vector of \( \mathbb{R}^3 \) while \( \vec{J} \) is a shorthand notation for the triplet consisting of the three matrices \( J_1, J_2, J_3 \). Equation (VI.13) illustrates the fact that the latter form a basis of generators, \( J_\vec{n} \) being the generator of rotations about the \( \vec{n} \)-direction.

Inserting Eq. (VI.13) into Eq. (VI.7), the rotation through \( \psi \) about the direction \( \vec{n} \) may be written as
\[\mathcal{R}_\vec{n}(\psi) = e^{-i\psi(n_1 J_1 + n_2 J_2 + n_3 J_3)}.
\]
(VI.14)

Since the generators \( J_k \) do not commute with each other, see Eqs. (VI.11), the exponential of a sum of matrices on the right hand side of this equation cannot be recast as a product of exponentials.

**Remarks:**

* The property \( \mathcal{R}_\vec{n}(\psi) = \mathcal{R}_{-\vec{n}}(-\psi) \) can again be checked on Eq. (VI.14).

* Parameterizing a rotation through Euler angles, one can use the decomposition (VI.1b) and write every rotation about the \( z \)- or \( y \)-direction in exponential form, which leads to
\[\mathcal{R}(\alpha, \beta, \gamma) = e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3}.
\]
(VI.14)

This time, one may not rewrite the product of exponentials as the exponential of the sum of the exponents.
VI.2 The group SU(2)

The group SU(2), consisting of the $2 \times 2$ complex matrices $U$ such that $U^\dagger U = UU^\dagger = \mathbb{I}_2$ and $\det U = 1$, is like SO(3) a three-parameter Lie group. After recalling a few results on SU(2), we shall study its Lie algebra (Sec. VI.2.1), and then discuss the relation between SU(2) and SO(3) in further detail (Sec. VI.2.2).

VI.2.1 The group SU(2) and its Lie algebra $\mathfrak{su}(2)$

VI.2.1 a Parameterization of a generic element of SU(2)

A generic element of SU(2) is of the form

$$U = \begin{pmatrix} u_4 + i u_3 & u_2 + i u_1 \\ u_2 - i u_1 & u_4 - i u_3 \end{pmatrix} \quad \text{with} \quad \begin{cases} u_1, u_2, u_3, u_4 \in \mathbb{R} \\ |u_1|^2 + |u_2|^2 + |u_3|^2 + |u_4|^2 = 1 \end{cases} \quad (VI.15)$$

Conversely, one checks that every matrix of this form is in SU(2), which shows the existence of a bijection between SU(2) and the unit sphere $S^3 \subset \mathbb{R}^4$. With the help of the Pauli matrices $\sigma_i$ (VI.16),

$$\begin{align*} 
\sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\end{align*} \quad (VI.16)$$

Eq. (VI.15) takes the form

$$U = u_4 \mathbb{I}_2 + i (u_1 \sigma_1 + u_2 \sigma_2 + u_3 \sigma_3) \equiv u_4 \mathbb{I}_2 + i \vec{u} \cdot \vec{\sigma}, \quad (VI.17)$$

where $\vec{u}$ resp. $\vec{\sigma}$ is a “vector” whose entries are the real numbers $u_1$, $u_2$, $u_3$ resp. the three Pauli matrices (VI.16), while their dot product is a shorthand notation — similar to that introduced in Eq. (VI.13).

Remark: The identity

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{I}_2 + i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k \quad \text{for all} \quad i, j \in \{1, 2, 3\} \quad (VI.18)$$

shows that the product of two matrices of the form (VI.17) is again of the same form — i.e. that the matrix multiplication is an internal composition law on SU(2).

Introducing the unit vector $\vec{n} \equiv -\vec{u}/|\vec{u}|$ and the “angle” $\psi \in [0, \pi]$ defined by $\cos \frac{\psi}{2} = u_4$ and $\sin \frac{\psi}{2} = |\vec{u}|$, the term on the right hand side of Eq. (VI.17) becomes

$$\left( \cos \frac{\psi}{2} \right) \mathbb{I}_2 - i \left( \sin \frac{\psi}{2} \right) \vec{n} \cdot \vec{\sigma} = e^{-i \psi \vec{n} \cdot \vec{\sigma}} / 2, \quad (VI.17)$$

where the identity follows from the definition of the exponential of a matrix under consideration of the identity $(\vec{n} \cdot \vec{\sigma})^2 = \mathbb{I}_2$. Denoting $U_{\vec{n}}(\psi)$ this matrix of SU(2), which is obviously entirely determined by the angle $\psi$ and the unit vector $\vec{n}$, we may thus write

$$U_{\vec{n}}(\psi) = e^{-i \psi \vec{n} \cdot \vec{\sigma}} / 2 = e^{-i \psi (n_1 \sigma_1 + n_2 \sigma_2 + n_3 \sigma_3) / 2}. \quad (VI.19)$$

The reader is invited to note the similarity between this expression for a generic matrix of SU(2), parameterized by $\psi \in [0, 2\pi]$ and $\vec{n} \in S^2$, and Eq. (VI.14).

VI.2.1 b Generators of $\mathfrak{su}(2)$

The Lie algebra $\mathfrak{su}(2)$ consists of the traceless antihermitian $2 \times 2$ matrices. A possible basis of generators, which is suggested by Eq. (VI.19), consists of the Pauli matrices. Conventionally, one
rather takes as generators the matrices \( T_i = \frac{1}{2} \sigma_i \) for \( i = 1, 2, 3 \):

\[
T^1 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad T^2 = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}, \quad T^3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{\sigma_3}{2}.
\]

(VI.20)

Denoting \( U_i(\psi) \) with \( i \in \{1, 2, 3\} \) the element of SU(2) obtained by taking the unit vector \( \vec{n} \) with components \( n_j = \delta_{ij} \) in Eq. (VI.19), one has

\[
\lim_{\psi \to 0} \frac{U_i(\psi) - I_2}{\psi} = \frac{\sigma_i}{2},
\]

which under consideration of Eq. (V.33) leads to the generators (VI.20).

Equation (VI.18) translates at once into

\[
T^i T^j = \delta_{ij} I_2 + \frac{i}{2} \sum_{k=1}^{3} \epsilon_{ijk} T^k \quad \text{for all } i, j \in \{1, 2, 3\},
\]

from which one deduces two properties of the generators (VI.20). On the one hand, they obey the "normalization"

\[
\text{Tr}(T^i T^j) = \frac{\delta_{ij}}{2} \quad \text{for all } i, j \in \{1, 2, 3\}.
\]

(VI.21)

On the other hand, their commutator is given by

\[
[T^i, T^j] = i \sum_{k=1}^{3} \epsilon_{ijk} T^k \quad \text{for all } i, j \in \{1, 2, 3\}.
\]

(VI.22)

The structure constants of the Lie algebra \( \mathfrak{su}(2) \) are thus \( f^{ij}_k = \epsilon_{ijk} \) — like the structure constants of \( \mathfrak{so}(3) \). Accordingly, any calculation carried out with the generators \( J_i \) [Eq. (VI.9)] or \( T^i \) [Eq. (VI.20)] which only involves commutation relations will hold for both Lie algebras. We shall exploit this result in Sec. VI.3 when discussing the representations of the Lie groups SO(3) and SU(2).

**VI.2.2 Relation between the Lie groups SU(2) and SO(3)**

Two-to-one group homomorphism from SU(2) to SO(3) — see exercise 38.iii

Kernel of the homomorphism is \( \{1_2, -1_2\} \cong \mathbb{Z}_2 \): isomorphism between \( SU(2)/\mathbb{Z}_2 \) and SO(3)

Both groups are connected. In addition, SU(2) is simply connected\(^{31}\) while SO(3) is not. SU(2) is the universal cover of SO(3) — i.e. SU(2) is the *spin group* Spin(3).

\(^{31}\)That is, all closed paths on SU(2) can be continuously contracted to a point.
VI.3 Representations of the Lie groups SU(2) and SO(3)

The reader has probably already encountered various physical objects whose mathematical representations transform differently under the rotations of three-dimensional space.

- scalar quantities, which remain invariant under rotations and are represented by a single number, as e.g. the electric charge or a scalar field like temperature;
- "vector" quantities, represented by a triplet of real numbers that transform like the Cartesian coordinates $x^i$ of the position, i.e.
  \[ V^{i'} = \sum_{i=1}^{3} \mathcal{R}^{i'}_i V^i \quad \text{for } i' \in \{1, 2, 3\}, \]  
  (VI.23)
  where the $\mathcal{R}^{i'}_i$ are the elements of a matrix of SO(3);
- "tensor" quantities, represented by a collection of real numbers $T^{i_1...i_s}$ labeled by two or more indices taking their values in the set $\{1, 2, 3\}$, such that each index of the tensor transforms like a "vector" index under rotations:
  \[ T^{i'_1...i'_s} = \sum_{i_1=1}^{3} \ldots \sum_{i_s=1}^{3} \mathcal{R}^{i'_1}_{i_1} \ldots \mathcal{R}^{i'_s}_{i_s} T^{i_1...i_s} \quad \text{for } i'_1, \ldots, i'_s \in \{1, 2, 3\}. \]  
  (VI.24)

For instance, a tensor of rank 2 like the inertia tensor consists of 9 real numbers.

In each of these cases, the mathematical objects (scalars, vectors, tensors) are elements of various vector spaces with dimension 1, 3 or more, which are in fact different representation spaces of the group of rotations of three-dimensional space — or equivalently of SO(3).

In this section, we systematically investigate the representations of SO(3) and of the companion group SU(2). Since both groups are compact Lie groups, we shall only study unitary representations, anticipating on results detailed in ??.. After arguing why it is convenient to focus on representations of the Lie algebra $\mathfrak{su}(2)$ (Sec. VI.3.1), we derive in Sec. VI.3.2 the irreducible representations of SU(2) — and in passing of SO(3). We then consider the direct product of such irreps. (Sec. VI.3.3).

Throughout the section, we shall often use the bra-ket notation of quantum mechanics. More generally, the reader will probably recognize results from the formalism of angular momentum a.k.a. "spin" in quantum mechanics.

VI.3.1 Representations of SU(2) and of the Lie algebra $\mathfrak{su}(2)$

Since SU(2) and SO(3) are connected Lie groups, they are entirely determined, through exponentiation, by their respective Lie algebras. Generalizing this idea, we argue that representations of the Lie groups can be deduced from representations of the Lie algebras, and more specifically of $\mathfrak{su}(2)$ (§ VI.3.1a), after which we discuss a few obvious representations (§ VI.3.1b).

VI.3.1 a Representations of the Lie algebra $\mathfrak{su}(2)$

Consider the linear mapping $\phi$ from $\mathfrak{su}(2)$ to $\mathfrak{so}(3)$, seen as vector spaces, that maps each generator $T^i \in \mathfrak{su}(2)$ to the generator $J_i \in \mathfrak{so}(3)$ for $i = 1, 2, 3$. This mapping is clearly bijective, for it maps a basis of the first space to a basis of the second one. Since $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ have the same structure constants $f^{ijk} = \epsilon_{ijk}$, see Eqs. (VI.11) and (VI.22), the mapping $\phi$ will also preserve the Lie brackets. All in all, $\psi$ is thus a Lie algebra isomorphism, i.e. $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic.

---

(32) We omit here the discussion of the transformation property under spatial parity, i.e. the point reflection $\vec{r} \to -\vec{r}$, since the latter is not in SO(3), but rather in O(3).

(33) In the case of a field, one has a single number at each point in space (and time).
VI.3 Representations of the Lie groups SU(2) and SO(3)

Therefore, each representation \( \hat{\mathcal{D}} \) of \( \mathfrak{su}(2) \) provides a representation \( \hat{\mathcal{D}} \circ \phi^{-1} \) of \( \mathfrak{so}(3) \) — and conversely every representation of \( \mathfrak{so}(3) \) is of this form. Accordingly, we shall from now on talk of representations of \( \mathfrak{su}(2) \), knowing that they may be used for \( \mathfrak{so}(3) \) as well.

A representation of dimension \( d \) of the Lie algebra \( \mathfrak{su}(2) \) is a linear mapping from \( \mathfrak{su}(2) \) to a matrix Lie algebra, whose elements are \( d \times d \) matrices, such that it preserves the commutator of two arbitrary matrices. Accordingly, that mapping should also preserve the commutator of the images of the generators \( T^a \) of \( \mathfrak{su}(2) \), which we shall denote \( \mathcal{J} \). That is, we shall look for a triplet of \( d \times d \) Hermitian matrices \( \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3 \) obeying

\[
[\mathcal{J}_i, \mathcal{J}_k] = i \sum_{k=1}^{3} \epsilon_{ikl} \mathcal{J}_l \quad \text{for all } i, k \in \{1, 2, 3\}.
\]

By exponentiation, the generators \( \{ \mathcal{J}_i \} \) generate unitary \( d \times d \) matrices

\[
\mathcal{U}(\theta_1, \theta_2, \theta_3) = e^{-i(\theta_1 \mathcal{J}_1 + \theta_2 \mathcal{J}_2 + \theta_3 \mathcal{J}_3)}
\]

that constitute a unitary representation of dimension \( d \) of SU(2).

Remark: Beware of the different meanings of “dimension”, which should not be confused with each other. On the one hand, there is the dimension of the Lie group, or of its Lie algebra, that is the number of real continuous parameters of the group (Sec. V.1.1), namely 3 in the case of \( \mathfrak{su}(2) \). This is also the number of generators of the Lie group/algebra. On the other hand, there is the dimension of the representation space for a given representation of the Lie group (see definition II.5), or of its Lie algebra, which was denoted \( d \) in the previous discussion.

VI.3.1b Low-dimensional representations of SU(2)

Before investigating the properties of an arbitrary irreducible representation of SU(2), let us note that we already know three of them.

First, there is the defining representation, in which every element \( U \in \text{SU}(2) \) is represented by itself! The corresponding generators are then the matrices \( \{ T^a \} \) introduced in Eq. (VI.20).

Since SU(2) consists of \( 2 \times 2 \) matrices, this is a two-dimensional representation, very often referred to as the “2” representation of SU(2) — as we shall see below, it is also called the “spin-\( \frac{1}{2} \)” representation and it is irreducible. Anticipating on the terminology introduced in ??, this is also the fundamental representation of SU(2).

Remark: As we shall discuss in § VI.3.2b the 2-representation of SU(2) is not a linear representation of SO(3).

According to Sec. VI.2.2 another representation of SU(2) is that which maps the unitary matrix \( \mathcal{U}_q(\psi) \in \text{SU}(2) \) to the orthogonal matrix \( \mathcal{A}_q(\psi) \in \text{SO}(3) \). This three-dimensional irreducible representation is the “3” representation of SU(2). Following the terminology of ??, this is the adjoint representation of SU(2) — and of SO(3) — since it is generated by the structure constants, see Eq. (VI.9b).

Note that as a representation of SU(2), it is not faithful, since the mapping SU(2) \( \rightarrow \) SO(3) is not injective.

Remark: From the point of view of SO(3), the 3-representation is also the defining and the fundamental representation.

Eventually, the “1” representation of SU(2) or SO(3) is the trivial representation, for which there is little to mention apart from the fact that the three corresponding generators are all equal to the \( 1 \times 1 \) matrix \( (0) \) — which is the only way to fulfill the commutation relation (VI.25) with such commuting matrices.
VI.3.2 Unitary representations of SU(2)

Let us now come to the general case of a unitary representation of SU(2) of arbitrary dimension \( d > 1 \) generated by Hermitian matrices \( J_1, J_2, J_3 \) obeying relation (VI.25). Instead of directly searching for the generators, we shall follow the classical construction and investigate the eigenvalues and eigenvectors of one of them (§ VI.3.2a), and thereafter work in the basis of these eigenvectors.

VI.3.2 a Diagonalization of the generators of SU(2)

Since two different generators \( J_i, J_k \) do not commute, they cannot be simultaneously diagonalizable. That is, in any given basis at most one generator can be diagonal\(^{(34)}\) say \( J_3 \), which we shall also denote \( J_z \).

As is customary, we introduce an auxiliary \( d \times d \) matrix, function of the generators \( \{J_i\} \), namely the (quadratic) Casimir element \( \vec{J}^2 \equiv J_1^2 + J_2^2 + J_3^2 \). (VI.26)

Since every generator \( J_i \) is Hermitian, so is \( \vec{J}^2 \). In addition, one quickly checks that the commutation relations (VI.25) lead to

\[
\left[ \vec{J}^2, J_i \right] = 0 \quad \forall i \in \{1, 2, 3\},
\]

(VI.27)

where 0 here denotes the zero \( d \times d \) matrix.

Remark: As we shall show below, all eigenvalues of the Casimir element are non-negative, so that \( \vec{J}^2 \) is not traceless. Therefore it does not belong to the Lie algebra generated by the \( \{J_i\} \), but to a larger algebra.

Eigenvectors of \( \vec{J}^2 \) and \( J_z \)

Since \( \vec{J}^2 \) and \( J_z \) commute, they are simultaneously diagonalizable, i.e. define a basis of simultaneous eigenvectors. We momentarily denote the latter \( \{|\lambda, m\rangle\} \), with

\[
\vec{J}^2 |\lambda, m\rangle = \lambda |\lambda, m\rangle, \quad J_z |\lambda, m\rangle = m |\lambda, m\rangle.
\]

(VI.28)

Since \( \vec{J}^2 \) and \( J_z \) are Hermitian, their eigenvalues \( \lambda \) and \( m \) are real numbers. In turn, the eigenvectors are \( d \)-dimensional complex vectors, i.e. elements of \( \mathbb{C}^d \), which are pairwise orthogonal for the Hermitian scalar product. We further assume that they are normalized to unity: \( |||\lambda, m||\rangle = 1 \).

From definition (VI.26) and the hermiticity of the generators follows that \( \lambda \) is non-negative. Indeed, one may consider scalar products with \( \langle \lambda, m \mid \vec{J}^2 |\lambda, m\rangle \) and write on the one hand

\[
\langle \lambda, m \mid \vec{J}^2 |\lambda, m\rangle = \sum_{i=1}^{3} \langle \lambda, m \mid J_i^2 |\lambda, m\rangle = \sum_{i=1}^{3} \langle \lambda, m \mid J_i^\dagger J_i |\lambda, m\rangle = \sum_{i=1}^{3} |||J_i|\lambda, m||\rangle^2 \geq 0,
\]

and on the other hand, using Eq. (VI.28) and the assumed normalization,

\[
\langle \lambda, m \mid \vec{J}^2 |\lambda, m\rangle = \lambda \langle \lambda, m |\lambda, m\rangle = \lambda.
\]

Both equations together then yield \( \lambda \geq 0 \). Since the mapping \( j \in \mathbb{R}^+ \mapsto j(j+1) \in \mathbb{R}^+ \) is bijective, one may thus write \( \lambda \equiv j(j+1) \) with \( j \geq 0 \). Replacing the notation \( |\lambda, m\rangle \) by \( |j, m\rangle \), the defining relations (VI.28) become

\[
\vec{J}^2 |j, m\rangle = j(j+1) |j, m\rangle \quad (\text{VI.29a})
\]
\[
J_z |j, m\rangle = m |j, m\rangle \quad (\text{VI.29b})
\]

\(^{(34)}\) Anticipating on the terminology of ??, SU(2) is said to be of rank 1.

\(^{(3)}\) H. Casimir, 1909–2000
with \( j \in \mathbb{R}_+ \) and \( m \in \mathbb{R} \). In turn, the orthonormality of the eigenvectors reads
\[
\langle j', m' | j, m \rangle = \delta_{j,j'} \delta_{m,m'} \quad \text{for all} \quad j, j' \in \mathbb{R}_+, m, m' \in \mathbb{R}.
\] (VI.30)

**“Ladder operators”**

With the help of the non-diagonal generators \( \mathcal{J}_1 \) and \( \mathcal{J}_2 \), one defines two \( d \times d \) matrices by
\[
\mathcal{J}_+ \equiv \mathcal{J}_1 + i \mathcal{J}_2 \quad \text{and} \quad \mathcal{J}_- \equiv \mathcal{J}_1 - i \mathcal{J}_2.
\] (VI.31)

In contrast to \( \mathcal{J}_1 \) and \( \mathcal{J}_2 \), the matrices \( \mathcal{J}_\pm \) are not Hermitian, but obey
\[
\mathcal{J}_+ \equiv \mathcal{J}_- , \quad \mathcal{J}_- \equiv \mathcal{J}_+.
\] (VI.32)

Using Eq. (VI.26), one first finds that the matrices \( \mathcal{J}_+ \) and \( \mathcal{J}_- \) commute with \( \mathcal{J}^2 \):
\[
[\mathcal{J}^2, \mathcal{J}_\pm] = [\mathcal{J}^2, \mathcal{J}_\mp] = 0,
\] but do not commute with \( \mathcal{J}_z \). Writing \([\mathcal{J}_z, \mathcal{J}_\pm] = [\mathcal{J}_z, \mathcal{J}_1] \pm i[\mathcal{J}_z, \mathcal{J}_2] = i \mathcal{J}_2 \pm i(-i \mathcal{J}_1)\), one eventually obtains
\[
[\mathcal{J}_z, \mathcal{J}_\pm] = ± \mathcal{J}_±.
\] (VI.34)

The commutation relations (VI.33), (VI.34) allow one to derive properties of the vectors \( \mathcal{J}_+ |j, m\rangle \) and \( \mathcal{J}_- |j, m\rangle \). Thus, one first finds
\[
\mathcal{J}_\pm |j, m\rangle = ± \mathcal{J}_\pm^2 |j, m\rangle = ± (j(j+1) |j, m\rangle) = j(j+1) \mathcal{J}_\pm |j, m\rangle.
\] This identity means that each vector \( \mathcal{J}_\pm |j, m\rangle \), \( \mathcal{J}_\pm |j, m\rangle \) is either eigenvector of the Casimir element \( \mathcal{J}^2 \) with the eigenvalue \( j(j+1) \), or equals the zero vector \( |\varnothing\rangle \).

In addition, Eq. (VI.34) yields
\[
\mathcal{J}_z \mathcal{J}_± |j, m\rangle = [\mathcal{J}_z, \mathcal{J}_±] |j, m\rangle + \mathcal{J}_± \mathcal{J}_z |j, m\rangle = ± \mathcal{J}_± |j, m\rangle + \mathcal{J}_± \mathcal{J}_z |j, m\rangle.
\] Using Eq. (VI.29a), the second term on the right hand side equals \( m \mathcal{J}_± |j, m\rangle \), which leads to
\[
\mathcal{J}_z \mathcal{J}_± |j, m\rangle = (m ± 1) \mathcal{J}_± |j, m\rangle.
\] (VI.35)

When it differs from the zero vector \( |\varnothing\rangle \), \( \mathcal{J}_± |j, m\rangle \) resp. \( \mathcal{J}_± |j, m\rangle \) is eigenvector of \( \mathcal{J}_z \) with the eigenvalue \( m+1 \) resp. \( m-1 \). The matrix \( \mathcal{J}_+ \) resp. \( \mathcal{J}_- \) thus increases resp. decreases the eigenvalue associated to the generator \( \mathcal{J}_z \) by 1, hence its denomination creation resp. annihilation operator in quantum-mechanical applications — together, they are referred to as the ladder operators.

**Eigenvalues of \( \mathcal{J}^2 \) and \( \mathcal{J}_z \)**

For any eigenvector \( |j, m\rangle \), the square norm \( \| \mathcal{J}_± |j, m\rangle \|^2 \) must be non-negative. Using the Hermitian conjugation (VI.32), this equals \( \langle j, m | \mathcal{J}_± \mathcal{J}_± |j, m\rangle \). Invoking the definitions (VI.31), the product \( \mathcal{J}_± \mathcal{J}_± \) can be rewritten as
\[
\mathcal{J}_± \mathcal{J}_± = (\mathcal{J}_± \mp i \mathcal{J}_y) (\mathcal{J}_± \pm i \mathcal{J}_y) = \mathcal{J}_±^2 + \mathcal{J}_y^2 \mp i[\mathcal{J}_z, \mathcal{J}_y].
\]

With the commutation relation (VI.25), this becomes
\[
\mathcal{J}_± \mathcal{J}_± = \mathcal{J}_±^2 + \mathcal{J}_±^2 \mp i(\mathcal{J}_z) = \mathcal{J}_±^2 - \mathcal{J}_±^2 \mp \mathcal{J}_z,
\] (VI.36)
so that the matrix element \( \langle j, m | \mathcal{J}_± \mathcal{J}_± |j, m\rangle \) takes the form
\[
\langle j, m | \mathcal{J}_± \mathcal{J}_± |j, m\rangle = \langle j, m | \mathcal{J}_±^2 |j, m\rangle - \langle j, m | \mathcal{J}_±^2 |j, m\rangle \mp \langle j, m | \mathcal{J}_z |j, m\rangle.
\]

Every term on the right hand side can be computed using Eqs. (VI.29) and the normalization of the eigenvector \( |j, m\rangle \):
\[
\langle j, m | \mathcal{J}_±^2 |j, m\rangle = \langle j, m | \mathcal{J}_±^2 |j, m\rangle \mp \langle j, m | \mathcal{J}_z |j, m\rangle = j(j+1) - m^2 \mp m = j(j+1) - m(m \pm 1),
\]
yielding eventually
\[ |\mathcal{J}_\pm |j, m\rangle|^2 = j(j+1) - m(m+1). \]  
(VI.37)
This is non-negative if and only if \(|m| \leq j\).

Let us now fix \(j \in \mathbb{R}_+\) and investigate the possible values of the eigenvalue \(m\). Let \(m_{\text{max}}\) be the largest one, which exists thanks to the condition \(|m| \leq j\); by definition, the corresponding eigenvector \(|j,m_{\text{max}}\rangle\) is non-zero. Invoking Eq. (VI.35) (with \(\mathcal{J}_+\)), the vector \(\mathcal{J}_+ |j,m_{\text{max}}\rangle\) must be the zero vector \(|\emptyset\rangle\) — otherwise it would be eigenvector of \(\mathcal{J}_-\) with the eigenvalue \(m_{\text{max}} + 1\), contradicting the definition of \(m_{\text{max}}\). Considering Eq. (VI.37) then yields \(m_{\text{max}} = j\), and thus the existence of the eigenvector \(|j, j\rangle\), characterized by

\[ |j,j\rangle \neq |\emptyset\rangle \quad \text{and} \quad \mathcal{J}_+ |j,j\rangle = |\emptyset\rangle. \]  
(VI.38a)
Similarly, one shows using \(\mathcal{J}_-\) that the smallest possible eigenvalue of \(m\) at given \(j\) is \(-j\), with a corresponding eigenvector obeying

\[ |j,-j\rangle \neq |\emptyset\rangle \quad \text{and} \quad \mathcal{J}_- |j,-j\rangle = |\emptyset\rangle. \]  
(VI.38b)

Starting now from the eigenvector \(|j,j\rangle\) and operating recursively with the matrix \(\mathcal{J}_-\), one finds for every \(p \in \mathbb{N}\)

\[ (\mathcal{J}_-)^p |j,j\rangle \propto |j,j-p\rangle, \]  
(VI.39)
with the usual convention that \((\mathcal{J}_-)\)^0 is the identity matrix. To ensure that the eigenvalue \(m = j-p\) remains bounded from below, \(j-p\) must for some value \(p_{\text{max}}\) equal the lower bound \(-j\), which gives \(2j = p_{\text{max}}\), i.e. \(2j \in \mathbb{N}\). Stated differently, the real number \(j\) characterizing the eigenvalue of the Casimir element must be an integer or a half-integer:

\[ j \in \{0, 1, 1/2, 3/2, \ldots \}. \]  
(VI.40a)
The other eigenvalue \(m\) then takes all values between \(-j\) and \(j\) in steps of 1:

\[ m \in \{-j, -j+1, \ldots, j-1, j\}, \]  
(VI.40b)
i.e. all in all \(2j+1\) different values.

Remark: According to Eq. (VI.35), \(\mathcal{J}_+ |j,m\rangle\) resp. \(\mathcal{J}_- |j,m\rangle\) is either the zero vector or eigenvector of \(\mathcal{J}_-\) with the eigenvalue \(m+1\) resp. \(m-1\), i.e. proportional to \(|j,m+1\rangle\) resp. \(|j,m-1\rangle\). In turn, Eq. (VI.37) yields the norms of those eigenvectors, which fixes the proportionality factors to the orthonormal basis vectors up to a phase. A usual choice — in particular in quantum mechanics — for the relative phase between \(|j,m\rangle\) und \(|j,m \pm 1\rangle\) is the so-called Condon-Shortley phase convention, namely

\[ \mathcal{J}_+ |j,m\rangle = \sqrt{j(j+1) - m(m+1)} |j,m+1\rangle = \sqrt{(j-m)(j+m+1)} |j,m+1\rangle \]  
(VI.41a)
and

\[ \mathcal{J}_- |j,m\rangle = \sqrt{j(j+1) - m(m-1)} |j,m-1\rangle = \sqrt{(j+m)(j-m+1)} |j,m-1\rangle. \]  
(VI.41b)

### VI.3.2b Irreducible representations of SU(2) and SO(3)

Consider the \(2j+1\)-dimensional space \(\mathcal{V}^{(j)}\) spanned by the eigenvectors \(|j,m\rangle\) where \(m\) takes all allowed values between \(-j\) and \(j\). By definition of the eigenvectors, \(\mathcal{V}^{(j)}\) is invariant under \(\mathcal{J}_\pm = \mathcal{J}_\pm\). According to Eq. (VI.39) and to the second of Eqs. (VI.38), it also invariant under \(\mathcal{J}_-\). Eventually, \(\mathcal{V}^{(j)}\) is also invariant under \(\mathcal{J}_+\), since we have shown \(\mathcal{J}_+ |j,m\rangle \propto |j,m+1\rangle\) for all \(m\) — unless \(m = j\), in which case \(\mathcal{J}_+ |j,j\rangle \propto |\emptyset\rangle\). Being invariant under both \(\mathcal{J}_+\) and \(\mathcal{J}_-\), \(\mathcal{V}^{(j)}\) is thus invariant under \(\mathcal{J}_1\) and \(\mathcal{J}_2\), and therefore under all three generators.

\(^{(aa)}\)E. U. CONDON, 1902–1974 \(^{(ab)}\)G. H. SHORTLEY, 1910–??
Denoting $J_i^{(j)}$ the restriction of $J_i$ to $\mathcal{Y}^{(j)}$ for every $i = 1, 2, 3$, one at once sees that $J_1^{(j)}$, $J_2^{(j)}$, $J_3^{(j)}$ map $\mathcal{Y}^{(j)}$ to itself and obey the commutation relation (VI.25), i.e. are the generators of a representation of the Lie algebra $\mathfrak{su}(2)$ on $\mathcal{Y}^{(j)}$, of dimension $2j + 1$. Exponentiation then yields a (matrix) representation $\mathcal{D}^{(j)}$ of the Lie group $\text{SU}(2)$ of the same dimension, which is irreducible.

In the basis of $\mathcal{Y}^{(j)}$ consisting of the eigenvectors $\{|j, m\rangle\}$, the matrices $U^{(j)}(\theta_3) \equiv e^{-i\theta_3 J_3^{(j)}}$ — which represent the elements of $\text{SU}(2)$ of the form $e^{-i\theta_3 J_3^{(j)}}$ — are diagonal, which shows that any proper subspace of $\mathcal{Y}^{(j)}$ invariant under the $U^{(j)}(\theta_3)$ is spanned by a subset of that basis. One easily checks that such subspaces are not preserved by either $J_1^{(j)}$ or $J_2^{(j)}$ (or both), and accordingly by the matrices $e^{-i\theta_1 J_1^{(j)}}$ or $e^{-i\theta_2 J_2^{(j)}}$ of the representation. \hfill $\square$

$\mathcal{D}^{(j)}$ is called the spin $j$ representation, and often simply denoted by its dimension (in these notes, in boldface): \textbf{2} (spin $\frac{1}{2}$), \textbf{3} (spin 1), and so on… One can actually show that for every integer or half-integer $j$ there exists an irreducible representation of $\text{SU}(2)$ of dimension $2j + 1$.

**Remark:** The Casimir operator $[J_i^{(j)}]^2$ commutes by definition with all generators $J_i^{(j)}$, and will thus commute with all matrices of the form $e^{-i(\theta_1 J_1^{(j)} + \theta_2 J_2^{(j)} + \theta_3 J_3^{(j)})}$, i.e. all matrices of the spin-$j$ representation of $\text{SU}(2)$. On the other hand, the previous calculations show that $[J_i^{(j)}]^2$ is diagonal on the representation space $\mathcal{Y}^{(j)}$ — it equals $j(j + 1)$ times the identity matrix. Since the spin-$j$ representation is irreducible, one recognizes here Schur’s lemma.

Let $\mathcal{D}^{(j)}(U)$ denote the $(2j + 1) \times (2j + 1)$ unitary matrix representing $U \in \text{SU}(2)$ in the spin-$j$ representation, and $\mathcal{D}^{(j)}_{m' m}(U)$ its matrix elements in the basis of the eigenvectors $\{|j, m\rangle\}$:

$$\mathcal{D}^{(j)}(U)|j, m\rangle = \sum_{m'= -j}^{j} \mathcal{D}^{(j)}_{m' m}(U)|j, m'\rangle.$$  \hfill (VI.42)

Choosing the parameterization $U_0(\psi)$ of the $\text{SU}(2)$ matrices introduced in Eq. (VI.19), with $\vec{n}$ a unit vector of $\mathbb{R}^3$ and $\psi$ an angle, we denote the corresponding matrix in the spin-$j$ representation $\mathcal{D}^{(j)}(\vec{n}, \psi)$ and its elements $\mathcal{D}^{(j)}_{m' m}(\vec{n}, \psi)$. By definition, the latter is

$$\mathcal{D}^{(j)}_{m' m}(\vec{n}, \psi) \equiv \langle j, m'| \mathcal{D}^{(j)}(\vec{n}, \psi) | j, m\rangle.$$  \hfill (VI.43)

Consider first the case $\vec{n} = \vec{e}_3 \equiv \vec{e}_z$, yielding $\mathcal{D}^{(j)}(\vec{e}_z, \psi) = e^{-i\psi J_3^{(j)}}$. Since the basis vectors are eigenvectors of $J_3^{(j)}$ and orthonormal, one obtains at once the matrix elements of $\mathcal{D}^{(j)}(\vec{e}_z, \psi)$:

$$\mathcal{D}^{(j)}_{m' m}(\vec{e}_z, \psi) \equiv \langle j, m'| e^{-i\psi J_3^{(j)}} | j, m\rangle = \delta_{m'm} e^{-im\psi}. \hfill (VI.44)$$

In turn, setting $\vec{n} = \vec{e}_2 \equiv \vec{e}_y$ in Eq. (VI.43) defines an element of Wigner’s small $d$-matrix:

$$\mathcal{D}^{(j)}_{m' m}(\vec{e}_y, \psi) \equiv \langle j, m'| e^{-i\psi J_2^{(j)}} | j, m\rangle \equiv d^{(j)}_{m' m}(\psi). \hfill (VI.45)$$

Tabulated values of these elements (for the small values of $j$) can be found in the literature, for instance in Ref. [2], Chap. 45.\hfill (36)

**Remarks:**

* An element of Wigner’s $D$-matrix is

$$d^j_{m' m}(\alpha, \beta, \gamma) \equiv \langle j, m'| e^{-i\alpha J_3^{(j)}} e^{-i\beta J_2^{(j)}} e^{-i\gamma J_1^{(j)}} | j, m\rangle = e^{-im\alpha} d^j_{m' m}(\beta) e^{-im\gamma}, \hfill (VI.46)$$

where the matrix $e^{-i\alpha J_3^{(j)}} e^{-i\beta J_2^{(j)}} e^{-i\gamma J_1^{(j)}}$ is the $(2j + 1)$-dimensional generalization of the representation (VI.14) of a rotation in three-dimensional Euclidean space with Euler angles.

\hfill (35) as well as $i = +, -, 0$, if need be


\hfill (ac) E. P. Wigner, 1902–1995
* As the reader can check, the matrix $\mathcal{D}^{(j)}(\vec{e}_z, \psi) \equiv e^{-i\psi J_z^{(j)}}$, obtaining by letting $\vec{n} = \vec{e}_1 \equiv \vec{e}_z$ in Eq. (VI.43), can actually be expressed through the matrices $\mathcal{D}^{(j)}(\vec{e}_z, \pm \frac{\pi}{2})$ and $\mathcal{D}^{(j)}(\vec{e}_y, \psi)$. That is, $\mathcal{D}^{(j)}(\vec{e}_z, \psi)$ is entirely determined by Eqs. (VI.44) and (VI.45).

Coming back to the matrix elements (VI.44) and setting $\psi = 2\pi$, one obtains for every $j$

$$
\mathcal{D}^{(j)}_{m'm}(\vec{e}_z, 2\pi) = \delta_{m'm} e^{-2im\pi} = \delta_{m'm}(-1)^{2m}.
$$

Since $j - m \in \mathbb{N}$ for every value of $m$, $2j - 2m$ is always an even integer, so that $(-1)^{2m} = (-1)^{2j}$ for all $m$, leading to

$$
\mathcal{D}^{(j)}_{m'm}(\vec{e}_z, 2\pi) = \delta_{m'm}(-1)^{2j} \quad \text{for all } m, m' \in \{-j, -j + 1, \ldots, -1, j\}. \quad (VI.47a)
$$

That is, the $(2j + 1)$-dimensional matrix representing the element $U(\vec{e}_z, 2\pi) = -\mathbb{I}_2$ of SU(2) is simply proportional to the identity matrix:

$$
\mathcal{D}^{(j)}(\vec{e}_z, 2\pi) = (-1)^{2j}\mathbb{I}_{2j+1}. \quad (VI.47b)
$$

In particular, for odd $j$, i.e. for a half-integral spin, $\mathcal{D}^{(j)}(\vec{e}_z, 2\pi)$ is the negative of the identity matrix. This is OK as far as SU(2) is concerned, but it means that the matrices $\{\mathcal{D}^{(j)}(\vec{n}, \psi)\}$ with $j$ a half-integer do not form a linear representation of the group SO(3) — since a rotation through an angle $\psi = 2\pi$ is the identity transformation, which must be mapped to the identity matrix.[38]

As the reader knows, there do exist physical objects with half-integer spin: elementary particles like electrons or quarks; protons and neutrons; atomic nuclei with an odd number of constituents...

**Remarks:**

* Conversely, Eq. (VI.47b) also means that every representation $\mathcal{D}^{(j)}$ of SU(2) with $j \in \mathbb{N}$ will not be faithful, since both $\mathbb{I}_2$ and $-\mathbb{I}_2$ are mapped to the $(2j + 1)$-dimensional identity matrix.

* One can show that every representation $\mathcal{D}^{(j)}$ with $j \in \mathbb{N}$, i.e. every irreducible representation of SO(3), is equivalent to a real representation.

This does not hold for the representations $\mathcal{D}^{(j)}$ of SU(2) with half-integer spin $j$. However, the latter can be shown to be equivalent to their complex conjugate $\mathcal{D}^{(j)*}$, i.e. to be self-conjugate. For instance, one quickly checks the property $\sigma_2 T^\dagger \sigma_2^{-1} = -(T^\dagger)^*$ for every generator of the defining representation $2$, which after multiplication by $-i$ and exponentiation gives

$$
\sigma_2 U_\pi(\psi) \sigma_2^{-1} = U_\pi(\psi)^*,
$$

illustrating the claim for the spin-$\frac{1}{2}$ case.

**VI.3.2c Characters of the spin-$j$ representation**

One can show that any matrix $U_\pi(\psi) = e^{-i\psi \vec{n} \cdot \vec{J}}$ of SU(2) is similar to the diagonal matrix $e^{-i\psi \sigma_3/2} = \text{diag}(e^{-i\psi/2}, e^{i\psi/2})$ with a similarity matrix which is itself an element of SU(2). This gives at once the conjugacy classes of SU(2), which are thus entirely characterized by the parameter $\psi \in [0, 2\pi]$.

In the spin-$j$ representation, the matrix $e^{-i\psi \vec{n} \cdot \vec{J}}$ representing $U_\pi(\psi)$ will then be similar to $e^{-i\psi \vec{J}_z^{(j)}} = \mathcal{D}^{(j)}(\vec{e}_z, \psi)$, i.e. they have the same trace:

$$
\text{Tr}(e^{-i\psi \vec{n} \cdot \vec{J}}) = \text{Tr}[\mathcal{D}^{(j)}(\vec{e}_z, \psi)] = \sum_{m=-j}^{j} e^{-im\psi},
$$

[37] The relation generalizes the conjugacy relation (VI.5) between matrices of SO(3).

[38] The half-integral spin representations are *projective representations* of SO(3), i.e. representations “up to a phase factor”. 
where the second identity uses the matrix elements (VI.44). The sum is easily computed and yields
\[
\chi^{(j)}(\psi) \equiv \text{Tr}(e^{-i\hat{n} \cdot \vec{J}^{(j)}}) = \begin{cases} 
2j + 1 & \text{if } \psi = 0 \\
\sin[(j + \frac{1}{2})\psi] & \text{for } \psi \neq 0,
\end{cases}
\]

where we have introduced the notation $\chi^{(i)}$ for the trace, which is obviously the character of the representation. Note that the formula given for $\psi \neq 0$ tends to $2j + 1$ in the limit $\psi \to 0$, so that the character is a continuous function of the parameter $\psi$ labeling the conjugacy classes of SU(2).

**Remark:** Following a similar reasoning, one shows that every matrix $e^{-i\hat{n} \cdot \vec{J}^{(j)}}$ of the spin-$j$ representation of SU(2) has determinant 1, which means that the generators $\vec{J}^{(j)}$ are traceless.

### VI.3.3 Tensor products of irreducible representations of SU(2)

#### VI.3.3.1 Clebsch–Gordan series

**Theorem VI.49.** Let $\mathcal{D}^{(j_1)}$ and $\mathcal{D}^{(j_2)}$ be two irreducible representations of the Lie group SU(2). Their tensor product is then decomposable, with the Clebsch–Gordan series

\[
\mathcal{D}^{(j_1)} \otimes \mathcal{D}^{(j_2)} = \bigoplus_{J=|j_1-j_2|}^{j_1+j_2} \mathcal{D}^{(J)},
\]

Applying both sides of Eq. (VI.49) to an element $\mathcal{U}_n(\psi)$ of SU(2) and taking the trace, one obtains an identity involving characters:

\[
\chi^{(j_1)}(\psi)\chi^{(j_2)}(\psi) = \sum_{J=|j_1-j_2|}^{j_1+j_2} \chi^{(J)}(\psi) \quad \text{for all } \psi,
\]

irrespective of the choice of the unit vector $\hat{n}$. This result can actually be shown using the expression (VI.48) for $\chi^{(j)}(\psi)$ and the usual addition / multiplication theorems for trigonometric functions, which in turn proves the decomposition (VI.49).

The reader interested in the proof can first show the identity

\[
\chi^{(j)}(\psi)\chi^{(j')}(\psi) = \chi^{(j+j')}(\psi) + \chi^{(j-j')}(\psi)\chi^{(j'+\frac{1}{2})}(\psi),
\]

and then apply it to $(j, j') = (j_1, j_2)$, and iteratively to $(j, j') = (j_1-k, j_2-k)$ for the successive values $k = 1, 2, \ldots, 2\min(j_1, j_2) - 1$, using eventually $\chi^{(0)}(\psi) = 1$.

**Remarks:**

* In the language of physicists, the tensor product of irreducible representations corresponds to the “addition of spins” (or “angular momenta”), and theorem VI.49 is the well-known rule for this addition, according to which the total spin $J$ of a system consisting of a spin $j_1$ and a spin $j_2$ can take the values $J = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, j_1 + j_2$.

* Every irreducible representation $\mathcal{D}^{(J)}$ with $J \in \{ |j_1-j_2|, \ldots, j_1+j_2 \}$ comes up with multiplicity 1 (and the others with multiplicity 0) in the Clebsch–Gordan series (VI.49), i.e. SU(2) is simply reducible (cf. definition III.54).

* Using Eq. (VI.49), one sees that every irreducible representation $\mathcal{D}^{(\frac{1}{2})}$ can be obtained by considering enough tensor products of the defining representation $\mathcal{D}^{(1)}$ with itself, which justifies the denomination “fundamental”.

\[\Box\]
**Example VI.51.** To illustrate the Clebsch–Gordan series \[^{(VI.49)}\], one may write down the tensor product of two spin-1 representations, which are also irreducible representations of SO(3):

\[
\mathcal{R}^{(1)} \otimes \mathcal{R}^{(1)} = \mathcal{R}^{(0)} \oplus \mathcal{R}^{(1)} \oplus \mathcal{R}^{(2)}. \tag{VI.51a}
\]

Denoting the representations by their respective dimension, this also reads

\[
3 \otimes 3 = 1 \oplus 3 \oplus 5, \tag{VI.51b}
\]

which shows that the dimensions on both sides of the equal sign indeed match.

Consider now two 3-dimensional real vectors \(\vec{v}, \vec{w}\) with coordinates \(v^i, w^j\); both transform under spin-1 representation of SO(3). One can construct a tensor (for three-dimensional rotations) of rank 2 consisting of the 9 entries \(v^i w^j\); the latter transforms under SO(3) under the reducible representation \(3 \otimes 3\). More precisely, one can write

\[
v^i w^j = \frac{1}{3} (\vec{v} \cdot \vec{w}) \delta^{ij} + \frac{1}{2} \left( v^i w^j - v^j w^i \right) + \frac{1}{2} \left[ v^i w^j + v^j w^i - \frac{2}{3} (\vec{v} \cdot \vec{w}) \delta^{ij} \right]. \tag{VI.51c}
\]

The first term on the right hand side is invariant under rotations, i.e. transforms under the trivial representation (1). The second, antisymmetric term — in which one recognizes the components of the cross product \(\vec{v} \times \vec{w}\) — transforms under the 3 representation. Eventually, the last term is symmetric and traceless — i.e. the sum over \(i = j\) equals 0 —, and can be shown to transform under the 5 representation of SO(3).

**VI.3.3b Clebsch–Gordan coefficients**

The Clebsch–Gordan series \[^{(VI.49)}\] involves two different “natural” orthonormal bases. The basis adapted to the left hand side consists of all possible tensor products of basis vectors \(|j_1, m_1\rangle\) and \(|j_2, m_2\rangle\), which will be denoted \(\{|j_1, j_2; m_1, m_2\}\rangle\), where only \(m_1\) and \(m_2\) are running while \(j_1\) and \(j_2\) are fixed. In turn, the right hand side rather involves the bases \(\{|J, M\}\rangle\) with \(J\) running from \(j_1 - j_2\) to \(j_1 + j_2\) and for each \(J\) in that set, \(M \in \{-J, \ldots, J\}\).

As in Sec. [III.4.3][39] [see Eq. (III.52)] , one defines the Clebsch–Gordan coefficients \(C_{j_1,j_2;m_1,m_2}^{J,M}\) as those entering the decomposition of the vector \(|J, M\rangle\) on the basis \(\{|j_1, j_2; m_1, m_2\}\rangle\):

\[
|J, M\rangle = \sum_{m_1, m_2} C_{j_1,j_2;m_1,m_2}^{J,M} |j_1, j_2; m_1, m_2\rangle. \tag{VI.52a}
\]

Exploiting fully the bra-ket notation of quantum mechanics, one also writes

\[
C_{j_1,j_2;m_1,m_2}^{J,M} \equiv \langle j_1, j_2; m_1, m_2 | J, M \rangle. \tag{VI.52b}
\]

The latter notation allows one to recognize the meaning of various properties of the coefficients at once, see e.g. Eqs. (VI.54) below.

The Clebsch–Gordan coefficients obey a number of properties, a few of which we now list:[39]

- The Clebsch–Gordan coefficients are real. Accordingly, one may also write

  \[
  C_{j_1,j_2;m_1,m_2}^{J,M} = \langle J, M | j_1, j_2; m_1, m_2 \rangle
  \]

  besides Eq. (VI.52b).

This “property” is actually a convention for the choice for the relative phase of the basis vectors \(\{|j, m\}\rangle\), when one explicitly constructs them by successive tensor product starting from the basis vectors \(\{|\frac{1}{2}, \pm \frac{1}{2}\}\rangle\) of the fundamental representation.

[39]Further results on the coefficients can be found in standard textbooks on Quantum Mechanics, as e.g. Ref. [3], chapter X (complement Bx) or Ref. [4], chapter 4.3.
VI.4 Behavior of quantum-mechanical observables under rotations

- The coefficient \( C_{j_1,j_2;m_1,m_2}^{J,M} \) is non-zero if and only if the numbers \( J, M, j_1, j_2, m_1, m_2 \) obey the "selection rules"

\[
M = m_1 + m_2
\]  

\[
|j_1 - j_2| \leq J \leq j_1 + j_2
\]

\[
J + j_1 + j_2 \in \mathbb{N}
\]

- The Clebsch–Gordan coefficients obey various "orthogonality relations", in particular

\[
\sum_{m_1,m_2} C_{j_1,j_2;m_1,m_2}^{J,M} C_{j_1,j_2;m_1,m_2}^{J',M'} = \delta_{JJ'} \delta_{MM'}
\]

and

\[
\sum_{J,M} C_{j_1,j_2;m_1,m_2}^{J,M} C_{j_1,j_2;m_1',m_2'}^{J,M} = \delta_{m_1 m_1'} \delta_{m_2 m_2'}
\]

VI.4 Behavior of quantum-mechanical observables under rotations

VI.4.1 Rotation of a quantum-mechanical system

Consider a quantum-mechanical system \( \Sigma \), whose physical state is described by a ket-vector \( |\Psi \rangle \) of a Hilbert space \( \mathcal{H} \). Let \( \hat{A} \) denote an observable of the system, i.e. a Hermitian operator on \( \mathcal{H} \). Figure VI.2 represents both \( \Sigma \) and an apparatus \( A \) for the measurement of the physical quantity associated to \( \hat{A} \), for various relative spatial configurations.

(a) In the reference configuration, in which the vector state of the system is \( |\Psi \rangle \), measurements of the expectation value of the observable \( \hat{A} \) yield \( \langle \Psi | \hat{A} | \Psi \rangle \).

(b) With respect to (a), the system \( \Sigma \) has been rotated: let \( \mathcal{R} \) denote the corresponding rotation (in three-dimensional Euclidean space). The latter induces a unitary transformation on the Hilbert space \( \mathcal{H} \) of the system (see discussion in § II.3.3b), such that the state vector becomes \( \mathcal{R} |\Psi \rangle \). Accordingly, the expectation value of \( \hat{A} \) is now \( \langle \Psi | \hat{A}^{\dagger} \mathcal{R} \mathcal{R}^{\dagger} | \Psi \rangle \).

Figure VI.2
(c) The system Σ has the same orientation as in (a), yet now the measurement apparatus has been rotated, by the same rotation \( R \) as that affecting Σ in (b). To account for the fact that the relative orientation of Σ and the apparatus has changed with respect to (a), whereas the state vector \( |Ψ⟩ \) remains unchanged, one has to consider that the rotation induces a transformation of the observable \( \hat{A} \) to a new Hermitian operator \( \hat{R}(\hat{A}) \) on \( \mathcal{H} \). In turn, the expectation value of measurements performed by the apparatus is now \( ⟨Ψ|\hat{R}(\hat{A})|Ψ⟩ \).

(d) Rotating now both the physical system Σ and the measurement apparatus by \( \hat{R} \), the state vector becomes \( \hat{R}|Ψ⟩ \) and the observable \( \hat{R}(\hat{A}) \), i.e. the expectation of measurements is given by \( ⟨Ψ|\hat{R}^\dagger\hat{R}(\hat{A})\hat{R}|Ψ⟩ \).

If space is isotropic, i.e. if physics is invariant under rotations, then the measurements in the configurations (a) and (d) should give the same results, since the relative orientations of the physical system and the measurement apparatus are identical in both cases. That is, one has

\[
⟨Ψ|\hat{R}^\dagger\hat{R}(\hat{A})\hat{R}|Ψ⟩ = ⟨Ψ|\hat{A}|Ψ⟩.
\]

Requiring that this identity hold for any state vector \( |Ψ⟩ \in \mathcal{H} \) yields \( \hat{R}^\dagger\hat{R}(\hat{A})\hat{R} = \hat{A} \), and thus

\[
\hat{R}(\hat{A}) = \hat{A}\hat{R}^\dagger.
\]

One easily checks that the transformation operator \( \hat{R}_2\hat{R}_1 \) acting on observables \( \hat{A} \) which corresponds to successive rotations \( \hat{R}_1 \) and \( \hat{R}_2 \) — i.e. to the rotation \( \hat{R}_2\hat{R}_1 \) — is the product of the operators \( \hat{R}_2 \) and \( \hat{R}_1 \) induced by each rotation:

\[
\hat{R}_2\hat{R}_1 = \hat{R}_2\hat{R}_1,
\]

which shows that the operators \( \hat{R} \) form a group.

**Remark VI.57.** The set of linear operators \( \hat{A} \) on a Hilbert space \( \mathcal{H} \) form a vector space \( \mathcal{L} \), called the \textit{Liouville} space. The operators \( \hat{R} \) introduced in the preceding discussion are operators on \( \mathcal{L} \), and relation (VI.56) means that they form a linear representation of the group of transformations of three-dimensional Euclidean space, or equivalently of the Lie group SO(3), on \( \mathcal{L} \).

Besides the (cumbersome) notation \( \hat{O} \) with a double hat for the operators on \( \mathcal{L} \), one may also introduce a convenient notation for its elements, in which an operator \( \hat{A} \in \mathcal{L} \) on \( \mathcal{H} \) is denoted \( |\hat{A}⟩⟩ \). The equations involving such “superkets” and “superoperators” then take a similar form as those for kets of \( \mathcal{H} \) and the operators acting on them, as e.g. \( \hat{O}|\hat{A}⟩⟩ \) instead of \( \hat{O}(\hat{A}) \).

**Infinitesimal rotations**

Let \( \hat{R}_{δφ} \) denote the rotation through an infinitesimal angle \( δφ \) about the direction \( \vec{n} \), where for the sake of brevity we introduced the shorthand notation \( δφ \equiv (δφ)\vec{n} \). The corresponding induced unitary operator on the Hilbert space \( \mathcal{H} \) reads [cf. Eq. (VI.6)]

\[
\hat{R}_{δφ} = 1_\mathcal{H} - i δφ \cdot \hat{J} + O(δφ^2),
\]

where the three operators \( \hat{J}_i \) with \( i = 1, 2, 3 \) are the Hermitian generators of a representation of the group SO(3) on \( \mathcal{H} \), while \( 1_\mathcal{H} \) is the identity operator on \( \mathcal{H} \).

Similarly, the operator \( \hat{R}_{δφ} \) on the Liouville space \( \mathcal{L} \) induced by the rotation \( \hat{R}_{δφ} \) takes the form

\[
\hat{R}_{δφ} = 1_\mathcal{L} - i δφ \cdot \hat{J} + O(δφ^2),
\]

where the “superoperators” \( \{\hat{J}_i\} \) are the generators of the representation of SO(3) on \( \mathcal{L} \).

\( ^{\text{(ad)}} \) J. Liouville, 1809–1882
Using Eq. (VI.55), the operator \( \hat{\mathcal{R}}_{\delta \phi} \) acts on an operator \( \hat{A} \in \mathcal{L} \) according to

\[
\hat{\mathcal{R}}_{\delta \phi}(\hat{A}) = \hat{\mathcal{R}}_{\delta \phi} \hat{A} \hat{\mathcal{R}}_{\delta \phi}^{-1},
\]

which under consideration of Eq. (VI.58) gives, up to terms of order \( \delta \phi^2 \)

\[
\hat{\mathcal{R}}_{\delta \phi}(\hat{A}) = \left( \mathbb{1}_H - i \delta \phi \cdot \hat{\mathcal{J}} \right) \hat{A} \left( \mathbb{1}_H + i \delta \phi \cdot \hat{\mathcal{J}} \right),
\]

where the hermiticity of the generators \( \hat{\mathcal{J}} \) was used. Thus, one finds

\[
\hat{\mathcal{R}}_{\delta \phi}(\hat{A}) = \hat{A} - i \left[ \delta \phi \cdot \hat{\mathcal{J}}, \hat{A} \right] + \mathcal{O}(\delta \phi^2).
\]

Comparing this result with the generic form (VI.59) shows that the generator \( \hat{\mathcal{J}}_i \) of the representation of SO(3) on the Liouville space acts on an operator \( \hat{A} \) on \( \mathcal{H} \) according to

\[
\hat{\mathcal{J}}_i(\hat{A}) = \left[ \hat{\mathcal{J}}_i, \hat{A} \right] \quad \text{for } i = 1, 2, 3.
\]

**Remark:** Using Eq. (VI.61), one can check that the generators \( \hat{\mathcal{J}}_i \) obey the commutation relation

\[
\left[ \hat{\mathcal{J}}_i, \hat{\mathcal{J}}_k \right] = i \sum_{l=1}^{3} \epsilon_{ikl} \hat{\mathcal{J}}_l
\]

of the group SO(3).

### VI.4.2 Irreducible tensor operators

#### VI.4.2a Definition and characterization

**Definition VI.63.** Let \( j \in \mathbb{N} \). A set of \( 2j + 1 \) operators \( \hat{T}^{(j)}_m \) with \( m \in \{-j, -j + 1, \ldots, j\} \) on a Hilbert space \( \mathcal{H} \) is said to form a spin-\( j \) irreducible tensor operator if they constitute the “standard” basis of a spin-\( j \) irreducible representation of SO(3) on the Liouville space \( \mathcal{L} \) of the linear operators on \( \mathcal{H} \).

This definition can be translated into various characterizations of the operators \( \{ \hat{T}^{(j)}_m \} \). First, denoting \( \mathcal{D}^{(j)}_{m'm}(\mathcal{R}) \) the entries of the \((2j+1) \times (2j+1)\)-matrix representing a given rotation \( \mathcal{R} \) [see Eq. (VI.42)], one has

\[
\hat{\mathcal{R}}(\hat{T}^{(j)}_m) = \sum_{m'=-j}^{j} \mathcal{D}^{(j)}_{m'm}(\mathcal{R}) \hat{T}^{(j)}_{m'}.
\]

Invoking Eq (VI.55), one may also write

\[
\hat{\mathcal{R}} \hat{T}^{(j)}_m \hat{\mathcal{R}}^{-1} = \sum_{m'=-j}^{j} \mathcal{D}^{(j)}_{m'm}(\mathcal{R}) \hat{T}^{(j)}_{m'},
\]

where \( \hat{\mathcal{R}} \) is the unitary operator on \( \mathcal{H} \) representing the rotation \( \mathcal{R} \).

Considering now infinitesimal rotations instead of finite ones, one deduces from the latter equation the equivalent identity

\[
\left[ \hat{\mathcal{J}}_i, \hat{T}^{(j)}_m \right] = \sum_{m'=-j}^{j} \left[ \mathcal{D}^{(j)}_{i'm'}(\mathcal{R}) \hat{T}^{(j)}_{m'} \right],
\]

where the operator \( \hat{\mathcal{J}}_i \) is one of the three Hermitian generators of the representation of SO(3) on \( \mathcal{H} \) while the coefficients \( \left[ \mathcal{D}^{(j)}_{i'm'}(\mathcal{R}) \right] \) are its matrix elements in the basis of eigenvectors \( \{ |j, m \rangle \} \), which are easily read off Eqs. (VI.29b) and (VI.41) under consideration of definitions (VI.31). Thus, one may write

\[
\left[ \hat{\mathcal{J}}_i, \hat{T}^{(j)}_m \right] = m \hat{T}^{(j)}_m
\]
and
\[ \hat{J}_\pm, \hat{T}^{(j)}_{m} = \sqrt{j(j+1) - m(m \pm 1)} \hat{T}^{(j)}_{m \pm 1}. \] (VI.65c)

Equations (VI.65) show that an irreducible tensor operator is fully characterized by its commutation relations with the generators of rotations \( \{ \hat{J}_i \} \), which is what we shall exploit in § VI.4.2b below.

Using Eq. (VI.61), the previous relations can be rewritten with the help of the generators \( \hat{J}_i \) of the representation of SO(3) on the Liouville space \( \mathcal{L} \):
\[ \hat{J}_i (\hat{T}^{(j)}_{m}) = \sum_{m'=-j}^j [\hat{J}_i]^{(j)}_{m'm} \hat{T}^{(j)}_{m'} \] (VI.66a)
\[ \hat{J}_z (\hat{T}^{(j)}_{m}) = m \hat{T}^{(j)}_{m} \] (VI.66b)
\[ \hat{J}_\pm (\hat{T}^{(j)}_{m}) = \sqrt{j(j+1) - m(m \pm 1)} \hat{T}^{(j)}_{m \pm 1}. \] (VI.66c)

**Remark:** The reader is invited to rewrite the equations (VI.66) with the superket notation \( |j,m \rangle \rangle \) (see remark VI.57) in lieu of \( \hat{T}^{(j)}_{m} \), so as to bring them in a form similar to that of relations already encountered in Sec. VI.3.

### VI.4.2b Examples

**Scalar operators**

A scalar operator \( \hat{A} \) is an operator that commutes with all three generators \( \{ \hat{J}_i \} \) of rotations,
\[ [\hat{J}_i, \hat{A}] = 0 \quad \text{for all} \quad i = 1, 2, 3 \] (VI.67)
corresponding to the case \( j = 0 \).

A first example of scalar operator is the quadratic Casimir operator \( \hat{\nabla}^2 \) [see Eqs. (VI.26) and (VI.27)]. Equation (VI.67) also characterizes the Hamilton operator of a system which is invariant under rotations.

**Vector operators**

Considering now the case \( j = 1 \), a vector operator \( \hat{V} \) is a set of three operators \( \hat{V}_x, \hat{V}_y, \hat{V}_z \) that obey the commutation relation
\[ [\hat{J}_a, \hat{V}_b] = i \sum_{c=x,y,z} \epsilon_{abc} \hat{V}_c \quad \text{for all} \quad a, b \in \{ x, y, z \}, \] (VI.68)
where as usual \( \hat{J}_x \equiv \hat{J}_1, \hat{J}_y \equiv \hat{J}_2, \) and \( \hat{J}_z \equiv \hat{J}_3 \). Introducing the standard components
\[ \hat{V}^{(1)}_0 \equiv -\frac{1}{\sqrt{2}}(\hat{V}_x + i\hat{V}_y), \quad \hat{V}^{(1)}_1 \equiv \hat{V}_z, \quad \hat{V}^{(1)}_{-1} \equiv \frac{1}{\sqrt{2}}(\hat{V}_x - i\hat{V}_y), \] (VI.69)
one checks that the latter obey Eqs. (VI.65b)–(VI.65c) with \( j = 1 \) and \( m = 1, 0, -1 \), respectively.

In his lectures on quantum mechanics, the reader should already have encountered a number of examples of vector operators, in particular the position operator \( \hat{\nabla} \), momentum operator \( \hat{p} \), and orbital angular-momentum operator \( \hat{L} \equiv \hat{\nabla} \times \hat{\nabla} \) of a particle, as well as the intrinsic-spin operator \( \hat{S} \) — which, even when its describes the spin degree of freedom of a spin-\( \frac{1}{2} \) particle, consists of three operators forming a vector operator!

### VI.4.2c Tensor product of irreducible tensor operators

see e.g. exercise 44. More later.
VI.4.3 Wigner–Eckart theorem

Consider an irreducible tensor operator $\hat{T}_m^{(j)}$. According to the Wigner–Eckart theorem, its matrix elements in the basis of the common eigenvectors of the generator $\hat{J}_z$ and the Casimir operator $\hat{J}^2$ take the form

$$\langle m'', m' | \hat{T}_m^{(j)} | m', m' \rangle = C_{j,j'; m, m'}^{j''} \langle j'' || \hat{J}^{(j)} || j' \rangle,$$

where $C_{j,j'; m, m'}^{j''}$ is the Clebsch–Gordan coefficient defined by Eqs. (VI.52) while $\langle j'' || \hat{J}^{(j)} || j' \rangle$, known as the reduced matrix element, is now independent of $m$, $m'$, and $m''$.

Viewing the matrix representation of $\hat{T}_m^{(j)}$ in the basis of eigenvectors $\{| j', m' \rangle \}$, it means that the $(2j' + 1) \times (2j'' + 1)$ block associated to fixed values $j'$ and $j''$ is proportional to the matrix of the same size consisting of the Clebsch–Gordan coefficients $C_{j,j'; m, m'}^{j''}$.

**Remark:** Throughout this chapter, we only consider transformations under SO(3) (and SU(2)), i.e. the tensor operators introduced in Sec. VI.4.2 are characterized by their commutation relations with generators of a representation of SO(3), while the Clebsch–Gordan coefficients are those of SO(3). The whole construction, including the Wigner–Eckart theorem, can be generalized to other groups as e.g. the special unitary groups SU($n$).

**Applications**

In physical applications, the Wigner–Eckart theorem in particular leads to selection rules for the transitions induced by an operator involving an irreducible tensor operator.

Consider for instance an atomic system invariant under rotations, whose energy eigenstates have definite values of the orbital and magnetic quantum numbers $\ell$ and $m \in \{-\ell, \ldots, \ell\}$. When the system is in an external (classical) electric field $\vec{E}$, its Hamilton operator is perturbed by a term which in the dipolar approximation takes the form $W = -\hat{P} \cdot \vec{E}$, where

$$\hat{P} = \sum_i q_i \hat{p}_i$$

is the electric dipole operator, which depends on the charges and positions of the particles (nucleus and electrons) constituting the system. This perturbation can induce a transition between an initial state $| i \rangle$ and a final state $| f \rangle$, with a rate proportional (to first order in perturbation theory) to the squared amplitude $| \langle f | W | i \rangle |^2$.

Expressing the scalar product $\hat{P} \cdot \vec{E}$ through the standard components $\hat{P}^{(1)}_m$ with $m = -1, 0, 1$ of the dipole operator, one sees that the computation of the transition rate involves that of the amplitude $\langle f | \hat{P}^{(1)}_m | i \rangle$. Invoking the Wigner–Eckart theorem, the latter is proportional to a Clebsch–Gordan coefficient:

$$\langle f | \hat{P}^{(1)}_m | i \rangle \propto C_{\ell_i m_i; \ell f, m_f}^{\ell f, m_f}$$

where $\ell_i, m_i$ resp. $\ell_f, m_f$ are the quantum numbers of the initial resp. final state. Using the selection rules (VI.53) obeyed by the Clebsch–Gordan coefficients, one deduces the following conditions on these quantum numbers to ensure a non-vanishing matrix element:

$$|\ell_i - \ell_f| \leq 1 \quad , \quad |m_i - m_f| \leq 1 \quad , \quad \ell_i + \ell_f \geq 1.$$

The first two conditions mean that a dipolar transition can at most change the angular momentum by one unit (of $\hbar$), and the third one, that the initial and final states cannot be both s-states (with $\ell = 0$).

(C. Eckart, 1902–1973)
As second example, consider the case of a spin-$j'$ system whose precise “polarization”, i.e. value of $m'$, is unknown — and where every single polarization is equally probable —, so that in the calculation of the expectation value of an observable $\hat{A}$, one needs to average over the different polarization states, i.e.

$$\langle \hat{A} \rangle = \frac{1}{2j'+1} \sum_{m'=-j'}^{j'} \langle j', m' | \hat{A} | j', m' \rangle.$$ 

In the case of a spin-$j$ tensor observable $\hat{T}_m^{(j)}$, the Wigner–Eckart theorem \[VI.70\] gives

$$\langle \hat{T}_m^{(j)} \rangle = \frac{1}{2j'+1} \sum_{m'=-j'}^{j'} \langle j', m' | \hat{T}_m^{(j)} | j', m' \rangle = \frac{1}{2j'+1} \sum_{m'=-j'}^{j'} C^{j';m'}_{j,j';m,m'} \langle j' \| \hat{T}^{(j)} \| j' \rangle,$$

where the reduced matrix element $\langle j' \| \hat{T}^{(j)} \| j' \rangle$ can actually be taken out of the sum. Now, one can show that the sum of the coefficients $C^{j';m'}_{j,j';m,m'}$ over all allowed $m'$ values actually vanishes unless $j = 0$ and $m = 0$ — in which case it trivially equals $2j' + 1$, yielding

$$\langle \hat{T}_m^{(j)} \rangle = \delta_{j0} \delta_{m0} \langle j' \| \hat{T}^{(j)} \| j' \rangle.$$

That is, the expectation value of an irreducible tensor operator in an “unpolarized state” vanishes when the operator is not scalar.
CHAPTER VII

Representations of GL$(n, \mathbb{C})$ and its continuous subgroups

integer $n \geq 2$

VII.1 Low-dimensional irreducible representations of GL$(n, \mathbb{C})$

VII.1.1 Representations of dimension 1

VII.1.1 a Trivial representation

As always, the mapping which maps an arbitrary element $M \in \text{GL}(n, \mathbb{C})$ to the complex number 1 constitutes the trivial representation of the group $\text{GL}(n, \mathbb{C})$, denoted $1$.

VII.1.1 b Determinant representation

Another representation of dimension 1 consists of the mapping which maps an given element $M \in \text{GL}(n, \mathbb{C})$ to its determinant $\det M \in \mathbb{C} \setminus \{0\}$.

In the case of the subgroups of $\text{GL}(n, \mathbb{C})$ consisting of matrices with determinant 1, like $\text{SL}(n, \mathbb{C})$, $\text{SL}(n, \mathbb{R})$, $\text{SU}(n)$ or $\text{SO}(n)$, the determinant representations obviously coincides with the trivial representation.

VII.1.2 Vector representation

Definition VII.1. The defining representation of $\text{GL}(n, \mathbb{C})$, i.e. the automorphism of $\text{GL}(n, \mathbb{C})$ which maps every matrix $M \in \text{GL}(n, \mathbb{C})$ to itself, is called vector representation and denoted $n$.

The latter notation corresponds to the dimension of the representation, which comes from the fact that the matrices of $\text{GL}(n, \mathbb{C})$ operate on complex vectors of dimension $n$, i.e. on the vector space $\mathbb{C}^n$. Since there is no proper subspace of $\mathbb{C}^n$ which is invariant under all elements of $\text{GL}(n, \mathbb{C})$, the vector representation is irreducible.

VII.1.3 Complex conjugate representation

Theorem & Definition VII.2. The mapping which to every matrix $M \in \text{GL}(n, \mathbb{C})$ associates the matrix $M^* \in \text{GL}(n, \mathbb{C})$ whose entries are complex conjugate to those of $M$ is a faithful irreducible representation of $\text{GL}(n, \mathbb{C})$ of dimension $n$, called the complex conjugate representation of the vector representation and denoted $\bar{n}$.

In the case of the subgroups of $\text{GL}(n, \mathbb{C})$ consisting of real matrices, i.e. of $\text{GL}(n, \mathbb{R})$ and its subgroups, the complex conjugate representation $\bar{n}$ coincides with the vector representation $n$. 
VII.1.4 Dual representation

Theorem & Definition VII.3. The mapping which to every matrix $M \in \text{GL}(n, \mathbb{C})$ associates the matrix $(M^{-1})^T \in \text{GL}(n, \mathbb{C})$ is a faithful irreducible representation of $\text{GL}(n, \mathbb{C})$ of dimension $n$, called the dual or contragredient representation of the vector representation and denoted $n^*$.

Considering that the matrix $M$ represents a regular linear map $\mathcal{V} \to \mathcal{W}$ in given bases, where $\mathcal{V}$ and $\mathcal{W}$ are $n$-dimensional complex vector spaces, then $M^{-1}$ represents a map $\mathcal{W} \to \mathcal{V}$, and $(M^{-1})^T$ a map $\mathcal{W}^* \to \mathcal{V}^*$, where $\mathcal{V}^*$ and $\mathcal{W}^*$ are the dual spaces to $\mathcal{V}$ and $\mathcal{W}$, respectively.

Property VII.4. In the case of the unitary group $U(n)$ and its subgroups $SU(n)$, $O(n)$, $SO(n)$..., the dual representation coincides with the complex conjugate representation.

A unitary matrix $\mathcal{U} \in U(n)$ obeys $\mathcal{U} \mathcal{U}^\dagger = \mathcal{U}^\dagger \mathcal{U} = 1_n$, i.e. $(\mathcal{U}^{-1})^T = (\mathcal{U}^\dagger)^T = \mathcal{U}^*$. □

Eventually, yet another irreducible representation of dimension $n$ of the group $\text{GL}(n, \mathbb{C})$ is the dual of the complex conjugate, denoted $\bar{n}^*$, which maps $M \in \text{GL}(n, \mathbb{C})$ to the Hermitian conjugate of its inverse, $(M^{-1})^\dagger$.

VII.2 Tensor representations of $\text{GL}(n, \mathbb{C})$

In this section...

VII.2.1 Tensor products of the vector representation

In Sec. VII.1.2 we introduced the vector representation $n$ of $\text{GL}(n, \mathbb{C})$: the corresponding $n \times n$-matrices operate on the vector space $\mathbb{C}^n$ according to

$$x \in \mathbb{C}^n \to x' = Mx \in \mathbb{C}^n,$$

i.e. component-wise

$$x^i \to x'^i = \sum_{j=1}^n M_{ij} x^j \quad \forall i \in \{1, \ldots, n\},$$

where the $\{x^i\}$ resp. $\{x'^i\}$ are the components of $x$ resp. $x'$ and $M_{ij}$ the entries of the matrix $M$, where as usual the left resp. right index stands for the line resp. column of the matrix.

Starting from the vector space $\mathbb{C}^n$, one constructs for every $p \in \mathbb{N}^*$ the tensor-product space $(\mathbb{C}^n)^\otimes p$ whose elements are tensors of order $p$, hereafter generically denoted $\mathbf{T}$, with components (in a given basis) traditionally characterized by $p$ indices:

$$\mathbf{T}^{i_1 \ldots i_p} \quad \text{with } i_1, i_2, \ldots, i_p \in \{1, \ldots, n\}.$$

Generalizing Eq. (VII.5), one defines the operation of $M \in \text{GL}(n, \mathbb{C})$ on such tensors:

Theorem & Definition VII.6. The mapping from $\text{GL}(n, \mathbb{C})$ into the group of regular linear applications of the tensor-product space $(\mathbb{C}^n)^\otimes p$ into itself which associates to a matrix $M$ with entries $M_{ij}$ the linear operator $M^\otimes p$ such that

$$M^\otimes p : \mathbf{T} \to \mathbf{T}' = (M \otimes \cdots \otimes M)\mathbf{T},$$

corresponding in terms of components to
VII.2 Tensor representations of GL\((n, \mathbb{C})\)

\[
T^{i_1i_2\ldots i_p} \rightarrow T'^{i_1i_2\ldots i_p} = \sum_{j_1=1}^{n} \cdots \sum_{j_p=1}^{n} M^{i_1j_1} M^{i_2j_2} \cdots M^{i_pj_p} T^{j_1j_2\ldots j_p},
\]

(VII.6b)

is a representation of \(GL(n, \mathbb{C})\) of dimension \(n^p\), called the tensor-product representation

\[
\underbrace{n \otimes \cdots \otimes n}_p \text{ copies}
\]

(VII.6c)

This representation is generally reducible when \(p > 1\).

**VII.2.2 Operations on the indices of a tensor**

**VII.2.2.a Operation of a permutation on a tensor**

**Definition VII.7.** Let \(\sigma \in S_p\) be a permutation of \(p\) elements. One associates with it an operator \(\hat{\sigma}\) acting on the tensor-product space \((\mathbb{C}^n)^{\otimes p}\) such that for every tensor \(T\) of order \(p\), the tensor \(\hat{\sigma}(T)\) has the components

\[
(\sigma T)^{i_1i_2\ldots i_p} = T^{i_{\sigma(1)}i_{\sigma(2)}\ldots i_{\sigma(p)}}.
\]

(VII.7)

That is, the permutation acts on the tensor indices.

**Remark:** One can actually show that this definition is independent of the choice of basis in \((\mathbb{C}^n)^{\otimes p}\).

**Example VII.8.** Letting \(p = 2\) and considering the transposition \(\tau_{12} = (1\ 2) \in S_2\), the associated operator \(\hat{\tau}_{12}\) is such that for every second-order tensor \(T\), the tensor \(\hat{\tau}_{12}(T)\) has the components

\[
(\tau_{12}T)^{ij} = T^{ji} \quad \text{for all } i, j \in \{1, \ldots, n\}.
\]

(VII.8)

**Property VII.9.** The operator \(\hat{\sigma}\) associated to a permutation \(\sigma \in S_p\) is linear.

**Theorem VII.10.** For every permutation \(\sigma \in S_p\) and every matrix \(M \in GL(n, \mathbb{C})\), the linear operators \(\hat{\sigma}\) and \(M^{\otimes p}\) on the tensor-product space \((\mathbb{C}^n)^{\otimes p}\) commute. That is, denoting the image by \(M^{\otimes p}\) with a \(\prime\) as in Eq. (VII.6a), one has

\[
\hat{\sigma}(T') = (\hat{\sigma}(T))'
\]

(VII.10)

for every tensor \(T\) of order \(p\).

**VII.2.2.b Young operators**

Any linear combination of linear operators on the tensor-product space \((\mathbb{C}^n)^{\otimes p}\) is again a linear operator. Accordingly, one can associate such operators to the elements of the group algebra \(\mathbb{C}S_p\) in the same manner as in definition VII.7. In particular, to each Young element \(\mathcal{Y}\) introduced in §III.5.1b corresponds a linear operator \(\mathcal{Y}\) on \((\mathbb{C}^n)^{\otimes p}\), generically called Young operators.

For instance, one associates to the symmetrizer (III.55a) and to the antisymmetrizer (III.55b) for \(p\) elements the operators

\[
\hat{\mathcal{S}} \equiv \frac{1}{p!} \sum_{\sigma \in S_p} \hat{\sigma}
\]

(VII.11a)

\[
\hat{\mathcal{A}} \equiv \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon(\sigma) \hat{\sigma},
\]

(VII.11b)

where we recall that \(\varepsilon(\sigma)\) denotes the signature of the permutation \(\sigma\). Invoking Eq. (VII.7), their respective actions on a tensor \(T \in (\mathbb{C}^n)^{\otimes p}\) of order \(p\) read...
\[
(\mathcal{T})^{i_1 i_2 \cdots i_p} = \frac{1}{p!} \sum_{\sigma \in S_p} T^{i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(p)}}, \quad (\mathcal{\bar{T}})^{i_1 i_2 \cdots i_p} = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon(\sigma) T^{i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(p)}}
\]

(VII.12a)

(VII.12b)

for every \(p\)-tuple \((i_1, i_2, \ldots, i_p)\). The tensor \(\mathcal{T}(\mathbf{T})\) resp. \(\mathcal{\bar{T}}(\mathbf{T})\) is then symmetric resp. antisymmetric under the exchange of any two indices.

In the case of order-2 tensors, these expressions become

\[
(\mathcal{T})^{ij} = \frac{1}{2} (T^{ij} + T^{ji}) \quad \text{and} \quad (\mathcal{\bar{T}})^{ij} = \frac{1}{2} (T^{ij} - T^{ji})
\]

(VII.13)

for every \(i, j \in \{1, \ldots, n\}\).

**Remark:** While the operators \(\sigma\) associated to permutations are regular, this is no longer the case of those associated to Young elements. For instance, the operation of the antisymmetrizer \((\text{VII.11b})\) on any fully symmetric tensor \[^{(40)}\] — i.e. on an element of the image of \((\mathbb{C}^n)^\otimes p\) by \(\mathcal{\bar{T}}\) — is the zero tensor.

**VII.2.3 Irreducible representations built from the vector representation**

**VII.2.3 a Irreducible representations and Young operators**

The theorem \((\text{VII.10})\) extends to the Young operators \(\mathcal{\hat{Y}}\), which all commute with every operator on \((\mathbb{C}^n)^\otimes p\) of the form \(M^\otimes p\) with \(M \in \text{GL}(n, \mathbb{C})\):

\[
\forall \mathbf{T} \in (\mathbb{C}^n)^\otimes p, \quad \mathcal{\hat{Y}}(M^\otimes p(\mathbf{T})) = M^\otimes p(\mathcal{\hat{Y}}(\mathbf{T})).
\]

One deduces thereof at once that the image of every Young operator \(\mathcal{\hat{Y}}\) is a subspace of the tensor-product space \((\mathbb{C}^n)^\otimes p\) which is invariant under all operators \(M^\otimes p\) of the tensor-product representation \(n \otimes \cdots \otimes n\).

Consider \(\mathbf{T} \in \text{im} \mathcal{\hat{Y}}\): there exists \(\mathbf{S} \in (\mathbb{C}^n)^\otimes p\) such that \(\mathbf{T} = \mathcal{\hat{Y}}(\mathbf{S})\). Then for any \(M \in \text{GL}(n, \mathbb{C})\) one may write

\[
M^\otimes p(\mathbf{T}) = M^\otimes p(\mathcal{\hat{Y}}(\mathbf{S})) = \mathcal{\hat{Y}}(M^\otimes p(\mathbf{S})),
\]

which shows that \(M^\otimes p(\mathbf{T}) \in \text{im} \mathcal{\hat{Y}}\). \(\square\)

Accordingly, the tensors of order \(p\) belonging to the image of any Young operator \(\mathcal{\hat{Y}}\) span a subspace of \((\mathbb{C}^n)^\otimes p\) on which the restrictions of the linear operators \(M^\otimes p\) constitute a linear representation of \(\text{GL}(n, \mathbb{C})\). One can show that this representation is irreducible. More precisely, one has the following result \[^{(41)}\]

**Theorem VII.14.** Given an integer \(p\) and a Young operator \(\mathcal{\hat{Y}}\) acting on order-\(p\) tensors, the image by \(\mathcal{\hat{Y}}\) of the tensor-product space \((\mathbb{C}^n)^\otimes p\) is the representation space of an irreducible linear representation \(\mathcal{\hat{Y}}(\mathcal{g})\) of the group \(\text{GL}(n, \mathbb{C})\), which is contained in the Clebsch–Gordan decomposition of the \(p\)-fold tensor-product representation \(n \otimes \cdots \otimes n\).

More generally, the representation \(\mathcal{\hat{Y}}(\mathcal{g})\) is also an irreducible representation of the subgroups \(\text{SL}(n, \mathbb{C}), \text{GL}(n, \mathbb{R}), \text{SL}(n, \mathbb{R}), \text{U}(n), \text{and SU}(n)\). On the other hand, it is in general reducible with respect to \(\text{GL}(n - 1, \mathbb{C})\) and its subgroups, and with respect to \(\text{O}(n), \text{SO}(n)\) and the indefinite orthogonal groups \(\text{O}(n - m, m), \text{SO}(n - m, m)\) where \(m \in \{1, \ldots, n - 1\}\).

\[^{(40)}\]or more generally, on any tensor which remains unchanged under the exchange of two of its indices.

\[^{(41)}\]A reference would be welcome!
VII.2 Tensor representations of GL($n$, $\mathbb{C}$)

VII.2.3 Young tableaux

Building on the correspondence introduced in Sec. [III.5.1] between Young elements of CS$_p$ and (standard) Young tableaux with $p$ boxes, we may now associate a specific Young tableau with the representation $\hat{\mathcal{D}}(\mathcal{Y})$ introduced in the previous paragraph.

More precisely, each irreducible representation contained in the Clebsch–Gordan decomposition of the $p$-fold tensor-product representation $\mathbf{n} \otimes \cdots \otimes \mathbf{n}$ built with the vector representation $\mathbf{n}$ of GL($n$, $\mathbb{C}$) is associated with a Young diagram with $p$ boxes, such that the number of rows of the diagram equals at most $n$.

Examples:

* The vector representation $\mathbf{n}$ is associated with the Young diagram with a single box $\square$.

* For tensors of order 2, there are two possible Young diagrams. As was mentioned in § III.5.1b, the diagram with a simple row corresponds to the symmetrizer, while that with a single column is associated to the antisymmetrizer:

  \[ \hat{\mathcal{S}} \equiv \frac{1}{2} (\text{Id} + (1 \ 2)) \qquad \square \square ; \]

  \[ \hat{\mathcal{A}} \equiv \frac{1}{2} (\text{Id} - (1 \ 2)) \qquad \square \square ; \] (VII.15a)

  \[ \qquad \square \square ; \]

  \[ \text{(VII.15b)} \]

* With 3 boxes, i.e. associated with tensors of order 3, we have found 4 standard Young tableaux (Fig. III.1). The one with a single line is associated to the symmetrizer, that with a single column with the antisymmetrizer, so we want to focus on those corresponding to the Young diagram $\square \square$. Replacing the numbers $\{1, 2, 3\}$ by the indices $\{i, j, k\}$, we have the consider the two tableaux

\[
\begin{array}{ccc}
  \square & \square & \square \\
  i & j & k \\
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
  \square & \square & \square \\
  i & k & j \\
\end{array}
\]

Let us in particular deal with the left one. Adapting what was found in § III.5.1b the corresponding Young element (of CS$_3$) is $\mathcal{Y} = \text{Id} + (i \ j) - (i \ k) - (i \ j \ k)$, up to an irrelevant factor. The associated Young operator $\mathcal{Y}$ is then such that its action on a tensor $T$ gives a tensor with components

\[ (\mathcal{Y}T)^{ijk} = T^{ijk} + T^{jik} - T^{kji} - T^{jki}, \]

i.e. symmetric under the exchange of $i$ and $j$, which are in the first row of the Young tableau, and antisymmetric under the exchange of $i$ and $k$, which both sit in the same column of the Young tableau.

The procedure outlined in the last example can be generalized. Starting from a given standard Young tableau with $p$ boxes — associated with a Young element $\mathcal{Y}$ following the recipe given in § III.5.1b —, one replaces one-to-one the numbers $1, \ldots, p$ in the boxes by indices $i_1, \ldots, i_p$. The expression of the component $(\mathcal{Y}T)^{i_1 \cdots i_p}$ of the tensor $\mathcal{Y} (T)$ can then be deduced from the expression of $\mathcal{Y}$ in terms of permutations of S$_p$, by letting the permutations act on the indices.

The antisymmetry with respect to the indices in a given column of the Young tableau explains why the latter can have at most $n$ rows. If it had more, there would always be at least two of the corresponding indices taking the same value, so that the antisymmetrization automatically would give 0: the associated tensor is the zero tensor of order $p$.

Remark: The above recipe gives the form of the tensors that constitute the representation space of a given irrep. contained in the $p$-fold tensor-product of the vector representation, but it does not specify the actual representation of a matrix $M$ of GL($n$, $\mathbb{C}$).
Dimension of the irrep. of GL\((n, \mathbb{C})\) associated with a Young tableau

Given a Young diagram with \(p\) boxes, the necessary symmetrization resp. antisymmetrization over indices in a same row resp. column leads to restrictions on the number of independent \(p\)-tuples of indices which are relevant — corresponding to the dimension of the associated representation space. To compute this dimension \(d\), one can combine the various combinatorial constraints into a simple recipe, which now give and illustrate on the example of the Young diagram

\[
\begin{array}{cccc}
& & & \\
& & & \\
& & & \\
\end{array}
\]

with 7 boxes.

- As first step, one fills the boxes of the Young diagram, with \(n\) in the upper left box and then increasing resp. decreasing by 1 for each step further right in a row resp. down in a column:

\[
\begin{array}{cccc}
n & n+1 & n+2 & n+3 \\
n-1 & n & & \\
n-2 & & & \\
\end{array}
\]

The \(p\) numbers thus entered in the Young diagram are multiplied with each other (yielding here \(n^2(n^2-1)(n^2-4)(n+3)\)).

- In a second step, one enters in another copy of the same Young diagram the respective hook lengths of the boxes, as introduced in §III.5.1a. As already illustrated on the same example in Fig. [III.3] this gives

\[
\begin{array}{cccc}
6 & 4 & 2 & 1 \\
3 & 1 & & \\
1 & & & \\
\end{array}
\]

These hook lengths are then multiplied with each other (yielding here \(6 \cdot 4 \cdot 2 \cdot 2\)).

- Eventually, the wanted dimension \(d\) is the ratio of the number found in the first step divided by that found in the second step.

Remark: If a column contains more than \(n\) boxes, then the entry in the \((n+1)\)-th box in the first step would be zero, resulting in a vanishing dimension for the associated representation: the recipe automatically “knows” about the maximal possible number of boxes in a column.

Using this recipe, one quickly finds the dimensions of the representations associated with the diagram with a single row resp. column with \(p\) boxes:

\[
\begin{align*}
\text{\(p\) boxes} & : \text{dimension } \frac{n(n+1) \cdots (n+p-1)}{p(p-1) \cdots 1} = \frac{(n+p-1)!}{(n-1)!p!} = \binom{n+p-1}{p}; \\
\text{\(p\) boxes} & : \text{dimension } \frac{n(n-1) \cdots (n-p+1)}{p(p-1) \cdots 1} = \frac{n!}{(n-p)!p!} = \binom{n}{p}.
\end{align*}
\]

Note two interesting special cases of the latter formula (which is only meaningful for \(p \leq n\)). First, if there are \(p = n\) boxes, the associated representation has dimension 1, and one easily sees
that the corresponding representation of $\text{GL}(n, \mathbb{C})$ is the determinant representation introduced in §VII.1.11.  

In the case of a single column with $p = n - 1$ boxes, Eq. (VII.17) shows that the corresponding irreducible representation has dimension $n$. If $n = 2$, this is clear with $n - 1$ boxes one obtains the vector representation $n$. For $n \geq 3$, the result is more interesting, and will be discussed at further length in Sec. VII.2.5 hereafter.

### VII.2.3 Tensor products of irreps. built from the vector representation

"outer product" introduced in Sec. III.5.2

### VII.2.4 Tensor products of the dual and conjugate representations

The construction of Secs. VII.2.1—VII.2.3 can readily be repeated to build representations of $\text{GL}(n, \mathbb{C})$ from tensor products of the three further $n$-dimensional irreducible representations (complex conjugate, dual, and dual to the complex conjugate) introduced in Secs. VII.1.3—VII.1.4.

#### VII.2.4 a Tensor products of the dual representation

As was mentioned in Sec. VII.1.4 the vector space on which the dual representation $n^*$ acts is actually the dual space of $\mathbb{C}^n$, i.e. the space of the linear forms $\mathbb{C}^n \to \mathbb{C}$. In matrix representation, such a linear form is represented by a row vector, which we shall denote $x^T$ (where $x$ stands for a column vector). In addition, the components of a linear form in a given basis will be denoted with subscript-indices.

If $M$ is an arbitrary element of $\text{GL}(n, \mathbb{C})$, the operation of the matrix $(M^{-1})^T$ of the dual representation on an arbitrary linear form $x^T$ reads

$$x^T \to x'^T = x^T (M^{-1})^T,$$

which indeed gives a linear form on $\mathbb{C}^n$. In terms of components, this becomes

$$x_i \to x'_i = \sum_{j=1}^{n} (M^{-1})^j_i x_j = \sum_{j=1}^{n} x_j ((M^{-1})^T)^j_i \quad \forall i \in \{1, \ldots, n\},$$

where $\{x_i\}$ resp. $\{x'_i\}$ denote the components of $x^T$ resp. $x'^T$ and $(M^{-1})^j_i = ((M^{-1})^T)^j_i$ the entry in the $i$-th row and $j$-th column of the matrix $M^{-1}$, i.e. in the $j$-th row and $i$-th column of the matrix $(M^{-1})^T$.

In complete similarity with what was done with the vector representation, one can now define for any $p \geq 2$ the $p$-fold tensor product representation

$$n^* \otimes \cdots \otimes n^*,$$

such that the action of an element $M \in \text{GL}(n, \mathbb{C})$ on a tensor of order $p$ built from linear forms, i.e. with components characterized by $p$ subscript indices

$$T_{i_1i_2\ldots i_p} \quad \text{with} \quad i_1, i_2, \ldots, i_p \in \{1, \ldots, n\},$$

reads

$$T_{i_1i_2\ldots i_p} \to T'_{i_1i_2\ldots i_p} = \sum_{j_1=1}^{n} \cdots \sum_{j_p=1}^{n} (M^{-1})^j_{i_1} (M^{-1})^j_{i_2} \cdots (M^{-1})^j_{i_p} j_{p} T_{j_1j_2\ldots j_p}$$

for every $p$-tuple $i_1, i_2, \ldots, i_p$.

#### Remarks:

* Subscript indices, associated with linear forms and their tensor products, are generally referred to as contravariant. In turn, superscript indices, associated with vectors, are called covariant.
A quantity which is traditionally denoted with \( n \) contravariant indices is the fully antisymmetric \( n \)-dimensional Levi-Civita symbol \( \epsilon_{i_1 i_2 \ldots i_n} \), defined by

\[
\epsilon_{i_1 i_2 \ldots i_n} = \begin{cases} 
+1 & \text{if } i_1, i_2, \ldots, i_n \text{ is an even permutation of } 1, 2, \ldots, n; \\
-1 & \text{if } i_1, i_2, \ldots, i_n \text{ is an odd permutation of } 1, 2, \ldots, n; \\
0 & \text{if at least two indices take the same value.}
\end{cases}
\] (VII.21)

Letting a matrix \( M \in \text{GL}(n, \mathbb{C}) \) act on it according to Eq. (VII.20) yields

\[
\epsilon_{i_1 i_2 \ldots i_n} \rightarrow \epsilon'_{i_1 i_2 \ldots i_n} = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} (M^{-1})_{i_1}^{j_1} (M^{-1})_{i_2}^{j_2} \cdots (M^{-1})_{i_n}^{j_n} \epsilon_{j_1 j_2 \ldots j_n} = \det(M^{-1}) \epsilon_{i_1 i_2 \ldots i_n}.
\] (VII.22)

If the determinant of \( M \) (and thus \( M^{-1} \)) differs from 1, the symbol is not invariant under the action of \( M \), and in particular loses its defining property \( \epsilon_{1 2 \ldots n} = 1 \). In that respect, the Levi-Civita symbol is not considered as a “tensor”.

On the other hand, \( \epsilon_{i_1 i_2 \ldots i_n} \) remains invariant under the action of all elements of \( \text{SL}(n, \mathbb{C}) \), and thus is a “valid” contravariant tensor for this group and its subgroups — in particular \( \text{SU}(n) \) or \( \text{SO}(n) \).

As in Sec. VII.2.3, one can let Young operators act on the contravariant tensors of order \( p \), which leads to irreducible representations of \( \text{GL}(n, \mathbb{C}) \) that are contained in the Clebsch–Gordan series of the \( p \)-fold tensor-product representation \( n^* \otimes \cdots \otimes n^* \).

In turn, these irreducible representations can be associated with Young tableaux, which specify the symmetry properties of the indices of the respective tensors. These tableaux are built from a fundamental building block \( \square \) which is no longer the vector representation \( n \), but now the dual representation \( n^* \).

Note that the recipe for the dimension of the irrep(s), built from the vector representation and associated with a given Young diagram still holds for the irreducible representations built from the dual representation.

**VII.2.4 b More tensor product representations**

Instead of using the vector or dual representation as fundamental building block, one may also use the complex conjugate \( \bar{n} \) or its dual \( \bar{n}^* \), and construct tensor-product representations

\[
\bar{n} \otimes \cdots \otimes \bar{n}, \quad \bar{n}^* \otimes \cdots \otimes \bar{n}^*
\] (VII.23)

that operate on respective tensors whose components involve a “new type” of indices for either possibility since they now transform with \( M^* \) or \( (M^{-1})^T \) instead of \( M \) or \( (M^{-1})^T \).

One may also construct “mixed” tensors, which constitute the objects that transform under representations built from more than one building block \( n, n^*, \bar{n}, \bar{n}^* \). For instance, the tensor-product representation \( n \otimes n^* \) acts on tensors with one covariant and one contravariant index:

\[
T^j \rightarrow T'^j = \sum_{k=1}^{n} \sum_{l=1}^{n} M^j_k (M^{-1})^l_j T^l_k \quad \text{for all } M \in \text{GL}(n, \mathbb{C}), \ i, j \in \{1, \ldots, n\},
\]

which combines Eqs. (VII.5b) and (VII.18b) in a natural way.

All in all, one can show that the finite-dimensional irreducible representations of \( \text{GL}(n, \mathbb{C}) \) indeed act on tensors whose components are labeled by 4 types of indices. However, these irreps. do not coincide with the set of all possible tensor products of \( n, n^*, \bar{n}, \bar{n}^* \) and of the irreducible representations built from each of these building blocks with the help of Young operators. For instance, one can already find that the representation \( n \otimes n^* \) is reducible in \( \text{GL}(n, \mathbb{C}) \) — it contains the trivial representation.

\footnote{NB: need a reference!}

\footnote{A common notation uses “dotted” indices, cf. Sec. ?? on the irreducible representations of \( \text{SL}(2, \mathbb{C}) \).}
VII.2.5 Representations of SL\((n, \mathbb{C})\) and SL\((n, \mathbb{R})\)

As mentioned in the second part of theorem VII.14, the finite-dimensional irreducible representations of GL\((n, \mathbb{C})\) are also irreducible with respect to its subgroups SL\((n, \mathbb{C})\), GL\((n, \mathbb{R})\), SL\((n, \mathbb{R})\), U\((n)\) and SU\((n)\). However, representations that are inequivalent with respect to GL\((n, \mathbb{C})\) may now become equivalent, which leads to a number of simplifications, as we now sketch in this section and the following.

We first quickly discuss the representations of the groups SL\((n, \mathbb{C})\), GL\((n, \mathbb{R})\), and SL\((n, \mathbb{R})\). Since these groups are not compact, their respective finite-dimensional irreducible representations are not unitary, and not equivalent to a unitary representation.

VII.2.5a Representations of SL\((n, \mathbb{C})\)

In the case of the group SL\((n, \mathbb{C})\), the dual representation \(n^*\) is equivalent to the irreducible representation acting on totally antisymmetric tensors of order \(n - 1\) with covariant indices, i.e. built on the vector representation \(n\) only:

\[
\begin{align*}
\begin{array}{c}
\n^* \\
\end{array} & \equiv \begin{array}{c}
\vdots \\
\end{array} \\
\end{align*}
\]

\(n - 1\) boxes

(Equation VII.24)

More generally, one shows the equivalence between the irreducible representations acting on totally antisymmetric covariant tensors of order \(p\) and on totally antisymmetric contravariant tensors of order \(n - p\), respectively.

The key to these equivalences is that the \(n\)-dimensional Levi-Civita symbol \(\epsilon_{i_1i_2...i_n}\) is invariant under SL\((n, \mathbb{C})\), whose matrices all have determinant 1. Therefore, one can introduce a correspondence

\[
T^{i_1...i_p} \rightarrow \bar{T}_{i_1...i_{n-p}} \equiv \sum_{i_{n-p+1}=1}^{n} \cdots \sum_{i_{n}=1}^{n} \epsilon_{i_1i_2...i_n} T^{i_{n-p+1}i_{n-p+2}...i_n}
\]

between both type of tensors which is preserved under the action of any matrix of SL\((n, \mathbb{C})\).

Obviously, similar equivalences hold between pairs of representations acting on totally antisymmetric tensors built from the complex conjugate representation \(\bar{n}\) and acting on its dual \(\bar{n}^*\).

Accordingly, the finite-dimensional irreducible representations of SL\((n, \mathbb{C})\) act on tensors whose components are labeled by two sets of indices only, instead of 4. Moreover, each of these sets is associated with a Young tableau — which characterizes the symmetry properties of the respective indices — with at most \(n - 1\) rows. A column with \(n\) boxes is indeed equivalent to the trivial representation, and thus irrelevant.

We shall illustrate these ideas on the example of SL\((2, \mathbb{C})\) in a later chapter.

VII.2.5b Representations of GL\((n, \mathbb{R})\) and SL\((n, \mathbb{R})\)

As already mentioned in Sec. VII.3 the vector and complex-conjugate representations \(n\) and \(\bar{n}\) of GL\((n, \mathbb{R})\) coincide. Similarly, the dual representation \(n^*\) and its complex conjugate \(\bar{n}^*\) are identical.

Accordingly, the tensors constituting the representation spaces of the finite-dimensional irreducible representations of GL\((n, \mathbb{R})\) are characterized by two sets of indices, covariant and contravariant, i.e. which transform respectively like a vector or like a linear form.

Since SL\((n, \mathbb{R})\) is a subgroup of GL\((n, \mathbb{R})\), the equalities of the \(n\) and \(\bar{n}\) and of the \(n^*\) and \(\bar{n}^*\) representations further hold. In addition, a further simplification comes from the fact that SL\((n, \mathbb{R})\) is also a subgroup of SL\((n, \mathbb{C})\), so that the results given in § VII.2.5a apply. In particular, the dual representation \(n^*\) is related to the vector representation \(n\), see Eq. (VII.24).
All in all, one thus finds that the finite-dimensional irreducible representations of $SL(n, \mathbb{R})$ act on tensors whose components involve indices of a single type, i.e. which can be built from the vector representation $n$ only.

### VII.2.6 Representations of SU($n$)

Let us now consider the representations of the special unitary group SU($n$). A first important property is that since SU($n$) is compact, its finite-dimensional irreducible representations are always equivalent to a unitary representation — as is obvious in the case of the defining representation $n$.

#### VII.2.6 a Relations between the vector representation and its dual and complex-conjugate

Since SU($n$) is a subgroup of SL($n, \mathbb{C}$), the results of § VII.2.5a hold, in particular the equivalence (VII.24) relating the dual representation $\bar{n}^*$ to the vector representation $n$. In addition, the property VII.4 valid for unitary groups gives a further relation between the dual and the complex-conjugate representation $\bar{n}$, which are identical. Altogether, this gives

\[
\bar{n} = n^* \cong \begin{array}{c}
\vdots \\
\end{array}, \\
\begin{array}{c}
\vdots \\
\end{array} \quad n - 1 \text{ boxes}, \\
\begin{array}{c}
\vdots \\
\end{array},
\]  

(VII.25)

where $\begin{array}{c}
\vdots \\
\end{array}$ stands for the vector representation $n$, and all four $n$-dimensional irreducible representations listed in Sec. VII.1.2–VII.1.4 are related.

In the case $n = 2$, relation (VII.25) means that the complex-conjugate representation $\bar{2}$ of SU(2) is equivalent to the defining representation 2. This is indeed the case, since one can check starting from Eq. (VI.17) that a generic matrix $U \in SU(2)$ obeys $U^* = \sigma_2 U \sigma_2^{-1}$, as already noted in the last remark of § VI.3.2b.

Starting from $n = 3$, the defining representation and its complex-conjugate (or dual) are inequivalent.

#### VII.2.6 b Adjoint representation of SU($n$)

Besides the vector representation $n$ and its complex conjugate $\bar{n}$, another irrep. of SU($n$), the adjoint representation, is often encountered in physical applications, which we now introduce.

Consider the tensor product $n \otimes \bar{n}$ of the vector representation and its complex conjugate. Representing them by Young diagrams and using relation (VII.25), one finds the Clebsch–Gordan series

\[
\begin{array}{c}
\vdots \\
\end{array} \otimes \begin{array}{c}
\vdots \\
\end{array} = \begin{array}{c}
\vdots \\
\end{array} \oplus \begin{array}{c}
\vdots \\
\end{array}, \\
\begin{array}{c}
\vdots \\
\end{array} \quad n - 1 \text{ boxes}.
\]  

(VII.26a)

where the first Young diagram on the right-hand side stands in fact for the trivial representation — for which there is no simpler diagram. Denoting the representations by their respective dimensions, this reads

\[
n \otimes \bar{n} = 1 \oplus n^2 - 1,
\]  

(VII.26b)

where the reader is invited to check the dimension $(n^2 - 1)$ of the rightmost irreducible representation in Eq. (VII.26a) with the help of the recipe given in § VII.2.3b. That is, the tensor product

\[\text{(44) See Sec. IX.3 for examples.}\]
representation \( n \otimes \bar{n} \) decomposes into the direct sum of two irreducible representations, one of which is the trivial representation. The second one, with dimension \( n^2 - 1 \), is called the \textit{adjoint representation}.

One can show that the adjoint representation is self-conjugate, i.e. that it is equivalent to the \((n^2 - 1)\)-dimensional representation in which every matrix is replaced by its complex conjugate.

Here are two possible lines of arguments to prove that result:

First, looking at the Young diagram, the column with \( n - 1 \) boxes, corresponding to the antisymmetrization of \( n - 1 \) covariant indices, is equivalent to a single contravariant index, while the column with only one box, which stands for a single contravariant index, is equivalent to a set of \( n - 1 \) antisymmetrized contravariant indices. That is, the Young diagram of the adjoint representation is the same as that of its dual — which is also its complex conjugate.

Alternatively, one may realize that the tensor product \( n \otimes \bar{n} \) is “obviously” self-conjugate, and that in its Clebsch–Gordan series (VII.26b) the trivial representation is also self-conjugate, so that the remaining adjoint representation must be self-conjugate as well.

As a consequence, the tensor product of the adjoint representation with itself contains the trivial representation in its Clebsch–Gordan series, as can be readily checked by performing the outer product of the associated Young diagrams.

\textbf{Example:} In the case of SU(2), for which we know the Clebsch–Gordan series (see theorem VI.49), the adjoint representation is the 3, which indeed appears in the tensor product of the self-conjugate 2-representation with itself. In a more physics-oriented language, the addition of two spins \( \frac{1}{2} \) (defining representation, 2) gives a triplet (spin 1, adjoint representation 3) and a singlet (spin 0, trivial representation 1).

Eventually, the addition of two spins 1 can give rise to a spin 0 (besides a spin 1 and a spin 2), i.e. the tensor product of two adjoint representations contains the trivial representation.

\textbf{Generators of the adjoint representation}

The dimension \( n^2 - 1 \) of the adjoint representation is also the number of real parameters of the Lie group SU\((n)\). As a matter of fact, the adjoint representation of SU\((n)\) can be deduced by the exponentiation from a representation of the Lie algebra su\((n)\), whose \( n^2 - 1 \) generators \( \{T^A\}_{A=1,...,n^2-1} \) are directly related to the structure constants \( f_{ab}^c \) of the group. More precisely, the \( BC \)-entry of the generator \( T^A \) is

\[
(T^A)^B = -i f^{AB} C \quad \text{for all } A, B, C \in \{1, \ldots, n^2 - 1\}.
\]  
(VII.27)

The Jacobi identity (V.30b) then expresses the property that the commutators of these generators obey the characteristic relation (V.29) involving the structure constants.

\textbf{Remark:} In Eq. (VII.27), we used the convenient (but far from universal!) convention in which labels running from 1 to \( n^2 - 1 \) are denoted with capital letters, while those running from 1 to \( n \) are denoted with lowercase indices.

For example, the 3 generators of the defining representation of SO(3), which is also the adjoint representation of SU(2), given in Eq. (VI.9) are precisely such that their matrix elements are related to the structure constants of SU(2) by in the above equation.
Consider the four-dimensional real vector space \( \mathbb{R}^4 \). Its vectors will generically be denoted in a sans-serif font, as e.g. \( x \). Assuming a basis has been chosen, the components of a vector will be denoted with Greek letters, running from 0 to 3: \( x^\mu \). The components \( x^1, x^2, x^3 \) will often be collectively represented by a three-vector \( \vec{x} \).

The introduction of a pseudo-Euclidean metric with signature \((-,+,+,+)\), corresponding to a metric tensor \( \eta \) with components

\[
\eta_{\mu\nu} \equiv \begin{cases} 
-1 & \text{for } \mu = \nu = 0 \\
+1 & \text{for } \mu = \nu \in \{1, 2, 3\} \\
0 & \text{for } \mu \neq \nu 
\end{cases}
\] (VIII.1a)

turns \( \mathbb{R}^4 \) into a pseudometric space, called Minkowski spacetime and hereafter denoted \( \mathcal{M}_4 \). It is convenient to associate with the metric tensor a \( 4 \times 4 \) matrix \( \eta \) defined as

\[
\eta = \text{diag}(-1, 1, 1, 1)
\] (VIII.1b)
i.e. such that its entries are precisely the components \( \eta_{\mu\nu} \).

### VIII.1 Definitions and first properties

With the metric tensor (VIII.1), the pseudo-distance between two infinitesimally close points with respective coordinates \( \{x^\mu\} \) and \( \{x^\mu + dx^\mu\} \) is the so-called line element

\[
ds^2 \equiv dx^\mu \eta_{\mu\nu} dx^\nu = -(dx^0)^2 + d\vec{x}^2,
\] (VIII.2a)

where \( d\vec{x}^2 \) denotes the squared Euclidean norm of the three-vector with components \( (dx^1, dx^2, dx^3) \).

Introducing the (column) four-vector \( dx \) with components \( dx^\mu \) and the transposed row-vector \( dx^\top \), one can equivalently write

\[
ds^2 = dx^\top \eta dx,
\] (VIII.2b)

where \( \eta \) is the matrix introduced in Eq. (VIII.1b).

#### VIII.1.1 Poincaré group

**Theorem & Definition VIII.3.** The transformations \( x \in \mathcal{M}_4 \rightarrow x' \in \mathcal{M}_4 \) that leave the infinitesimal line element (VIII.2) invariant, i.e. the Minkowski spacetime isometries, form a group, called the Poincaré group.

Requiring that the transformations should be at least twice continuously differentiable, one can show that they are necessarily affine \((45)\), i.e. of the form

\[x' = \Lambda x + \mathbf{a},\]

where \( \Lambda \) is a Lorentz transformation and \( \mathbf{a} \) is a constant.

\[(45) A\ proof\ is\ presented\ in\ Appendix\ VIII.A\ to\ this\ chapter.
\]

\[(46) H.\ Minkowski,\ 1864–1909\ ]\ (47) H.\ Poincaré,\ 1854–1912\]
\[ x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \quad \forall \mu \in \{0,1,2,3\}, \quad (VIII.4) \]

where the real numbers \( \{a^\mu\} \) are arbitrary, while the coefficients \( \Lambda^\mu_\nu \) obey the relation

\[ \Lambda^\mu_\rho \eta_{\rho\sigma} \Lambda^\nu_\sigma = \eta_{\mu\sigma} \quad \forall \rho, \sigma \in \{0,1,2,3\}. \quad \text{(VIII.5a)} \]

Equation (VIII.4) leads to

\[ dx'^\mu = \Lambda^\mu_\nu dx^\nu, \]

which in turn gives for the infinitesimal line element

\[ ds'^2 = dx'^\mu \eta_{\mu\nu} dx'^\nu = dx^\mu \eta_{\mu\nu} \Lambda^\nu_\sigma dx^\sigma = dx^\rho \eta_{\rho\sigma} \Lambda^\nu_\sigma dx^\nu. \]

By definition this should equal

\[ ds^2 = dx^\rho \eta_{\rho\sigma} dx^\sigma, \]

which is only possible for every \( dx^\mu \) if and only if Eq. (VIII.5a) holds.

Viewing the numbers \( \Lambda^\mu_\nu \) as the entries of a \( 4 \times 4 \) matrix \( \Lambda \), Eq. (VIII.5) is equivalent to the matrix identity

\[ \Lambda^T \eta \Lambda = \eta. \quad \text{(VIII.5b)} \]

where \( \eta \) was introduced in Eq. (VIII.1b). This condition can directly be derived by using the infinitesimal line element in its form (VIII.2b) and writing \( dx' = \Lambda dx \).

**Theorem VIII.6.** The spacetime translations \( x \rightarrow x' = x + a \), i.e. component-wise

\[ x^\mu \rightarrow x'^\mu = x^\mu + a^\mu \quad \forall \mu \in \{0,1,2,3\} \quad (VIII.6) \]

with \( a^\mu \in \mathbb{R} \), form a normal subgroup of the Poincaré group.

Another important subgroup of the Poincaré group, which however is not a normal subgroup, is that introduced in the following section.

**VIII.1.2 Lorentz group**

**Theorem & Definition VIII.7.** The linear transformations

\[ x \in \mathbb{M}_4 \rightarrow x' = \Lambda x \in \mathbb{M}_4 \quad \text{(VIII.7a)} \]

or equivalently

\[ x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu \quad \forall \mu \in \{0,1,2,3\} \quad (VIII.7b) \]

that preserve the infinitesimal line element (VIII.2) are called Lorentz transformations and form a group, the Lorentz group, denoted \( O(3,1) \).

**Remarks:**

* As the notation suggests, the Lorentz group is one of the indefinite orthogonal groups introduced in § V.1.2d.

* Obviously, the Lorentz transformations are the Minkowski spacetime isometries that leave a point of \( \mathbb{M}_4 \) invariant, i.e. such that \( a^\mu = 0 \) in Eq. (VIII.4). Accordingly, the Lorentz group is a subgroup of the Poincaré group.

\(^{(ah)}\)H. A. Lorentz, 1853–1926
VIII.1.2a Examples of Lorentz transformations

The matrices of the form

$$\Lambda_R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & R & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{with } R \in \text{SO}(3) \quad (VIII.8)$$

represent a first class of Lorentz transformations, whose action on a four-vector is a rotation of the spatial components while keeping the time-component unchanged: $(x^0, \vec{x}) \rightarrow (x^0 = x^0, \vec{x}' = R\vec{x})$. For instance, the rotation through $\psi$ around the $x_1$-axis is represented by the matrix

$$\Lambda_R(\vec{e}_1, \psi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \psi & -\sin \psi \\ 0 & 0 & \sin \psi & \cos \psi \end{pmatrix}. \quad (VIII.9)$$

**Theorem VIII.10.** The matrices of the form (VIII.8) form a subgroup of the Lorentz group. This subgroup is actually a (faithful and reducible) representation of $\text{SO}(3)$.

A second large class of Lorentz transformations consists of the so-called Lorentz boosts (or special Lorentz transformations). For such a boost with (reduced) velocity $\beta \in ]-1, 1[$ along the direction with unit vector $\vec{n} \in S^2$, the corresponding transformation reads

$$\begin{cases} x^0 \rightarrow x'^0 = x^0 \cosh \xi + (\vec{n} \cdot \vec{x}) \sinh \xi \\ \vec{x} \rightarrow \vec{x}' = \vec{x} + (\cosh \xi - 1)(\vec{n} \cdot \vec{x})\vec{n} + (\sinh \xi)x^0\vec{n}, \end{cases} \quad (VIII.11)$$

where $\xi \equiv \text{artanh } \beta \in \mathbb{R}$ is the rapidity of the boost. For a Lorentz boost along the $x^1$-direction, the associated transformation matrix is

$$\Lambda = \begin{pmatrix} \cosh \xi & \sinh \xi & 0 & 0 \\ \sinh \xi & \cosh \xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (VIII.12)$$

The Lorentz boosts do not form a group — successive boosts along non-parallel directions do not yield a boost, but the combination of a boost and and spatial rotation. However, the Lorentz boosts along a fixed (arbitrary) direction $\vec{n}$ do form a subgroup of the Lorentz group, which is isomorphic to $(\mathbb{R}, +)$.

Let us eventually mention three particular Lorentz transformations, which will help us characterize the connected components of the Lorentz group in the following paragraph.

- **The space inversion transformation** $(x^0, \vec{x}) \rightarrow (x'^0 = x^0, \vec{x}' = -\vec{x})$ reads in matrix representation

  $$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (VIII.13)$$

- **The time reversal transformation** $(x^0, \vec{x}) \rightarrow (x'^0 = -x^0, \vec{x}' = \vec{x})$ is represented by the matrix

  $$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (VIII.14)$$

\(^{(46)}\) The reader is invited to note the similarity between the form of this transformation and Rodrigues’ formula (VI.3) for rotations.
VIII.1 Definitions and first properties

- Eventually, the composition of the previous two transformations transforms $x$ into $x' = -x$, i.e. the corresponding matrix is the negative of the identity, $\Lambda_P \Lambda_T = \Lambda_T \Lambda_P = -I_4$.

These three transformations are obviously their own inverses, and one easily checks that the matrices $\{I_4, \Lambda_P, \Lambda_T, \Lambda_P \Lambda_T\}$ form a subgroup which is isomorphic to the Klein group $V_4$.

VIII.1.2b Structure of the Lorentz group

Taking the determinant of both sides of Eq. (VIII.5b) gives $(\det \Lambda)^2 \det \eta = \det \eta$ for every matrix $\Lambda$ of the Lorentz group, that is, since $\det \eta$ is non-zero,

$$\det \Lambda = \pm 1.$$  \hspace{1cm} (VIII.15)

Theorem & Definition VIII.16. The Lorentz transformations with determinant equal to 1, called proper Lorentz transformations, form a subgroup of the Lorentz group denoted $SO(3,1)$.

In contrast, the Lorentz transformations with determinant $-1$ are called improper, as e.g. the spatial-parity and time-reversal transformations (VIII.13)–(VIII.14).

Consider now the property of a Lorentz transformation for the case $\rho = \sigma = 0$:

$$\Lambda^\mu_{\nu} \eta_{\mu\nu} = \eta_{00} = -1.$$  \hspace{1cm} (VIII.17)

valid for all $\Lambda \in O(3,1)$. Accordingly, the coefficient $\Lambda^0_{0}$ is either greater than or equal to $+1$, or smaller than or equal to $-1$. More precisely, one easily shows the following two results:

Theorem & Definition VIII.18. The Lorentz transformations such that $\Lambda^0_{0} \geq 1$, called orthochronous Lorentz transformations, form a subgroup of the Lorentz group denoted $O^+(3,1)$.

Theorem & Definition VIII.19. The proper, orthochronous Lorentz transformations, i.e. such that $\det \Lambda = +1$ and $\Lambda^0_{0} \geq 1$, form a normal subgroup of the Lorentz group denoted $SO^+(3,1)$ and called restricted Lorentz group.

Theorem VIII.20. The most general element of the restricted Lorentz group $SO^+(3,1)$ is a product of a Lorentz boost (VIII.11) and a rotation transformation (VIII.8).

Property VIII.21. $SO^+(3,1)$ is the connected component of the identity in $O(3,1)$.

In fact, the Lorentz group — like every indefinite orthogonal group — consists of four connected components, which can be characterized with the help of the spatial inversion (VIII.13) and time reversal transformation (VIII.14):

- The connected component of the identity is the restricted Lorentz group $SO^+(3,1)$ introduced in definition VIII.19. Note that this component contains the subgroup of rotations (VIII.8).

- The connected component of the space inversion transformation (VIII.13) consists of the matrices of the form $\Lambda = \Lambda_P \Lambda_T^\dagger$, with $\Lambda_T^\dagger \in SO^+(3,1)$. These matrices $\Lambda$ are precisely the orthochronous ($\Lambda^0_{0} \geq 1$) improper ($\det \Lambda = -1$) Lorentz transformations.

- The connected component of the time reversal transformation (VIII.14) consists of the matrices of the form $\Lambda = \Lambda_T \Lambda_P^\dagger$ with $\Lambda_P^\dagger \in SO^+(3,1)$. These are the Lorentz transformations with $\det \Lambda = -1$ and $\Lambda^0_{0} \leq -1$ (“antichronous”).

- Eventually, the Lorentz transformations with $\det \Lambda = 1$ and $\Lambda^0_{0} \leq -1$ are of the form $-\Lambda_T^\dagger$, with $\Lambda_T^\dagger \in SO^+(3,1)$ and constitute the connected component of $-I_4$. 
Besides the topological meaning of the four subsets \( \text{SO}^+(3, 1), \Lambda_P \text{SO}^+(3, 1), \Lambda_T \text{SO}^+(3, 1), \) and \( \Lambda_P \Lambda_T \text{SO}^+(3, 1) \) mentioned above, one can also note that these are the cosets of the restricted Lorentz group in \( \text{O}(3, 1) \), i.e.

\[
\text{O}(3, 1)/\text{SO}^+(3, 1) \cong \{ \mathbb{I}_4, \Lambda_P, \Lambda_T, \Lambda_P \Lambda_T \} \cong \mathbb{V}_4,
\]

where the second isomorphism was already mentioned at the end of § VIII.1.2a.

### VIII.2 The Lorentz and Poincaré groups as Lie groups

According to theorem (VIII.20), any element of the restricted Lorentz group \( \text{SO}^+(3, 1) \) can be written as the product of a Lorentz boost and a three-dimensional rotation. As the reader knows from Sec. VI.1 the latter are characterized by three real parameters. In turn, a general Lorentz boost (VIII.11) also depends on three real parameters, namely the rapidity \( \xi \) and the two parameters associated with the unit three-vector along which the boost takes place.

All in all, the Lorentz group \( \text{O}(3, 1) \) is thus a 6-parameter group. Since the rapidity can take any value in \( \mathbb{R} \), this is a non-compact group.

Turning to the Poincaré group, one deduces from the generic form (VIII.4) of a Minkowski spacetime isometry that it is a non-compact 10-parameter group. Thus, 6 real parameters correspond to the Lorentz transformations \( (\Lambda^\mu_\nu) \), and 4 to the spacetime translations \( (a^\mu) \).

In this section, we shall investigate the generators of both groups, starting with those of the (restricted) Lorentz group (Sec. VIII.2.1), and dealing afterwards with the Poincaré group (Sec. VIII.2.2). Note, however, that since \( \text{O}(3, 1) \) is a subgroup of the Poincaré group, one may instead begin with the latter and consider the former afterwards.

#### VIII.2.1 Generators of the Lorentz group

Starting from the explicit form of spatial rotation transformations (VIII.8)–(VIII.9) and Lorentz boosts (VIII.11)–(VIII.12), and invoking Eq. (V.33), one can derive an expression for the associated generators.

Thus, the spatial rotations around the \( x^1, x^2 \) and \( x^3 \) axes will be respectively generated by

\[
J_1 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & -i & 0 \\
0 & i & 0 & 0
\end{pmatrix}, \quad J_2 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & i \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}, \quad J_3 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & -i & 0 \\
0 & i & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix},
\]

(VIII.23)

which trivially generalizes the three-dimensional generators (VI.9a). In turn, the generators of the Lorentz boosts along the \( x^1, x^2 \) and \( x^3 \) directions read

\[
K_1 = \begin{pmatrix}
0 & i & 0 & 0 \\
i & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}, \quad K_2 = \begin{pmatrix}
0 & 0 & i & 0 \\
0 & 0 & 0 & 0 \\
i & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}, \quad K_3 = \begin{pmatrix}
0 & 0 & 0 & i \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
i & 0 & 0 & 0
\end{pmatrix},
\]

(VIII.24)

**Remark VIII.25.** While the generators of rotations (VIII.23) are Hermitian, i.e. lead after exponentiation to unitary — in fact, orthogonal — matrices, those of Lorentz boosts (VIII.24) are not: they are antihermitian, and thus generally yield non-unitary matrices.

The generators (VIII.23)–(VIII.24) obey the commutation relations

\[
[J_i, J_j] = i \sum_{k=1}^3 \epsilon_{ijk} J_k \quad \forall i, j \in \{1, 2, 3\},
\]

(VIII.26a)
which merely expresses the fact that the matrices \( \{ J_i \} \) are generators of a representation of SO(3), as well as

\[
[K_i, K_j] = -i \sum_{k=1}^{3} \epsilon_{ijk} J_k \quad \forall i, j \in \{1, 2, 3\},
\]

(VIII.26b)

and

\[
[J_i, K_j] = i \sum_{k=1}^{3} \epsilon_{ijk} K_k \quad \forall i, j \in \{1, 2, 3\}.
\]

(VIII.26c)

The latter relation means that the matrices \( \{ K_i \} \) form a vector operator under SO(3) rotations, cf. §VI.4.2b. In turn, the commutation relation (VIII.26b) means that the product of Lorentz boosts along different directions is not a pure Lorentz boost, but also involves a rotation — i.e. that the Lorentz boosts do not form a subgroup of the Lorentz group.

### VIII.2.2 Generators of the Poincaré group

#### VIII.2.2 a Infinitesimal Lorentz and Poincaré transformations

Consider the generic expression (VIII.4) of a Minkowski spacetime isometry. For a transformation lying infinitesimally close to the identity, one may write

\[
x \to x' = (1 + \omega) x + \epsilon,
\]

(VIII.27a)

or equivalently, in terms of components

\[
x^\mu \to x'^\mu = (\delta^\mu_\nu + \omega^\mu_\nu) x^\nu + \epsilon^\mu \quad \forall \mu \in \{0, 1, 2, 3\},
\]

(VIII.27b)

where the (real) numbers \( \omega^\mu_\nu, \epsilon^\mu \) are much smaller than 1 in norm. In particular, the entries of an infinitesimal Lorentz transformation — which is necessarily an orthochronous, proper Lorentz transformation — take the form

\[
\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu,
\]

(VIII.28)

where the properties of the numbers \( \omega^\mu_\nu \), equivalent to Eq. (VIII.5a), still need to be specified.

With the above ansatz, one finds

\[
\Lambda^\rho_\mu \eta_{\mu\nu} \Lambda^\nu_\sigma = (\delta^\rho_\mu + \omega^\rho_\mu) \eta_{\mu\nu} (\delta^\nu_\sigma + \omega^\nu_\sigma) \simeq \delta^\rho_\mu \eta_{\mu\nu} \delta^\nu_\sigma + \delta^\rho_\mu \eta_{\mu\sigma} \omega^\nu_\sigma + \omega^\rho_\mu \eta_{\mu\sigma} \delta^\nu_\sigma,
\]

where in the second identity the quadratic contribution in \( \omega \) was neglected. The first term on the right hand side is precisely \( \delta^\rho_\mu \eta_{\mu\nu} \delta^\nu_\sigma = \eta_{\rho\sigma} \), so that Eq. (VIII.5a) becomes

\[
\omega_{\rho\sigma} + \omega_{\sigma\rho} = 0 \quad \forall \rho, \sigma \in \{0, 1, 2, 3\}.
\]

(VIII.29)

That is, the \( 4 \times 4 \) matrix with entries \( \omega_{\rho\sigma} \) — which is not the matrix \( \omega \) introduced in Eq. (VIII.27a) — is antisymmetric, and thus entirely determined by the entries in the triangle above the diagonal, i.e. the 6 elements with \( \sigma > \rho \). One thus recovers the fact that a generic matrix of the restricted Lorentz group depends on 6 real parameters — and in turn, adding the 4 extra parameters \( \epsilon^\mu \), that the Poincaré group is a 10-parameter group.

#### VIII.2.2 b Generators

Equation (VIII.27b) together with the condition (VIII.29) allow one to easily re-derive the set of six \( 4 \times 4 \) generators of the Lorentz group already found in Sec. VIII.2.1, yet leave the generators of translations aside.

To find a set of generators that also includes the translations, we look at operators acting on functions of the spacetime coordinates \( x \) — i.e. physically on fields. Considering an infinitesimal transformation (VIII.27) of the coordinates, a trivial Taylor expansion to first order gives for any sufficiently differentiable function \( f \) defined on \( \mathcal{M}_4 \)

\[
f(x') = f(x^\mu + \epsilon^\mu + \omega^\mu_\nu x^\nu) \simeq f(x) + (\epsilon^\mu + \omega^\mu_\nu x^\nu) \frac{\partial f(x)}{\partial x^\mu},
\]

(VIII.30)
which can be recast in the form
\[ f(x') = \left( \text{Id} + i\epsilon^{\mu}P_\mu - \frac{i}{2}\omega^{\mu\nu}J_{\mu\nu} \right) f(x) \]
where we have defined differential operators
\[ P_\mu \equiv -i\frac{\partial}{\partial x^\mu} \equiv -i\partial_\mu \quad \text{(VIII.30a)} \]
and
\[ J_{\mu\nu} \equiv -i(x_\mu \partial_\nu - x_\nu \partial_\mu) \quad \text{(VIII.30b)} \]
for all \( \mu, \nu \in \{0, 1, 2, 3\} \). These operators are the generators of the action of the Poincaré group on the functions on \( \mathbb{R}^4 \). One then easily finds that they obey the commutation relations
\[ [P_\mu, P_\nu] = 0 \quad \text{for all } \mu, \nu \in \{0, 1, 2, 3\}, \quad \text{(VIII.31a)} \]
which expresses that commutativity of translations, together with
\[ [J_{\mu\nu}, P_\rho] = i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu) \quad \text{for all } \mu, \nu, \rho \in \{0, 1, 2, 3\} \quad \text{(VIII.31b)} \]
and eventually
\[ [J_{\mu\nu}, J_{\rho\sigma}] = i(\eta_{\mu\rho}J_{\nu\sigma} - \eta_{\nu\rho}J_{\mu\sigma} + \eta_{\nu\sigma}J_{\mu\rho} - \eta_{\mu\sigma}J_{\nu\rho}) \quad \text{for all } \mu, \nu, \rho, \sigma \in \{0, 1, 2, 3\} \quad \text{(VIII.31c)} \]

Let us for instance check Eq. \( \text{(VIII.34b)} \): given a function \( f \) on \( \mathbb{R}^4 \), one has
\[
[J_{\mu\nu}, P_\rho] f = -i(x_\mu \partial_\nu - x_\nu \partial_\mu)(-i\partial_\rho)f + i\partial_\rho[-i(x_\mu \partial_\nu - x_\nu \partial_\mu)] f
\]
\[
= -x_\mu \partial_\nu \partial_\rho f + x_\nu \partial_\mu \partial_\rho f - x_\nu \partial_\mu \partial_\rho f - x_\nu \partial_\mu \partial_\rho f.
\]
Using the product rule in the last two terms together with \( \partial_\nu x_\mu = \eta_{\mu\nu} \), one finds
\[
[J_{\mu\nu}, P_\rho] f = -x_\mu \partial_\nu \partial_\rho f + x_\nu \partial_\mu \partial_\rho f + \eta_{\mu\rho} \partial_\nu f + x_\nu \partial_\mu \partial_\rho f - \eta_{\nu\rho} \partial_\mu f - x_\nu \partial_\mu \partial_\rho f.
\]
The first and fourth terms compensate, as do the second and sixth, leaving
\[
[J_{\mu\nu}, P_\rho] f = \eta_{\mu\rho} \partial_\nu f - \eta_{\nu\rho} \partial_\mu f = i\eta_{\mu\rho}(-i\partial_\nu f) - i\eta_{\nu\rho}(-i\partial_\mu f) = i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu) f.
\]
\[\square\]

**Remark VIII.32.** Restricting oneself to the components \( i, j = 1, 2, 3 \), the definitions \( \text{(VIII.30)} \) read
\( P_i = -i\partial_i \) and \( J_{ij} = -i(x_i \partial_j - x_j \partial_i) \), which are up to a factor \( \hbar \) the linear-momentum and angular-momentum operators acting on wave functions in spatial representation in non-relativistic quantum mechanics.

In turn, \( P^0 = -P_0 = i\partial_0 \) can be interpreted — up to irrelevant factors of \( \hbar \) and \( c \) — as the Hamilton operator. Accordingly, we shall use \( P^0 \) rather than \( P_0 \) in equations which will be given hereafter.

Let us now define for \( i = 1, 2, 3 \) the operators
\[ J_i \equiv \frac{1}{2} \sum_{j,k=1}^{3} \epsilon_{ijk} J_{jk}, \quad K_i \equiv J_{0i}, \quad \text{(VIII.33a)} \]
where the first definition can be “inverted” as
\[ J_{ij} = \sum_{k=1}^{3} \epsilon_{ijk} J_k. \quad \text{(VIII.33b)} \]
With these definitions, the commutation relations \((\text{VIII.31})\) become on the one hand

\[
\begin{align*}
[J_i, J_j] &= i \sum_{k=1}^{3} \epsilon_{ijk} J_k, \\
[K_i, K_j] &= -i \sum_{k=1}^{3} \epsilon_{ijk} J_k, \\
[J_i, K_j] &= i \sum_{k=1}^{3} \epsilon_{ijk} K_k \quad \forall i, j \in \{1, 2, 3\},
\end{align*}
\] (VIII.34a)

which are precisely the relations \((\text{VIII.26})\) already found for generators of the Lorentz group. In addition, one also finds

\[
\begin{align*}
[J_i, P_j] &= i \sum_{k=1}^{3} \epsilon_{ijk} P_k \quad \forall i, j \in \{1, 2, 3\},
\end{align*}
\] (VIII.34b)

which means that the three generators \(P_1, P_2, P_3\) form a vector representation of the subgroup \(SO(3)\) of the Poincaré group, as well as

\[
\begin{align*}
[K_i, P_j] &= i \delta_{ij} P^0 \quad \forall i, j \in \{1, 2, 3\},
\end{align*}
\] (VIII.34c)

where we used \(P^0 = -P_0\). One further has

\[
\begin{align*}
[P^0, J_i] &= 0 \quad \text{and} \quad [P^0, P_i] = 0 \quad \forall i \in \{1, 2, 3\},
\end{align*}
\] (VIII.34d)

which may be viewed as expressing momentum and angular-momentum conservation. Eventually, one has

\[
\begin{align*}
[P^0, K_i] &= -i P_i \quad \forall i \in \{1, 2, 3\}.
\end{align*}
\] (VIII.34e)
Appendix to Chapter VIII

**VIII.A Proof of the linearity of Minkowski spacetime isometries**

In this appendix, we show that the transformations \( x^\mu \to x'^\mu \) that preserve the infinitesimal line element \( ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \) are affine transformations.

Let \( x^\mu \to x'^\mu(x^\nu) \) be such a transformation. As we shall see hereafter, the functions \( x'^\mu(x) \) should be at least twice continuously differentiable to ensure the validity of the proof — physically, this requirement is however quite reasonable. Using the chain rule, the differential of the coordinate function \( x'^\rho \) reads

\[
d x'^\rho = \frac{\partial x'^\rho}{\partial x^\mu} dx^\mu ,
\]

which allows one to write the infinitesimal line element expressed in terms of the primed components:

\[
ds'^2 \equiv \eta_{\rho\sigma} dx'^\rho dx'^\sigma = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} dx^\mu dx^\nu .
\]

The condition \( ds'^2 = ds^2 \) with \( ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \) thus translates into

\[
\eta_{\mu\nu} = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \text{ for all } \mu, \nu \in \{0, 1, 2, 3\} .
\]

One can now differentiate both sides of this equations with respect to the coordinate \( x^\lambda \) with \( \lambda \in \{0, 1, 2, 3\} \), which thanks to the product rule yields

\[
0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\mu \partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu} + \frac{\partial x'^\rho}{\partial x^\nu} \frac{\partial^2 x'^\sigma}{\partial x^\mu \partial x^\lambda} \right) \forall \lambda, \mu, \nu \in \{0, 1, 2, 3\} .
\]

Exchanging the dummy indices \( \lambda \) and \( \mu \) resp. \( \lambda \) and \( \nu \), this equation gives

\[
0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\mu \partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu} + \frac{\partial x'^\rho}{\partial x^\nu} \frac{\partial^2 x'^\sigma}{\partial x^\mu \partial x^\lambda} \right) \forall \lambda, \mu, \nu .
\]

resp.

\[
0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\nu \partial x^\sigma} \frac{\partial x'^\lambda}{\partial x^\mu} + \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial^2 x'^\lambda}{\partial x^\nu \partial x^\sigma} \right) \forall \lambda, \mu, \nu .
\]

One can now add Eqs. \((\text{VIII.38a})\) and \((\text{VIII.38b})\) and subtract Eq. \((\text{VIII.38c})\). The second term of Eq. \((\text{VIII.38a})\) cancels out with that of Eq. \((\text{VIII.38c})\) while, thanks to the identity \( \eta_{\rho\sigma} = \eta_{\sigma\rho} \), the first terms of Eq. \((\text{VIII.38b})\) and \((\text{VIII.38c})\) compensate. Eventually, the two remaining terms are equal:

\[
0 = 2 \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\nu} \frac{\partial^2 x'^\sigma}{\partial x^\mu \partial x^\lambda} \forall \lambda, \mu, \nu \in \{0, 1, 2, 3\} .
\]

This equation can be interpreted as representing the product from left to right of a row vector \( \partial^2 x'^\rho/\partial x^\mu \partial x^\lambda \) (with four components \( \rho = 0, 1, 2, 3 \)) with two \( 4 \times 4 \)-matrices with respective entries \( 2 \eta_{\rho\sigma} \) and \( \partial x'^\sigma/\partial x^\nu \). If the latter matrix is regular, i.e. if the transformation \( x^\mu \to x'^\mu \) can be inverted, one can multiply the above equation with the inverse matrix, which leads to the condition

\[
\frac{\partial^2 x'^\rho}{\partial x^\mu \partial x^\lambda} = 0 \forall \lambda, \mu, \rho \in \{0, 1, 2, 3\} .
\]
After twofold integration, this equation leads first to $\partial x'^\rho/\partial x^\mu = \Lambda^\rho_\mu$ with $\Lambda^\rho_\mu \in \mathbb{R}$, then to the affine transformation

$$x'^\rho = \Lambda^\rho_\mu x^\mu + a^\rho,$$

i.e. to the announced result (VIII.4).
CHAPTER IX
Symmetries in classical field theory

intro

IX.1 Basics of classical field theory

Consider independent scalar fields \( \varphi_a(x) \), \( a = 1, \ldots, N \), defined at each point \( x \) of a \( D \)-dimensional space-time \( \mathcal{M}_D \) on which a metric with signature \((-+, +, +, \cdots , +)\) is defined — the obvious case, which we consider in the following, being that of the 4-dimensional Minkowski\(^{(a)}\) space \( \mathcal{M}_4 \) of special relativity. As we shall illustrate on a few examples, the fields \( \varphi_a(x) \) may be real- or complex-valued, and may represent the components in a given reference frame of a vector or tensor field.

IX.1.1 Definitions

The dynamical behavior of the fields is described by a Lagrange density

\[
\mathcal{L}[\{\varphi_a(x)\}, \{\partial_\mu \varphi_a(x)\}],
\]

which is a functional of the fields and of their derivatives \( \partial_\mu \varphi_a(x) \) with \( \mu \in \{0, 1, 2, 3\} \), where \( \partial_\mu \) denotes the partial derivative with respect to the coordinate \( x^\mu \) of the space-time point \( x \) in a given reference frame.

Integrating the Lagrange density over the spatial coordinates yields the Lagrangian

\[
\mathcal{L} = \int \mathcal{L}[\{\varphi_a(x)\}, \{\partial_\mu \varphi_a(x)\}] d^3 \vec{r}.
\]

In turn, integrating the Lagrangian over a time interval \([t_1, t_2]\) gives the action

\[
S = \int_{t_1}^{t_2} L \, dt = \int \mathcal{L}[\{\varphi_a(x)\}, \{\partial_\mu \varphi_a(x)\}] d^4x
\]

with \( d^4x \) the 4-volume element, \( d^4x = dx^0 dx^1 dx^2 dx^3 \).

Remarks:

* If the Lagrange density of a system of fields is invariant under the transformations of the Lorentz or Poincaré groups, then this also holds for the resulting action, since \( d^4x \) is Lorentz-invariant.

* The Lagrange density for a given physical system is not unique. Anticipating on next section, two Lagrange densities \( \mathcal{L} \) and \( \mathcal{L}' \) that only differ by a 4-divergence \( \partial_\mu \mathcal{M}^\mu(x) \), with \( \mathcal{M} \) a 4-vector field, are equivalent: their integrals give respective actions \( S, S' \), which lead to the same equations of motion.

\(^{(a)}\)H. Minkowski, 1864–1909
**IX.1.2 Euler–Lagrange equations**

According to the principle of the stationary action, or Hamilton’s principle, the equations of motion obeyed by a system of fields can be obtained by demanding that the action $S$ should be stationary when the field configuration is varied, assuming that the fields do not vary at the boundaries of the space-time region over which the action is computed.

Considering a (small) variation $\delta \varphi_a$ of each field $\varphi_a$, which results in a variation $\delta \partial_\mu \varphi_a = \partial_\mu \delta \varphi_a$ of its partial derivatives, the ensuing variation of the action (IX.1c) reads

$$\delta S = \int \sum_a \left[ \frac{\partial L}{\partial \varphi_a} \delta \varphi_a + \frac{\partial L}{\partial \partial_\mu \varphi_a} \delta \partial_\mu \varphi_a \right] d^4x,$$

where the second term involves a non-written sum over the Lorentz indices $\mu$. When integrating that term by partial integration, the contribution from the integrated term is zero, thanks to the assumption that $\delta \epsilon_a$ vanishes on the boundaries of the integration region. One thus obtains

$$\delta S = \int \sum_a \left[ \frac{\partial L}{\partial \varphi_a} - \partial_\mu \frac{\partial L}{\partial \partial_\mu \varphi_a} \right] \delta \varphi_a d^4x.$$

The stationarity of the action is realized when $\delta S = 0$ for arbitrary variations $\delta \varphi_a$, i.e. when each field $\varphi_a$ and its derivatives obey the Euler–Lagrange equations

$$\frac{\partial L}{\partial \varphi_a} = \partial_\mu \frac{\partial L}{\partial \partial_\mu \varphi_a}.$$

These represent the equations of motion of the fields.

**Example:** The simplest possible example is that of a single real scalar field $\varphi$ with the Lagrange density

$$L[\varphi, \partial_\mu \varphi] = -\frac{1}{2} \left( \partial_\mu \varphi \right) \left( \partial^\mu \varphi \right) - \frac{1}{2} m^2 \varphi^2.$$  

Straightforward calculations, where one should not forget the identity $\partial^\mu \varphi = \eta^\mu_\nu \partial_\nu \varphi$, yield

$$\frac{\partial L}{\partial \varphi} = -m^2 \varphi \quad \text{and} \quad \frac{\partial L}{\partial \partial_\mu \varphi} = -\partial_\mu \varphi,$$

so that the Euler–Lagrange equation gives $-m^2 \varphi = -\partial_\mu \partial^\mu \varphi$, that is, restoring the $x$-dependence of the field

$$\partial_\mu \partial^\mu \varphi(x) - m^2 \varphi(x) = 0$$

which is known as the Klein&(aj)–Gordon&(ak) equation.

**Remark:** In analogy to the construction in analytical mechanics, one associates to each field $\varphi_a$ governed by the Lagrange density $L$ a **conjugate momentum**

$$\pi_a = \frac{\partial L}{\partial \partial_0 \varphi_a}.$$  

With its help, one defines the **Hamilton density**

$$\mathcal{H}[\{\pi_a\}, \{\varphi_a\}] = \sum_a \pi_a \partial_0 \varphi_a - L,$$

whose integral over the spatial coordinates gives the Hamilton function of the system.

For the real scalar field of the previous example, one at once finds $\pi = -\partial^0 \varphi = \partial_0 \varphi$, which yields $\mathcal{H} = \frac{1}{2} \left( \partial_0 \varphi \right)^2 + \frac{1}{2} \left( \nabla \varphi \right)^2 + \frac{1}{2} m^2 \varphi^2$.

(a)O. Klein, 1894–1977  (ak)W. Gordon, 1893–1939
Symmetries in classical field theories and conserved quantities

An important feature of classical field theory lies in its relation to symmetries. On the one hand, implementing a symmetry into a given theory is rather simple, as we shall see in the following section. On the other hand, if a continuous symmetry is present, then there exist corresponding conserved quantities (or set of conserved quantities), as we shall now see.

Symmetries in a field theory

Using the Lagrange formalism, one can easily study the behavior of a system of classical fields under various types of transformations. For physical applications, two big classes of such transformations are of interest, which may possibly be applied simultaneously.

On the one hand, one can consider space-time transformations, affecting the “position” of the field system. These are generically of the form

\[ x \rightarrow x' = f(x, \{\theta_c\}) \tag{IX.7a} \]

where \( f \) denotes a sufficiently regular function depending on real parameters \( \{\theta_c\} \). Introducing Minkowski coordinates, the transformation reads

\[ x^\mu \rightarrow x'^\mu = f^\mu(\{x^\nu\}, \{\theta_c\}) \tag{IX.7b} \]

for \( \mu = 0, 1, 2, 3 \) (or up to \( D-1 \) in the more general case of a \( D \)-dimensional space-time \( \mathcal{M}_D \)).

On the other hand, one may also consider “internal” transformations of the fields, of the form

\[ \varphi_a(x) \rightarrow \varphi'_a(x) = g_a(\{\varphi_b(x)\}, \{\Lambda_c(x)\}) \tag{IX.8} \]

where \( g_a \) is some function depending on real parameters \( \{\Lambda_c(x)\} \). An important possibility, corresponding to so-called gauge transformations, is that the latter parameters indeed depend on the space-time position of the fields. In case of space- and time-independent parameters, one also talks of global transformations.

A group of transformations is said to be a symmetry of a physical system described by fields if the associated action \( \{\text{IX.1c}\} \) is invariant under the transformations.

Noether theorem

Let us now show that every continuous group of symmetry transformations of the action \( S \) of a system leads to the existence of a continuous quantity, which constitutes the Noether\( ^{(a)} \) theorem.

Result and discussion

More precisely, we shall find that if the integral of a Lagrange density \( \mathcal{L}[\{\varphi_a(x)\}, \{\partial_\mu \varphi_a(x)\}] \) over a 4-volume \( \Omega \) — which will be further specified hereafter — is invariant under simultaneous infinitesimal transformations

\[ x^\mu \rightarrow x'^\mu = x^\mu + \delta x^\mu \quad \forall \mu = 0, 1, 2, 3, \tag{IX.9a} \]

\[ \varphi_a(x) \rightarrow \varphi'_a(x') = \varphi_a(x) + \delta \varphi_a(x) \quad \forall a \tag{IX.9b} \]

of the space-time coordinates and the fields, then the Noether current with components

\[ J^\mu \equiv \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \delta \varphi_a + T^{\mu\nu}_{\text{can}} \delta x^\nu \tag{IX.10} \]

obeys

\[ \int_\Omega \partial_\mu J^\mu(x) \, d^4 x = 0, \tag{IX.11} \]

\((a)\) E. Noether, 1882–1935
where the \textit{canonical energy-momentum tensor} is defined by

\begin{equation}
T_{\text{can}}^{\mu
u} \equiv \eta^{\mu\nu} \mathcal{L} - \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \partial^\nu \phi_a. \tag{IX.12}
\end{equation}

In the above, the 4-volume $\Omega$ is delimited in the temporal direction by two constant-time hyperplanes $t = t_1$ and $t = t_2$, while it extends to infinity in the spatial directions, in agreement with Eqs. (IX.1b) and (IX.1c). That is, the boundary $\partial \Omega$ of $\Omega$ consists of the two hyperplanes $t = t_1$ and $t = t_2$ and of the hypersurface $\partial \mathbb{R}^3$ at spatial infinity for every $t \in [t_1, t_2]$ — where one assumes that every field vanishes together with its derivatives. Accordingly, an integral over $\partial \Omega$ actually reduces to integrals over the spatial coordinates $\bar{\mathbf{r}} \in \mathbb{R}^3$ at $t = t_1$ and $t = t_2$. In particular, the divergence theorem reads for any 4-vector-field $B^\mu$

\[
\int_{\Omega} \partial_\mu B^\mu(\mathbf{x}) \, d^4\mathbf{x} = \int_{\partial \Omega} B^\mu(\mathbf{x}) \, d^3\sigma_\mu = \left[ \int_{\mathbb{R}^3} B^\mu(t, \bar{\mathbf{r}}) \, d^3\bar{\mathbf{r}} \right]_{t = t_2}^{t = t_1},
\]

where we used the fact that on the hyperplane $t = t_1$ resp. $t = t_2$, the outwards-oriented hypersurface element $d^3\sigma_\mu$ is purely along the time direction with $d^3\sigma_0 = d^3\bar{\mathbf{r}}$ resp. $-d^3\bar{\mathbf{r}}$.

Applying the latter result to the relation (IX.11) obeyed by the Noether current (IX.10) yields

\[
\int_{\Omega} \partial_\mu J^\mu(\mathbf{x}) \, d^4\mathbf{x} = 0 = \left[ \int_{\mathbb{R}^3} J^0(t, \bar{\mathbf{r}}) \, d^3\bar{\mathbf{r}} \right]_{t = t_2}^{t = t_1},
\]

which shows that the \textit{Noether charge}

\[
Q \equiv \int_{\mathbb{R}^3} J^0(t, \bar{\mathbf{r}}) \, d^3\bar{\mathbf{r}}
\]

is a conserved quantity, a constant of motion.

\section*{Remarks:}

* Using $\eta^{00} = -1$ and $\partial^0 \phi_a = -\partial_0 \phi_a$, definition (IX.12) shows that the 00-component of the canonical energy-momentum tensor coincides with the Hamilton density (IX.6): $T_{\text{can}}^{00} = \mathcal{H}$.

* Strictly speaking, $J^\mu$ is a current density.

* Instead of the invariance of the action, one often finds in the literature the stronger requirement of having an invariant Lagrange density $\mathcal{L}$ — which of course automatically leads to an invariant action. This stronger condition has the advantage that it leads to the "local conservation equation" $\partial_\mu J^\mu(\mathbf{x}) = 0$ in lieu of Eq. (IX.11), yet it is not needed for the existence of the conserved Noether charge.

\section*{IX.2.2 Proof}

Let us first rewrite the expression of action of the fields

\[
S = \int_{\Omega} \mathcal{L} \left[ \{ \phi_a(\mathbf{x}) \}, \{ \partial_\mu \phi_a(\mathbf{x}) \} \right] d^4\mathbf{x} \tag{IX.14}
\]

valid for the transformed fields and coordinates (IX.9). When transforming the latter, one simultaneously modifies the integration 4-volume: $\Omega \to \Omega'$. Accordingly, the transformed action reads

\[
S' = \int_{\Omega'} \mathcal{L} \left[ \{ \phi'_a(\mathbf{x}') \}, \{ \partial_\mu \phi'_a(\mathbf{x}') \} \right] d^4\mathbf{x}'.
\]

Introducing the boundary $\partial \Omega$ of the 4-volume $\Omega$ — which is also, up to terms of order $\delta x^\mu$, the boundary of $\Omega'$ —, one may decompose the integral over $\Omega'$ into the sum of an integral over $\Omega$ and of an integral over the layer of thickness $\delta x^\mu$ that corresponds to the difference $\Omega' \setminus \Omega$:
\[
S' = \int_{\Omega} \mathcal{L} \{\varphi_a'(x)\}, \{\partial_\mu \varphi_a'(x)\}\,d^4x + \oint_{\partial \Omega} \mathcal{L} \{\varphi_a(x)\}, \{\partial_\mu \varphi_a(x)\}\,\delta x^\mu \,d^3\sigma_\mu, \tag{IX.15}
\]

where \(d^3\sigma_\mu\) is a normal (hyper)surface 4-vector at each point of the boundary \(\partial \Omega\). Note that every space-time position is now expressed in terms of the “old” positions \(x\), which allows us to suppress it from now on for the sake of brevity. Accordingly, we introduce the variation of the field \(\varphi_a\) at a fixed point:

\[
\delta \varphi_a(x) = \varphi_a'(x) - \varphi_a(x),
\]

which will generally differ from \(\delta \varphi_a(x)\) if \(x'\) and \(x\) are different, i.e. if the space-time coordinates are transformed: to first order in the variation \(\delta x\), one has

\[
\delta \varphi_a = \varphi_a'(x') - \varphi_a(x) \simeq \delta x_\mu \partial^\mu \varphi_a(x) - \varphi_a(x)
\]

i.e.

\[
\delta \varphi_a \simeq \delta \varphi_a(x) - \delta x_\mu \partial^\mu \varphi_a(x). \tag{IX.16}
\]

Coming back to Eq. (IX.15), one may write \(\varphi_a'(x) = \varphi_a(x) + \delta \varphi_a(x)\) in the arguments of the Lagrange density and expand to first order in the variations \(\delta x\), \(\delta \varphi_a\), and \(\delta (\partial_\mu \varphi_a) = \partial_\mu (\delta \varphi_a)\):

\[
S' \simeq \int_{\Omega} \mathcal{L} \{\varphi_a\}, \{\partial_\mu \varphi_a\}\,\sum_a \left( \frac{\partial \mathcal{L}}{\partial \varphi_a} \delta \varphi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \partial_\mu (\delta \varphi_a) \right) d^4x
\]

\[
+ \oint_{\partial \Omega} \mathcal{L} \{\varphi_a\}, \{\partial_\mu \varphi_a\}\,\delta x^\mu \,d^3\sigma_\mu.
\]

Note that in the integral over the boundary \(\partial \Omega\), we have suppressed terms stemming from the expansion of the Lagrange density that yield contributions of at least second order in the variations. From now on, we shall also omit the argument \(\{\varphi_a\}, \{\partial_\mu \varphi_a\}\) of the Lagrange density and its (functional) derivatives, since it is the same in every term.

One may now subtract the action (IX.14) from the new expression of \(S'\), which is straightforward. Performing an integration by parts for the term in the 4-volume integral involving the partial derivatives \(\partial_\mu (\delta \varphi_a)\), one obtains

\[
\delta S = \oint_{\partial \Omega} \sum_a \left( \frac{\partial \mathcal{L}}{\partial \varphi_a} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \right) \right) \delta \varphi_a d^4x + \oint_{\partial \Omega} \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \delta \varphi_a + \oint_{\partial \Omega} \mathcal{L} \delta x^\mu \,d^3\sigma_\mu.
\]

Thanks to the Euler–Lagrange equations (IX.2), the integrand of the 4-volume integral identically vanishes, since we assume that the field configuration \(\{\varphi_a(x)\}\) is a solution of the equations of motion. In turn, expressing \(\delta \varphi_a\) through \(\delta \varphi_a\) [Eq. (IX.16)] in the integral over \(\partial \Omega\) yields

\[
\delta S = \oint_{\partial \Omega} \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} (\delta \varphi_a - \delta x_\mu \partial^\mu \varphi_a) + \mathcal{L} \delta x^\mu \,d^3\sigma_\mu.
\]

After some straightforward rewriting, this becomes

\[
\delta S = \oint_{\partial \Omega} \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \delta \varphi_a + \left( \eta^\mu \mathcal{L} - \sum_a \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_a)} \partial^\mu \varphi_a \right) \delta x_\mu \,d^3\sigma_\mu. \tag{IX.17}
\]

Using the definitions of the Noether current (IX.10) and the canonical energy-momentum tensor (IX.12), this becomes

\[
\delta S = \oint_{\partial \Omega} J^\mu \,d^3\sigma_\mu.
\]

The action is stationary for any infinitesimal transformation of the fields and space-time coordinates if its variation \(\delta S\) vanishes, which invoking the divergence theorem precisely gives the announced result (IX.11) — which concludes the proof.
IX.2.3 Examples

IX.2.3 a Energy and momentum of a field system

Considering a Lagrange density which is invariant under space-time translations

\[ x^\mu \rightarrow x'^\mu = x^\mu + a^\mu \]  (IX.18)

with \( a^\mu \in \mathbb{R} \) for every \( \mu \in \{0, 1, 2, 3\} \), one finds by setting \( \delta \phi_a = 0 \) and \( \delta x_\nu = a_\nu \) in the expression (IX.10) of the Noether current that the latter becomes proportional to the canonical energy-momentum tensor. Accordingly, the four components

\[ P^\nu = \int_{\mathbb{R}^3} T^0_\text{can}(t, \vec{r}) \, d^3 \vec{r} \]

for \( \nu = 0, 1, 2, 3 \),

are conserved quantities, which corresponds to the energy (\( P^0 \)) and the momentum of the field system, and form a four-vector.

IX.2.3 b Complex scalar field

Let \( \phi(x) \) denote a complex scalar field, with the Lagrange density

\[ \mathcal{L}[\phi, \phi^*, \partial_\mu \phi, \partial_\mu \phi^*] = -\left( \partial_\mu \phi \right) \left( \partial^\mu \phi^* \right) - m^2 \phi \phi^*, \]

where \( \phi \) and \( \phi^* \) are to be considered as independent fields.

The Euler–Lagrange equations

\[ \frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \phi^*} = \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)^*} \]

are easily computed and lead to respective Klein–Gordon equations

\[ \partial_\mu \partial^\mu \phi^*(x) - m^2 \phi^*(x) = 0 \quad \text{and} \quad \partial_\mu \partial^\mu \phi(x) - m^2 \phi(x) = 0, \]

(IX.20)

which are naturally equivalent if \( m \in \mathbb{R} \).

One sees at once that the Lagrange density (IX.19) is invariant under the "global U(1) transformation"

\[
\begin{align*}
\phi(x) &\rightarrow \phi'(x) = e^{-i \Lambda \phi(x)} \\
\phi^*(x) &\rightarrow \phi'^*(x) = e^{i \Lambda \phi(x)}
\end{align*}
\]

(IX.21)

where \( \Lambda \in \mathbb{R} \) is invariant of the space-time position. In the case of an infinitesimally small \( \Lambda \), these transformations become

\[ \phi'(x) \simeq \phi(x) - i \Lambda \phi(x) \quad \text{and} \quad \phi'^*(x) \simeq \phi^*(x) + i \Lambda \phi^*(x), \]

corresponding to infinitesimal variations \( \delta \phi = -i \Lambda \phi \), \( \delta \phi^* = i \Lambda \phi^* \) of the independent fields. Inserting the latter in the expression (IX.10) of the Noether current yields

\[ J^\mu = i \left( \phi^* \partial^\mu \phi - \phi \partial^\mu \phi^* \right), \]

(IX.22)

where we have divided by \( \Lambda \). One easily checks with the equations of motion (IX.20) that this current, whose spatial components are reminiscent of the probability current of non-relativistic quantum mechanics, obeys the local conservation equation \( \partial_\mu J^\mu = 0 \).

Remarks:

\* One can add to the Lagrange density (IX.19) a "potential term" \(-V(|\phi|^2)\) depending only on the squared modulus \( |\phi|^2 \) of the field. This modifies the equations of motion, yet the theory remains invariant under the global U(1) symmetry (IX.21), and the corresponding Noether current is still given by Eq. (IX.22).

\* The reader is encouraged to go through the above example by replacing \( \phi \) by \( \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \), where \( \phi_1 \) and \( \phi_2 \) are now real scalar fields. The symmetry transformations (IX.21) are then replaced by SO(2) transformations, i.e. rotations in the \( (\phi_1, \phi_2) \)-plane — which are of course totally equivalent.
IX.3 Gauge field theories


APPENDIX A

Elements of topology

A.1 Topological space

open / closed set; continuity; (path-)connectedness; compacity; homotopy group(?)

homeomorphism: bijective map such that $f$ and $f^{-1}$ are continuous
diffeomorphism: homeomorphism such that $f$ and $f^{-1}$ are infinitely differentiable
analytic diffeomorphism:

A.2 Manifolds

Charts, atlas

Definition A.1. A manifold $\mathcal{M}$ is a Hausdorff topological space with an atlas such that the open sets $U_i$ provide a covering of $\mathcal{M}$ and that the mappings $\varphi_j \circ \varphi_i^{-1}$ are infinitely differentiable.

The dimension $n$ of the manifold is the dimension of the vector space into which the charts $\varphi_i$ map.


