SSHEP 2018: "Symmetries in HEP" — Version 1.0

Vasja Susič, University of Basel

These lecture notes are based on an introductory 4-hour course on "Symmetries in High Energy Physics" that was given at the *Sarajevo School of High Energy Physics 2018* in October 2018. It is intended for graduate and advanced undergraduate students not yet familiar with this topic.

The course focuses on the underlying mathematics of "symmetries", i.e. group theory, but it also briefly touches on some applications in High Energy Physics (HEP). The prerequisites have been kept to a minimum, requiring in principle only knowledge of linear algebra and calculus from mathematical topics, while some basic familiarity with Quantum Mechanics and (quantum) Field Theory is recommended for understanding the applications in HEP. The material covered in the lectures was limited by time constraints; concepts were introduced to the degree that they were needed, sometimes simply by example, and no exact proofs are given, but results typically are at least motivated beforehand.

The writing style of these notes is quite verbose; it is hoped that this makes them more pedagogical and improves conceptual understanding. The Einstein summation convention applies throughout these notes. Colored text was used for improved readability, with blue signifying definitions, and red signifying statements/results/theorems.

If typos are found, or the reader has questions regarding the material in these notes, they can contact me at vasja.susic@unibas.ch. A reader further interested in group theory in the context of physics can for example check references [1,2], but is also invited to have a look at this link, where a cloud service hosts all materials related to a one semester course on this topic that I gave at the University of Basel in the Fall Semester 2017.

[1] H. Georgi, "Lie Algebras In Particle Physics. From Isospin To Unified Theories," Front. Phys. **54** (1982) 1.

[2] A. Zee, "Group Theory in a Nutshell for Physicists," Princeton University Press (2016). ISBN 0691162697.

Contents

1 Basics of group theory

1.1 Basic notions in group theory

A group is a set of elements G equipped with a binary operation $G \times G \to G$ (it prescribes to any 2 elements $x, y \in G$ a product xy); the operation has to satisfy the following properties:

1. *Associativity*: $(xy)z = x(yz)$ for all $x, y, z \in G$.
2. *Existence of identity element*: $\exists e \in G$, such that $ex = xe = x$ for all $x \in G$.

   Notation: we shall label the identity element e by 1.
3. *Existence of inverse elements*: $\forall x \; \exists y$: $xy = yx = 1$.

   Notation: we label the inverse element of x by x^{-1}.

Comment: in physics, one can imagine an element of a group as a "transformation". In this context, "symmetries" of a system are non-trivial transformations which leave the system unchanged.

Notation: $\mathbb{F}$ is a generic label for either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. We label the set of all matrices of size $m \times n$ with entries in $\mathbb{F}$ by $\mathbb{F}^{m,n}$.

Example: the integers $\mathbb{Z}$ with addition form a group. The operation between $n, m \in \mathbb{Z}$ is $m + n$, the identity element is $0 \in \mathbb{Z}$ and the inverse of n is $-n$.

We list a number of basic concepts associated to groups below:

- A subgroup H of G is a subset $H \subseteq G$, which closes under the operation ($xy \in H$ for any $x, y \in H$) and is itself a group (H must contain 1 and inverses for all its elements).

  Example: even integers $2\mathbb{Z}$ are a subgroup of the integers $\mathbb{Z}$.
- A group homomorphism between two groups G_1 and G_2 is a map $\phi : G_1 \to G_2$, which preserves the multiplication rule: for all $x, y \in G_1$ we have

  $$\phi(x)\phi(y) = \phi(xy). \tag{1}$$

  A group isomorphism is an invertible group homomorphisms.

  Comment: if two groups are isomorphic, they are "equivalent" from the point of view of group theory.
- A group is called Abelian if its operation is commutative: $xy = yx$ for all $x, y \in G$.
- A group is finite if it has a finite number of elements.
- If a group G has elements, which can be parametrized by n parameters, it is called a Lie group. We will consider only cases where the parameters are real, i.e. with real Lie groups. The number of parameters n is called the dimension of the Lie group.

  Notation: unless otherwise specified, we shall label a group element by $\mathbf{A}(\vec{\alpha})$, where $\vec{\alpha} = (\alpha^1, \alpha^2, \ldots, \alpha^n)$ are the parameters. We use the convention where the parameters α^a have an upper index a. Different choices of $\vec{\alpha}$ pick different elements in the group.

Example: rotations in $\mathbb{R}^2$ can be written as 2×2 matrices. A rotation by φ is represented by the matrix

$$\mathbf{O}(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{2}$$

The elements are parametrized by a real parameter φ ("the angle") in the 0 to 2π range. We can extend to definition for φ to the entire real line $\mathbb{R}$ due to the periodicity of the functions sin and cos, but the elements with an angle which differs by $2\pi n$ for $n \in \mathbb{Z}$ are the same: $\mathbf{O}(\varphi) = \mathbf{O}(\varphi + 2\pi n)$. Multiplication: $\mathbf{O}(\varphi_1)\mathbf{O}(\varphi_2) = \mathbf{O}(\varphi_3)$, where the multiplication rule in this case is simply $\varphi_3(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2$. Since this Lie group has 1 real parameter, its dimension is 1.

Comment: in a general Lie group, the multiplication rule between elements would satisfy $\mathbf{A}(\vec{\alpha}_1)\mathbf{A}(\vec{\alpha}_2) = \mathbf{A}(\vec{\alpha}_3)$, where $\vec{\alpha}_3 = \vec{\alpha}_3(\vec{\alpha}_1, \vec{\alpha}_2)$ is in general a very complicated function.

Comment: for our purposes in this lecture, we can imagine a Lie group to always be a group of matrices of some sort of rotations, i.e. Lie group elements $\mathbf{A}(\vec{\alpha})$ are square matrices.

- If a group H is a subgroup of a Lie group, but the set of parameter values would correspond only to a discrete set, the group is called a discrete group.

  Example: finite groups are discrete. $\mathbb{Z}$ with addition is also discrete.

We list below examples of groups relevant for this course:

1. Integers $\mathbb{Z}$ with addition $+$; $-x$ is the inverse of x and 0 is the identity element.

   Properties: Abelian, discrete.

2. Cyclic group $\mathbb{Z}_n$ (also labeled C_n): every $n \in \mathbb{N}$ defines its own cyclic group. The set of $\mathbb{Z}_n$ is $\{0, 1, 2, \ldots n-1\}$, and the operation between two elements x and y is

   $$x \dot{+} y := x + y \pmod{n}, \tag{3}$$

   where $\bmod n$ gives the remainder when dividing by n; such a remainder is indeed an integer between 0 and $n-1$.

   Example: in $\mathbb{Z}_7$ we have $5\dot{+}4 = 2$, since $5 + 4 = 9$ and $9 \bmod 7 = 2$.

   Properties: Abelian, discrete.

3. Some matrix groups (operation is always matrix multiplication; identity element is the unit matrix $\mathbb{1}$): for each $n \in \mathbb{N}$ we define

name	label	definition
general linear group	$\mathrm{GL}(n, \mathbb{F})$	$\{\mathbf{A} \in \mathbb{F}^{n,n};\ \det \mathbf{A} \neq 0\}$
special linear group	$\mathrm{SL}(n, \mathbb{F})$	$\{\mathbf{A} \in \mathrm{GL}(n, \mathbb{F});\ \det \mathbf{A} = 1\}$
orthogonal group	$\mathrm{O}(n)$	$\{\mathbf{O} \in \mathrm{GL}(n, \mathbb{R});\ \mathbf{O}\mathbf{O}^T = \mathbb{1}\}$
special orthogonal group	$\mathrm{SO}(n)$	$\{\mathbf{O} \in \mathrm{GL}(n, \mathbb{R});\ \mathbf{O}\mathbf{O}^T = \mathbb{1},\ \det \mathbf{O} = 1\}$
unitary group	$\mathrm{U}(n)$	$\{\mathbf{U} \in \mathrm{GL}(n, \mathbb{C});\ \mathbf{U}\mathbf{U}^\dagger = \mathbb{1}\}$
special unitary group	$\mathrm{SU}(n)$	$\{\mathbf{U} \in \mathrm{GL}(n, \mathbb{C});\ \mathbf{U}\mathbf{U}^\dagger = \mathbb{1},\ \det \mathbf{U} = 1\}$

Properties: Lie groups and non-Abelian, exceptions for $n = 1$ or SO(2).

Comment: $\mathrm{GL}(n, \mathbb{F})$ consist of all $n \times n$ invertible matrices with entries in $\mathbb{F}$. $\mathrm{SO}(n)$ consists of all (proper) rotations in n-dimensional real space $\mathbb{R}^n$, and $\mathrm{SU}(n)$ consists of all (proper) complex rotations in n-dimensional complex space $\mathbb{C}^n$.

1.2 Lie groups, Lie algebras and the Exponential Map

Suppose G is a Lie group of dimension n; furthermore, suppose the elements are parametrized by $\mathbf{A}(\vec{\alpha})$, where $\vec{\alpha}$ represents the vector of parameters. We label the components of $\vec{\alpha}$ by α^a, where a goes from 1 to n.

In line with our (limited) definition of Lie groups, we can imagine $\mathbf{A}(\vec{\alpha})$ as a matrix. We consider parameterizations that include the identity element in their domain, i.e. $\mathbf{A}(\vec{\alpha}) = \mathbb{1}$ for some $\vec{\alpha}$; we can then assume without loss of generality that $\mathbf{A}(\vec{0}) = \mathbb{1}$ (otherwise we reparametrize by a translation in parameter space).

We define the infinitesimal generator corresponding to the parameter α^a to be

$$\mathbf{T}_a := \frac{1}{i} \frac{\partial \mathbf{A}(\vec{\alpha})}{\partial \alpha^a} \bigg|_{\vec{\alpha}=0}, \tag{4}$$

which corresponds to computing the transformation "infinitesimally close to the identity" in the direction represented by the parameter α^a; the factor $1/i$ is there as part of convention (we shall see the reason later). There are as many generators as there are parameters, and each $\mathbf{T}_a$ is again a matrix. A different parametrization $\mathbf{A}$ would in general give different generators, which would however be simply linear combinations of the old ones.

Example: the generator corresponding to φ in Eq. (2) is

$$\mathbf{T} = -i \frac{\partial}{\partial \varphi} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \bigg|_{\varphi=0} = -i \begin{pmatrix} -\sin\varphi & -\cos\varphi \\ \cos\varphi & -\sin\varphi \end{pmatrix} \bigg|_{\varphi=0} = -i \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{5}$$

There are always n generators in G, the same number as the number of parameters; they are linearly independent because the parameters are independent. Making use of the generators, it is possible to define another parametrization $\tilde{\mathbf{A}}(\vec{\alpha})$ of group elements by the exponential map:

$$\tilde{\mathbf{A}}(\vec{\alpha}) := e^{i\alpha^a \mathbf{T}_a}, \tag{6}$$

where $\alpha^a \mathbf{T}_a$ is a linear combination of the generator matrices, and the exponential e for a matrix is simply defined by its power series

$$e^{\mathbf{X}} = \sum_{k=1}^{\infty} \tfrac{1}{k!} \mathbf{X}^k = \mathbb{1} + \mathbf{X} + \tfrac{1}{2!}\mathbf{X}^2 + \cdots. \tag{7}$$

This series always converges. The product of two exponentials can be written as another exponential due to the Baker-Campbell-Hausdorff (BCH) formula:

$$e^{\mathbf{X}} e^{\mathbf{Y}} = e^{\mathbf{X}+\mathbf{Y}+\frac{1}{2}[\mathbf{X},\mathbf{Y}]+\cdots}, \tag{8}$$

where the commutator $[.,.]$ is defined by $[\mathbf{X}, \mathbf{Y}] := \mathbf{XY} - \mathbf{YX}$, and the dots in the equation represent higher order terms in $\mathbf{X}$ and $\mathbf{Y}$, which can be written using exclusively nested

commutators (such as $[[\mathbf{X},\mathbf{Y}],\mathbf{X}]$), which are higher than second order in $\mathbf{X}$ and $\mathbf{Y}$. Crucially, usual matrix multiplication is not required, only the commutator operation is. The BCH formula thus indicates that it is the commutator operation that is crucial at the level of generators $\mathbf{T}_a$, since it completely determines the product rule between the elements of the group.

From now on, we shall always view the elements of a Lie group in an exponential parameterization, i.e. as exponentials $e^{i\alpha^a \mathbf{T}_a}$. We call the vector space formed by all possible linear combinations $i\alpha^a \mathbf{T}_a$ the Lie algebra of the Lie group G, and we label it by $\mathfrak{g}$. The Lie algebra is a vector space equipped with the commutator, which maps two elements (matrices) of the Lie algebra into one, as defined earlier. The commutator satisfies the following properties:

1. Bilinearity: $[\alpha\mathbf{X}+\beta\mathbf{Y},\mathbf{Z}] = \alpha[\mathbf{X},\mathbf{Z}] + \beta[\mathbf{Y},\mathbf{Z}]$ and analogous in the second factor.
2. Antisymmetry: $[\mathbf{X},\mathbf{Y}] = -[\mathbf{Y},\mathbf{X}]$.
3. Jacobi identity (cyclic permutations): $[[\mathbf{X},\mathbf{Y}],\mathbf{Z}] + [[\mathbf{Y},\mathbf{Z}],\mathbf{X}] + [[\mathbf{Z},\mathbf{X}],\mathbf{Y}] = 0$.

These properties are trivial to check for the usual definition of the commutator operation $[\mathbf{X},\mathbf{Y}] := \mathbf{XY} - \mathbf{YX}$.

We now give some further concepts relevant for Lie algebras:

- Due to bilinearity, the commutator is uniquely defined if we know its value between the generators. The result can then be expanded as a linear combination of generators (Lie algebra closes under the commutator, so we simply expand the result in the basis of generators). We define the structure constants f_{abc} as the (real) coefficients in the equation

$$[\mathbf{T}_a,\mathbf{T}_b] = if_{abc}\mathbf{T}_c, \tag{9}$$

  where the right side is summed over c. Due to the antisymmetry of the commutator, f_{abc} is antisymmetric in a and b, and it can be made completely antisymmetric for a suitable choice of generators $\mathbf{T}_a$. Specifying f_{abc} is thus sufficient to define the Lie algebra.

- A subalgebra is a vector subspace, which closes under the commutator: in a subalgebra $\mathfrak{h}$ of the Lie algebra $\mathfrak{g}$ we have the condition that for all $\mathbf{X},\mathbf{Y}\in\mathfrak{h}$ also $[\mathbf{X},\mathbf{Y}]\in\mathfrak{h}$.

- The Cartan subalgebra is a maximal commuting subalgebra. "Commuting" means that all its elements commute, i.e. $[\mathbf{X},\mathbf{Y}] = 0$ for all Cartan subalgebra elements $\mathbf{X}$ and $\mathbf{Y}$, and it is a maximal algebra with this property (there is no bigger commuting subalgebra). Its importance lies in the fact that mutually commuting matrices can be simultaneously diagonalized; using a suitable basis of rows and columns, all elements of the Cartan subalgebra are thus diagonal matrices. We shall call the basis of the Cartan algebra the Cartan generators.

  Comment: the choice of a Cartan subalgebra is not unique, but all choices are equivalent from the point of view of group theory. The choice will be dictated by convention.

- The dimension of a Lie algebra is its dimension as a vector space (number of generators). The rank of a Lie algebra is defined as the dimension of its Cartan subalgebra. We shall generically label the rank by k.

- The Lie algebras of the matrix groups we defined are labeled by *fraktur* letters, e.g. the Lie algebra of $\mathrm{GL}(n,\mathbb{F})$ is labeled by $\mathfrak{gl}(n,\mathbb{F})$. We observe that for any square matrix $\mathbf{X}$, the exponential $e^{\mathbf{X}}$ is invertible, namely $(e^{\mathbf{X}})^{-1} = e^{-\mathbf{X}}$ (the reader is invited to show this by using the BCH formula). Writing $\mathbf{A} = e^{i\mathbf{X}}$, invertibility of $\mathbf{A}$ gives no restrictions on $\mathbf{X}$, and thus the Lie algebra $\mathfrak{gl}(n,\mathbb{F})$ consists of all $n \times n$ matrices: $\mathfrak{gl}(n,\mathbb{F}) = \mathbb{F}^{n,n}$.

  Conditions on $\mathbf{A}$ in the other groups (special,orthogonal,unitary) are translated to conditions on the Lie algebra elements $\mathbf{X}$ in the following way:

$$\mathbf{A}\mathbf{A}^\dagger = \mathbb{1} \quad \Longrightarrow \quad \mathbf{X} = \mathbf{X}^\dagger, \tag{10}$$
$$\mathbf{A}\mathbf{A}^T = \mathbb{1} \quad \Longrightarrow \quad \mathbf{X} = -\mathbf{X}^T, \tag{11}$$
$$\det \mathbf{A} = 1 \quad \Longrightarrow \quad \operatorname{Tr} \mathbf{X} = 0. \tag{12}$$

  The first condition can be demonstrated by expanding $\mathbf{A}$ to linear order in $\mathbf{X}$:

$$\begin{aligned} \mathbb{1} = \mathbf{A}\mathbf{A}^\dagger = (e^{i\mathbf{X}})(e^{i\mathbf{X}})^\dagger = e^{i\mathbf{X}}e^{-i\mathbf{X}^\dagger} = (\mathbb{1} + i\mathbf{X} + \cdots)(\mathbb{1} - i\mathbf{X}^\dagger + \cdots) = \\ = \mathbb{1} + i(\mathbf{X} - \mathbf{X}^\dagger) + \cdots, \end{aligned} \tag{13}$$

  and thus $\mathbf{X} = \mathbf{X}^\dagger$. This was the reason for the conventional factor $1/i$ in Eq. (4), since it implies that the generators of a unitary matrix are Hermitian (more convenient) rather than anti-Hermitian. An analogous computation gives the second condition. The third condition is the result of the determinant-trace formula, which holds for any square matrix $\mathbf{X}$:

$$\det(e^{\mathbf{X}}) = e^{\operatorname{Tr}\mathbf{X}}. \tag{14}$$

  Summarizing the considerations above, the definitions of the Lie algebras corresponding to the Lie groups we defined earlier are the following:

$$\mathfrak{gl}(n,\mathbb{F}) := \mathbb{F}^{n,n}, \tag{15}$$
$$\mathfrak{sl}(n,\mathbb{F}) := \{i\mathbf{X} \in \mathbb{F}^{n,n};\ \operatorname{Tr}\mathbf{X} = 0\}, \tag{16}$$
$$\mathfrak{o}(n) = \mathfrak{so}(n) := \{i\mathbf{X} \in \mathbb{R}^{n,n};\ \mathbf{X}^T = -\mathbf{X}\}, \tag{17}$$
$$\mathfrak{u}(n) := \{i\mathbf{X} \in \mathbb{C}^{n,n};\ \mathbf{X}^\dagger = \mathbf{X}\}, \tag{18}$$
$$\mathfrak{su}(n) := \{i\mathbf{X} \in \mathbb{C}^{n,n};\ \mathbf{X}^\dagger = \mathbf{X},\ \operatorname{Tr}\mathbf{X} = 0\}, \tag{19}$$
$$\tag{20}$$

  Observe that the orthogonal and special orthogonal Lie algebras are the same, since antisymmetry implies a zero trace of $\mathbf{X}$.

1.3 Representations

A very important concept for physics is that of a representation of a group (and also a representation of a Lie algebra). We again list the relevant concepts below:

- A representation of a group G is a homomorphism $\Phi : G \to \mathrm{GL}(m,\mathbb{F})$ for some $m \in \mathbb{N}$. It maps group elements into invertible $m \times m$ matrices; since it is a homomorphism, it preserves the multiplication rule for any two elements $x, y \in G$:

$$\Phi(x)\Phi(y) = \Phi(xy). \tag{21}$$

A representation is thus a "realization" of a particular symmetry group as matrices, which in turn can be understood as linear transformations on an underlying vector space (columns).

Comment: in physics jargon we typically refer to the elements $\Phi(x)$ as a representation of a group, and not to the map Φ itself. The underlying vector space that these matrices act on is called the space of states.

Notation: we shall sometimes make use of shorthand notation where we denote an element $\Phi(x)$ in a representation Φ simply by x (the representation used is inferred from the context, or preferably stated explicitly). For "abstract" group elements, before mapping them via a representation map, we use a hat on top of the element: $\hat{x}$. Any equation between abstract hatted objects is valid in any representation (without the hats), since representations preserve the multiplication rule. Also, when referring to representations, we shall sometimes label them with a subscripted letter R or by their size m rather than the map Φ.

Example: every group has the trivial representation $\Phi(x) = 1$. The multiplication law of Eq. (21) is trivially satisfied.

- The number m (the size of the matrices $\Phi(x)$) above is called the dimension of a representation.

  Example: for the group $\mathbb{Z}_n$ we define $\Phi(k) = e^{i\,2\pi k/n}$, where $k \in \{0, 1, \ldots, n-1\}$. This is a representation of dimension 1 (1×1 complex matrices are complex numbers). The reader is invited to check that Φ indeed satisfies the $\mathbb{Z}_n$ multiplication law, i.e. $\Phi(k)\Phi(l) = \Phi((k+l) \bmod n)$.

- A representation of a Lie algebra $\mathfrak{g}$ of dimension m is a linear map $\tilde{\Phi} : \mathfrak{g} \to \mathfrak{gl}(m, \mathbb{F})$ (into matrices $\mathbb{F}^{m,m}$) matrices, such that the commutator is preserved:

$$\tilde{\Phi}([\mathbf{X}, \mathbf{Y}]) = [\tilde{\Phi}(\mathbf{X}), \tilde{\Phi}(\mathbf{Y})]. \tag{22}$$

  Comment: a Lie algebra representation $\tilde{\Phi}$ induces a Lie group representation Φ, since we get the group elements by exponentiation: $\Phi(e^{i\mathbf{X}}) = e^{i\,\tilde{\Phi}(\mathbf{X})}$.

  Notation: similar to groups, we use a hat on top of a symbol to denote an "abstract element" of the algebra, i.e. by $\hat{\mathbf{X}}$, while $\tilde{\Phi}(\mathbf{X})$ is then denoted simply as $\mathbf{X}$. In this notation, all relations/equations which hold at the abstract level (with hats) hold also in any specific representation (without hats).

- A unitary representation is a representation R with map Φ, for which all group elements $x \in G$ are represented by unitary matrices $\Phi(x)$: $\Phi(x)\Phi(x)^\dagger = \mathbb{1}$. At the infinitesimal level, this happens when all generators in that representation are Hermitian matrices: $\mathbf{T}_a = \mathbf{T}_a^\dagger$.

- We define the following operations between representations:

  1. If R_1 and R_2 are two group representations of G, defined by the maps Φ_1 and Φ_2, respectively, we define the direct sum representation $R_1 \oplus R_2$ by the map

$$\Phi(x) = \begin{pmatrix} \Phi_1(x) & \\ & \Phi_2(x) \end{pmatrix}. \tag{23}$$

Comment: in the direct sum, the element x is represented by a matrix where the two representations are combined (the space of states is the direct sum of vector spaces $V_1 \oplus V_2$), with the representations Φ_1 and Φ_2 coming in as diagonal blocks.

2. If R is a representation, we label its conjugate representation by R^*, and it consists of complex-conjugated matrices; if $\Phi(x)$ belongs to R, then $\Phi(x)^*$ belongs to R^*. At the level of the Lie algebra: an element $\mathbf{U}(\vec{\alpha}) = e^{i\alpha^a \mathbf{T}_a}$ in R is written in R^* as

$$\mathbf{U}(\vec{\alpha})^* = (e^{i\alpha^a \mathbf{T}_a})^* = e^{-i\alpha^a \mathbf{T}_a^*} = e^{i\alpha^a (-\mathbf{T}_a^*)}. \tag{24}$$

The last expression shows that if the generators of R are the matrices $\mathbf{T}_a$, then the generators of R^* are $-\mathbf{T}_a^*$, but using the same real parameters α^a for the group element as in the original representation R.

3. The product representation $R_1 \otimes R_2$ is defined in the following way: if the vector space of states for R_i is labeled by V_i, then the space of states of $R_1 \otimes R_2$ is the tensor product vector space $V_1 \otimes V_2$. Suppose we have a state in the product space, which we can write as $v_1 \otimes v_2$, $v_1 \in V_1$ and $v_2 \in V_2$. Then the group element x acts on this state as $(\Phi_1(x)v_1) \otimes (\Phi_2(x)v_2)$.

Comment: when writing the group element $\Phi(x)$ in $R_1 \otimes R_2$ as a matrix, it turns out to be the Kronecker product of two matrices: $\Phi(x) = \Phi_1(x) \otimes \Phi_2(x)$. We shall later present one other technique, where the states of the product representation can be written as components of tensors (multi-dimensional arrays).

- Two representations R_1 and R_2 with maps Φ_1 and Φ_2 are equivalent, if there exists an invertible matrix $\mathbf{P}$, such that $\Phi_1(x) = \mathbf{P}\Phi_2(x)\mathbf{P}^{-1}$ for all $x \in G$. This means that matrices in equivalent representations differ only in the choice of basis for the space of states.

- A representation R is irreducible if it (or a representation equivalent to it) cannot be written as a direct sum of two smaller representations, $R_1 \oplus R_2$. A representation is reducible if it can be written as such a direct sum.

  Comment: a representation is thus reducible if it is possible to simultaneously block diagonalize $\Phi(x)$ for all group elements $x \in G$ into a form $\begin{pmatrix} \Phi_1(x) & \\ & \Phi_2(x) \end{pmatrix}$. If that happens, the space of states splits, and group transformations mix states in the first representation R_1 independently of the states in R_2. Irreducible representations are thus the "building blocks" for all representations, so it is sufficient to study irreducible representations.

- In Lie groups, we have 2 important representations, which are always available (and can be used to construct new representations with the 3 operations on representations mentioned earlier):

  1. The defining representation (also called the fundamental representation) is the one by which we defined the Lie groups in these notes. The defining representation of SU(n) is n-dimensional (unitary $n \times n$ matrices with determinant 1).
  2. The adjoint representation is a representation, which uses the Lie algebra itself as a space of states, so it is of the same dimension as the Lie algebra. The generators of the adjoint representation act as commutators on the basis states,

i.e. $\hat{\mathbf{T}}_a\,\mathbf{T}_b := [\mathbf{T}_a, \mathbf{T}_b]$. The generator $\mathbf{T}_a$ in the adjoint representation is given by the antisymmetric structure constants:

$$(\mathbf{T}_a)_{bc} = -i\, f_{abc}. \tag{25}$$

The reader is invited to prove that generators defined in this way indeed satisfy the commutation relations with the help of the Jacobi identity when plugging-in $\mathbf{X} = \mathbf{T}_a$, $\mathbf{Y} = \mathbf{T}_b$ and $\mathbf{Z} = \mathbf{T}_c$.

2 More on Lie groups through examples

2.1 The group U(1)

The group U(1) by definition consists of 1×1 unitary complex matrices. Matrices of size 1 are just numbers, and unitarity for a complex number a implies $aa^* = |a|^2 = 1$. This means that the elements of the group U(1) are elements of the unit circle in the complex plane:

$$\mathrm{U}(1) := \{e^{i\varphi}; \varphi \in [0, 2\pi)\}. \tag{26}$$

The group has the topology of the circle, since an angle of 0 is equivalent to 2π. We now list some properties of this group:

- The group U(1) is isomorphic to the group SO(2), as can be seen by mapping the group elements in the following way:

$$e^{i\varphi} \mapsto \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{27}$$

  This implies the following mapping between the underlying vector spaces of states, on which the above transformations act: $x + iy \mapsto (x, y)^T$.

- The group is Abelian (since the multiplication of complex numbers is commutative), and it has dimension 1 since it has only one parameter φ. Using the exponential map, an "abstract" group element $\hat{\mathbf{U}}(\varphi)$ corresponding to the particular value of the parameter φ, can be written as

$$\hat{\mathbf{U}}(\varphi) = e^{i\varphi\hat{\mathbf{Q}}}, \tag{28}$$

  where $\hat{\mathbf{Q}}$ is called the charge operator. For the defining representation used above, we simply have $\mathbf{Q} = 1$.

- Classification of irreducible representations of U(1): it is sufficient for the classification to list the possible values that the charge operator $\mathbf{Q}$ can take. Since U(1) is an Abelian group, all irreducible representation are of dimension 1 (we provide no proof for that here). Since $\mathbf{U}(0) = \mathbf{U}(2\pi)$ for elements in every representation, $\mathbf{Q} = n \in \mathbb{Z}$: charges are integer. The trivial representation is the one with $\mathbf{Q} = 0$.

- Cyclic group are subgroups of U(1). One can obtain the elements of $\mathbb{Z}_n$ by taking only discrete values for the angles: $\varphi = 2\pi k/n$, where $k = 0, \ldots, n-1$. The representations of U(1) thus also become representations of $\mathbb{Z}_n$ when considering only elements corresponding to discrete angle values.

- Invariants: we are interested in what kind of terms in the Lagrangian stay invariant under U(1) transformations (equivalent to saying that they transform under the trivial/singlet representation). If $\phi(x)$ is a complex field with charge n under the U(1) group, then a transformation with the group element $\hat{\mathbf{U}}(\varphi)$ causes the change $\phi(x) \mapsto \phi'(x) = e^{in\varphi}\phi(x)$. Suppose now we have a product term in the Lagrangian; for concreteness, we choose for example 3 fields: $\phi_1\phi_2\phi_3$, where ϕ_i transforms with the representation of charge n_i under U(1). Then a U(1) rotation transform this term into

$$\phi_1\phi_2\phi_3 \mapsto \phi_1'\phi_2'\phi_3' = (e^{in_1\varphi}\phi_1)(e^{in_2\varphi}\phi_2)(e^{in_3\varphi}\phi_3) = e^{i(n_1+n_2+n_3)\varphi}\phi_1\phi_2\phi_3. \tag{29}$$

  We see that the term is invariant for all φ if and only if $n_1 + n_2 + n_3 = 0$. This holds true in general: terms are invariant under U(1) if the sum of the charges of the fields is zero.

2.2 The SU(2) group

The group SU(2) is defined by 2×2 unitary matrices $\mathbf{U}$, which satisfy $\det \mathbf{U} = 1$. We write group elements as exponents of the $\mathfrak{su}(2)$ Lie algebra elements $\mathbf{X}$ by $\mathbf{U} = e^{i\mathbf{X}}$. Unitarity and unit determinant of $\mathbf{U}$ then translate to the conditions $\mathbf{X} = \mathbf{X}^\dagger$ and $\operatorname{Tr} \mathbf{X} = 0$ for $\mathbf{X} \in \mathbb{C}^{2,2}$. The basis of all 2×2 complex matrices is

$$\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right),\ \left(\begin{smallmatrix}i&0\\0&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&i\\0&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&0\\i&0\end{smallmatrix}\right),\ \left(\begin{smallmatrix}0&0\\0&i\end{smallmatrix}\right). \tag{30}$$

Hermiticity implies that diagonal elements must be real, while the upper triangular part determines the lower triangular part. Tracelessness further imposes one condition. Taken together, they imply that the basis for the $\mathfrak{su}(2)$ Lie algebra can be chosen to consist of the Pauli matrices:

$$\boldsymbol{\sigma}_1 := \begin{pmatrix}0&1\\1&0\end{pmatrix}, \qquad \boldsymbol{\sigma}_2 := \begin{pmatrix}0&-i\\i&0\end{pmatrix}, \qquad \boldsymbol{\sigma}_3 := \begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{31}$$

By convention we then define the SU(2) generators as

$$\mathbf{T}_a := \boldsymbol{\sigma}_a/2, \tag{32}$$

where $a = 1, 2, 3$. These generators then have the normalization

$$\operatorname{Tr}(\mathbf{T}_a\mathbf{T}_b) = \tfrac{1}{2}\delta_{ab}. \tag{33}$$

The generators satisfy the commutation relations

$$[\mathbf{T}_a, \mathbf{T}_b] = i\,\varepsilon_{abc}\,\mathbf{T}_c, \tag{34}$$

so the structure constants are simply the Levi-Civita tensor ε_{abc}, the unique completely antisymmetric tensor defined by $\varepsilon_{123} = 1$.

The above considerations imply the following properties of the SU(2) group:

- The dimension of SU(2) is 3, since there are 3 generators.
- The defining representation has dimension 2 (Pauli matrices are 2×2 matrices). The adjoint representation has dimensions 3; the generators in the adjoint representation are $(\mathbf{T}_a)_{bc} = -if_{abc} = -i\varepsilon_{abc}$, or explicitly

$$\mathbf{T}_1 = (-i)\begin{pmatrix}0&&\\&&1\\&-1&\end{pmatrix}, \quad \mathbf{T}_2 = (-i)\begin{pmatrix}&&-1\\&0&\\1&&\end{pmatrix}, \quad \mathbf{T}_3 = (-i)\begin{pmatrix}&1&\\-1&&\\&&0\end{pmatrix}. \tag{35}$$

- A general element of the group (at least if it is sufficiently close to the identity matrix) can be written as

$$\mathbf{U}(\vec{\alpha}) = e^{i\alpha^a \mathbf{T}_a}, \tag{36}$$

where $\vec{\alpha}$ is the parameter vector consisting of components α^a, where $a = 1, 2, 3$. Due to the properties of the Pauli matrices, it is possible to evaluate such an exponential analytically. We decompose the parameter vector into its size $\alpha := |\vec{\alpha}|$ and direction $\vec{n} := \vec{\alpha}/|\vec{\alpha}|$, such that $\vec{\alpha} = \alpha\,\vec{n}$. We observe that for a vector $\vec{n} = (n_x, n_y, n_z)^T$ of unit size, the following identity holds:

$$(\vec{n}\cdot\vec{\boldsymbol{\sigma}})^2 = \mathbb{1}, \qquad \text{where} \quad \vec{n}\cdot\vec{\boldsymbol{\sigma}} := n^a \boldsymbol{\sigma}_a = \begin{pmatrix} n_z & n_x - i n_y \\ n_x + i n_y & -n_z \end{pmatrix}. \tag{37}$$

Expanding the exponential in Eq. (36) and using the identity in Eq. (37), it is possible to derive (the reader is invited to show this) that

$$\begin{aligned} \mathbf{U}(\vec{\alpha}) = e^{i\alpha/2\,\vec{n}\cdot\vec{\boldsymbol{\sigma}}} &= \cos(\alpha/2)\,\mathbb{1} + i\sin(\alpha/2)\,\vec{n}\cdot\vec{\boldsymbol{\sigma}} \\ &= \begin{pmatrix} \cos(\alpha/2) + i n_z \sin(\alpha/2) & (i n_x + n_y)\sin(\alpha/2) \\ (i n_x - n_y)\sin(\alpha/2) & \cos(\alpha/2) - i n_z \sin(\alpha/2) \end{pmatrix}. \end{aligned} \tag{38}$$

The group SU(2) is simple enough that we were able to analytically evaluate the exponential; this is typically not possible in a general Lie group.

- The Cartan subalgebra of $\mathfrak{su}(2)$ is of dimension 1. By convention, we choose it to be generated by $\hat{\mathbf{T}}_3$. The other two generators $\hat{\mathbf{T}}_1$ and $\hat{\mathbf{T}}_2$ (or any linear combination thereof) indeed do not commute with $\hat{\mathbf{T}}_3$.

- We use the non-Cartan generators $\hat{\mathbf{T}}_1$ and $\hat{\mathbf{T}}_2$ to define raising and lowering operators:

$$\hat{\mathbf{T}}_\pm := \hat{\mathbf{T}}_1 \pm i\hat{\mathbf{T}}_2. \tag{39}$$

This definition is at the abstract level, so it applies to all representations; in the defining representation this amounts to

$$\mathbf{T}_+ = \mathbf{T}_1 + i\mathbf{T}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{T}_- = \mathbf{T}_1 - i\mathbf{T}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{40}$$

Using the new definitions, the commutation relations can be rewritten as (the reader is invited to confirm them, e.g. using the defining representation):

$$[\hat{\mathbf{T}}_3, \hat{\mathbf{T}}_\pm] = \pm\hat{\mathbf{T}}_\pm \qquad [\hat{\mathbf{T}}_+, \hat{\mathbf{T}}_-] = 2\hat{\mathbf{T}}_3, \tag{41}$$

Using the concepts of the Cartan subalgebra and raising/lowering operators from above, we now see how to treat the states inside any (irreducible) representation. We choose in each representation a basis of states such that $\mathbf{T}_3$ is diagonal; in the defining representation for example we already have such a basis, while the basis in Eq. (35) for the adjoint would need to be changed. We label any basis state by a ket (as in "bra-ket notation") containing its $\mathbf{T}_3$ eigenvalue m: the state is labeled as $|m\rangle$, where

$$\mathbf{T}_3\,|m\rangle = m\,|m\rangle. \tag{42}$$

A very important observation is that the states $\mathbf{T}_\pm|m\rangle$, if they are non-zero vectors, have $\mathbf{T}_3$ eigenvalues equal to $m \pm 1$.

Proof:

$$\begin{aligned}\mathbf{T}_3\ \mathbf{T}_\pm|m\rangle &= ([\mathbf{T}_3, \mathbf{T}_\pm] + \mathbf{T}_\pm\mathbf{T}_3)\ |m\rangle = (\pm\mathbf{T}_\pm + \mathbf{T}_\pm\mathbf{T}_3)|m\rangle \\ &= (\pm\mathbf{T}_\pm + \mathbf{T}_\pm\, m)|m\rangle = (m \pm 1)\ \mathbf{T}_\pm|m\rangle. \end{aligned} \tag{43}$$

Assuming some normalization convention for the states, we thus have $\mathbf{T}_\pm|m\rangle \propto |m \pm 1\rangle$, if such a state exists. □

This motivates the classification of irreducible representations in SU(2) (stated here without proof):

1. An irreducible representation can be uniquely classified (up to representation equivalence) by a non-negative integer or half-integer number l: $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ The number l is called spin and the representation is labeled by $\{l\}$.

2. The representation $\{l\}$ has $2l + 1$ states. We label them by $|l; m\rangle$, where m runs from l to $-l$ in integer steps: $m = l, l-1, l-2, \ldots, -l$. As before m is the $\mathbf{T}_3$ eigenvalue of the basis state, but we also add the label l to signify which representation this state is part of.

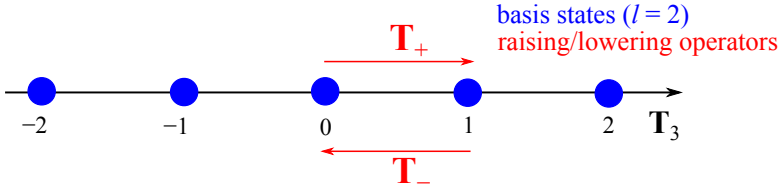
3. The generators act on the basis states in $\{l\}$ by

$$\hat{\mathbf{T}}_3\,|l; m\rangle = m\,|l; m\rangle, \tag{44}$$
$$\hat{\mathbf{T}}_+\,|l; m\rangle = N_{m+1}\,|l; m+1\rangle, \tag{45}$$
$$\hat{\mathbf{T}}_-\,|l; m\rangle = N_m|l; m-1\rangle, \tag{46}$$

where the normalization factor is $N_m = \sqrt{(l+m)(l-m+1)}$ (depends both on the state m and the representation l). With this definition, the basis states in $\{l\}$ are orthonormal: $\langle l; m|l; m'\rangle = \delta_{mm'}$.

Comment: Eq. (44)–(46) are sufficient to reconstruct the generators in the representation $\{l\}$ as $(2l + 1) \times (2l + 1)$ matrices, since the equations describe exactly where each basis state $|l; m\rangle$ is mapped. The generator $\mathbf{T}_3$ can have non-zero entries only on the diagonal (the values m), the generator $\mathbf{T}_+$ only on the superdiagonal and $\mathbf{T}_-$ on the subdiagonal.

Given the classification above, it is possible to draw the basis states of a representation on a line as dots placed at their $\mathbf{T}_3$ eigenvalues. We show below a picture of the $l = 2$ representation; it has 5 states. It is possible to hop between any two neighboring states by applying the raising/lowering operators, as indicated by the red arrows.



Comment: As a final remark we discuss how the representations of the group SO(3) are related to the SU(2) representations. We note that at the Lie algebra level $\mathfrak{su}(2)$ is isomorphic

to $\mathfrak{so}(3)$. At the group level, however, there is a difference, since an angle of $\alpha = 360°$ does not give the identity element in Eq. (38) as it should if we imagine a 360° real rotation in 3D space. For this reason only representations with an integer l are allowed for the group SO(3): $l = 0, 1, 2, 3, \ldots$

2.3 The SU(3) group

The group SU(3) has the Lie algebra $\mathfrak{su}(3)$, which consists of traceless Hermitian 3×3 complex matrices $\mathbf{X}$: $\mathbf{X} = \mathbf{X}^\dagger$, $\operatorname{Tr} \mathbf{X} = 0$. Analogously to the Pauli matrices used as a basis in the 2×2 case, we now define the Gell-Mann matrices:

$$\boldsymbol{\lambda}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$\boldsymbol{\lambda}_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \boldsymbol{\lambda}_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{47}$$

We define the generators of $\mathfrak{su}(3)$ (in the defining representation) as

$$\mathbf{T}_a = \tfrac{1}{2} \boldsymbol{\lambda}_a, \tag{48}$$

where a goes from 1 to 8. An element of $\mathfrak{su}(3)$ is then written as $i\alpha^a \mathbf{T}_a$, where α^a are real coefficients. The SU(3) group element (in the defining representation) can then be written as $e^{i\alpha^a \mathbf{T}_a}$. In this representation the following normalization for the generators holds:

$$\operatorname{Tr}(\mathbf{T}_a \mathbf{T}_b) = \tfrac{1}{2} \delta_{ab}. \tag{49}$$

Analogously to SU(2) we now list some of the properties of SU(3):

- The dimension of the group/algebra is 8. The dimension of the defining representation is 3, the dimension of the adjoint is 8.
- The Cartan subalgebra is 2-dimensional: SU(3) thus has rank 2. It is convenient to take for its basis the generators $\mathbf{T}_3$ and $\mathbf{T}_8$, which are already diagonal in the defining representation (and thus commute). To show that the $\{\mathbf{T}_3, \mathbf{T}_8\}$ generators form a maximal commuting sabalgebra, consider a general linear combination of the other generators written in the defining representation:

$$\mathbf{X} := \sum_{a \neq 3,8} \alpha^a \mathbf{T}_a = \begin{pmatrix} 0 & \alpha^1 - i\alpha^2 & \alpha^4 - i\alpha^5 \\ \alpha^1 + i\alpha^2 & 0 & \alpha^6 - i\alpha^7 \\ \alpha^4 + i\alpha^5 & \alpha^6 + i\alpha^7 & 0 \end{pmatrix}. \tag{50}$$

  Assuming that $\mathbf{X}$ is also in the Cartan subalgebra, it must commute with both $\mathbf{T}_3$ and $\mathbf{T}_8$; the following result follows for the commutator with $\mathbf{T}_3$ by explicit computation:

$$[\mathbf{T}_3, \mathbf{X}] = \tfrac{1}{4} \begin{pmatrix} 0 & 2(\alpha^1 - i\alpha^2) & \alpha^4 - i\alpha^5 \\ 2(-\alpha^1 - i\alpha^2) & 0 & -\alpha^6 + i\alpha^7 \\ -\alpha^4 - i\alpha^5 & \alpha^6 + i\alpha^7 & 0 \end{pmatrix} = \mathbf{0}, \tag{51}$$

  where $\mathbf{0}$ is the zero matrix. This is possible only if $\alpha^1 = \alpha^2 = \alpha^4 = \alpha^5 = \alpha^6 = \alpha^7 = 0$, so the Cartan algebra indeed consists of only $\mathbf{T}_3$ and $\mathbf{T}_8$.

- Since the Cartan generators $\mathbf{T}_3$ and $\mathbf{T}_8$ commute, they can be simultaneously diagonalized. We choose such a common eigenbasis in every irreducible representation. The states can then be denoted by their m_3 and m_8 eigenvalues under $\mathbf{T}_3$ and $\mathbf{T}_8$, respectively:

$$\hat{\mathbf{T}}_3 \,|m_3, m_8\rangle = m_3 \,|m_3, m_8\rangle, \qquad \hat{\mathbf{T}}_8 \,|m_3, m_8\rangle = m_8 \,|m_3, m_8\rangle. \tag{52}$$

  Example: in the defining representation, the generators are the diagonal matrices

$$\mathbf{T}_3 = \frac{1}{2}\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \qquad \mathbf{T}_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}. \tag{53}$$

  The basis states in this representation are thus labeled by

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \equiv \left|\tfrac{1}{2}, \tfrac{1}{2\sqrt{3}}\right\rangle, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv \left|-\tfrac{1}{2}, \tfrac{1}{2\sqrt{3}}\right\rangle, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \equiv \left|0, -\tfrac{1}{\sqrt{3}}\right\rangle. \tag{54}$$

- The 6 non-Cartan generators can again be combined into 3 pairs of raising/lowering operators:

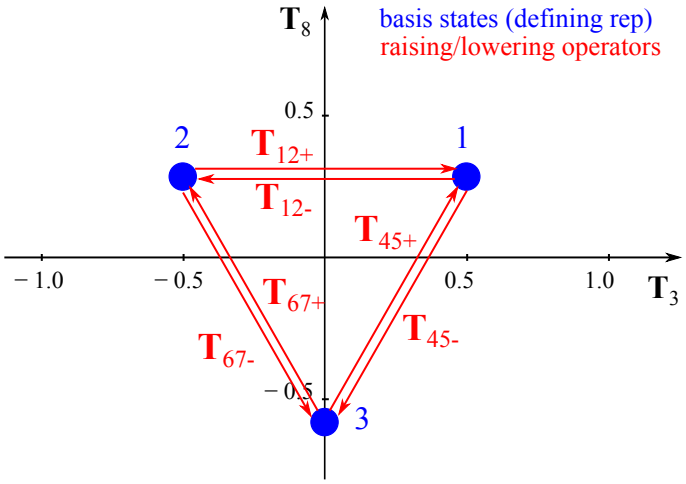
$$\hat{\mathbf{T}}_{12\pm} := \hat{\mathbf{T}}_1 \pm i\hat{\mathbf{T}}_2, \qquad \hat{\mathbf{T}}_{45\pm} := \hat{\mathbf{T}}_4 \pm i\hat{\mathbf{T}}_5, \qquad \hat{\mathbf{T}}_{67\pm} := \hat{\mathbf{T}}_6 \pm i\hat{\mathbf{T}}_7. \tag{55}$$

  By explicit computation in the defining representation, we can show for example that

$$\left.\begin{aligned} \left[\hat{\mathbf{T}}_3, \hat{\mathbf{T}}_{67\pm}\right] &= \pm\left(-\tfrac{1}{2}\right)\hat{\mathbf{T}}_{67\pm} \\ \left[\hat{\mathbf{T}}_8, \hat{\mathbf{T}}_{67\pm}\right] &= \pm\left(\tfrac{\sqrt{3}}{2}\right)\hat{\mathbf{T}}_{67\pm}, \end{aligned}\right\} \Rightarrow \quad \hat{\mathbf{T}}_{67\pm}\,|m_3, m_8\rangle \propto \left|m_3 \pm (-1/2), m_8 \pm \sqrt{3}/2\right\rangle. \tag{56}$$

  This is an analogous relation to the raising/lowering operators in SU(2), except that now both eigenvalues of a state are changed simultaneously, and by a number different than 1 (due to the modified commutation relations).

Analogous to SU(2), we can draw the basis states as dots in the eigenvalue space of the Cartan generators. We now have 2 Cartan generators, so the picture is 2D. The raising/lowering operators change the eigenvalues as indicated by the arrows.



When drawing a different representation (we shall not classify them for SU(3) in these notes), the red arrows change the eigenvalues by the same amounts (since this is dictated by the commutation relations), but the blue dots representing the states of the representation form a different pattern in the plane; this pattern is called the weight lattice.

3 Applications of symmetries to HEP

In this section we shall look at some simple applications of group theory in high energy physics.

3.1 Symmetries of a Hamiltonian

Within the context of quantum mechanics, the Hamiltonian operator $\hat{\mathbf{H}}$ is a Hermitian operator acting on an underlying Hilbert space of states.[1] The underlying space is where the wave function $|\Psi(t)\rangle$ lives. The Hamiltonian operator is the energy operator.

The wave function satisfies the Schrödinger equation (with $\hbar = 1$):

$$\hat{\mathbf{H}} \, |\Psi(t)\rangle = i\partial_t \, |\Psi(t)\rangle \, . \tag{57}$$

If the Hamiltonian does not explicitly depend on time, the solution to this equation can be simply written as

$$|\Psi(t)\rangle = e^{-i\hat{\mathbf{H}}t} \, |\Psi(0)\rangle \, , \tag{58}$$

where the exponential is defined by the power series. The reader is invited to confirm that Eq. (58) is indeed the solution by plugging it into Eq. (57). We define the factor in front of the initial wave function at time $t = 0$ to be the time evolution operator, which evolves the wave function forward in time:

$$\hat{\mathbf{U}}(t) := e^{-i\hat{\mathbf{H}}t}. \tag{59}$$

We make the following observations:

- Since $\hat{\mathbf{H}}$ is Hermitian, $\hat{\mathbf{U}}(t)$ is unitary for all t.
- The time evolution operator forms a one parameter subgroup in the group of all unitary operators on the Hilbert space. The multiplication law is simply

  $$\hat{\mathbf{U}}(t_1)\hat{\mathbf{U}}(t_2) = \hat{\mathbf{U}}(t_1 + t_2), \tag{60}$$

  and the inverse is

  $$\hat{\mathbf{U}}(t)^{-1} = \hat{\mathbf{U}}(-t). \tag{61}$$

  The generator of this Lie group is $-\hat{\mathbf{H}}$; the Hamiltonian can thus be viewed as an infinitesimal generator of time translations.

We now discuss the role symmetries of the Hamiltonian play.

Suppose we have a set of linear operators (matrices) $\{\hat{\mathbf{A}}\}$ on the Hilbert space, which form a group. We say that this is the symmetry group of the Hamiltonian if

$$\hat{\mathbf{H}}\hat{\mathbf{A}} = \hat{\mathbf{A}}\hat{\mathbf{H}} \tag{62}$$

[1] A Hilbert space is by definition also a vector space. Although it may be infinite-dimensional, we consider this as a mere technical complication, which does not change the general picture presented in this chapter.

for all operators $\hat{\mathbf{A}}$ in the group. For Lie groups, it is clearly sufficient for the Hamiltonian to be symmetrical if it commutes with all the infinitesimal generators:

$$[\hat{\mathbf{H}}, \hat{\mathbf{T}}_a] = 0. \tag{63}$$

Suppose now that $|E\rangle$ is an energy eigenstate with energy E:

$$\hat{\mathbf{H}}\,|E\rangle = E\,|E\rangle. \tag{64}$$

The state $\hat{\mathbf{A}}|E\rangle$ is then also an energy eigenstate with the same energy:

$$\hat{\mathbf{H}}\;\hat{\mathbf{A}}|E\rangle = \hat{\mathbf{A}}\;\hat{\mathbf{H}}|E\rangle = \hat{\mathbf{A}}\;E|E\rangle = E\;\hat{\mathbf{A}}|E\rangle. \tag{65}$$

Since repeated application of group elements $\hat{\mathbf{A}}$ on the state $|E\rangle$ eventually spans the irreducible representation of the symmetry group to which $|E\rangle$ belongs (if some states were inaccessible, the representation would not be irreducible), the entire irreducible representation consists of energy eigenstates with the energy E.

In this way, we see that symmetries manifest themselves as degeneracies in the energy states. An analogous analysis can be extended to symmetries of any observable (Hermitian operator) on the Hilbert space; symmetries of that observable imply degeneracies in its eigenvalues.

3.2 Symmetries of a Lagrangian

The Lagrangian is symmetric under a group of transformations (acting on the fields) if it is invariant under such transformations, i.e. if $\mathcal{L} = \mathcal{L}'$, where $\mathcal{L}'$ has the transformed fields inserted. This symmetry can be imposed *a priori* (for example as part of the formulation of the theory), or it can be discovered only later through relations between observable quantities.

The symmetry of a Lagrangian has a number of interesting consequences, such as Noether's theorem in the case of Lie groups, which implies the existence of conserved quantities. We shall not study these consequences here, but will instead focus on how to build an invariant Lagrangian.

The most general invariant Lagrangian can be written by writing it in terms of invariants, which are those polynomial functions in the fields that do not change under group transformations. In the case of the group U(1), we already discovered that an invariant term is built by taking a product of fields such that the sum of their charges is zero. We now describe how to build invariant terms in an arbitrary SU(n) group.

We make use of tensor methods, where the underlying states are written as components of tensors. The rules for building such invariants are summarized below:

- Components of a state in the defining representation (dimension n in SU(n), we label it by $\mathbf{n}$) are labeled by an upper index, i.e. $X^i(x)$, where i goes from 1 to n and X labels the field (we could also have used the label ϕ instead of X; it will be easier to label fields transforming in different representations with latin letters X,Y,Z,etc.). The components transform under a group element $\hat{\mathbf{U}} \in$ SU(n) by "transforming their index":

$$X^i \mapsto \quad X'^i = (\mathbf{U})^i{}_j X^j. \tag{66}$$

  The Einstein summation convention applies. Notice the placement of the indices on the matrix $\mathbf{U}$: the first is upper and the second lower, such that the contraction in the index j happens diagonally, while the upper index i remains.

- The conjugate representation of the defining representation is also n-dimensional in SU(n), we label it by $\bar{\mathbf{n}}$. A state transforming under this representation has components carrying a lower index, i.e. X_i, where i goes from 1 to n. The components transform under $\hat{\mathbf{U}} \in \mathrm{SU}(n)$ as

$$X_i \mapsto \quad X_i' = (\mathbf{U}^*)_i{}^j X_j. \tag{67}$$

  Notice the placement of indices in $\mathbf{U}^*$: the second index j is upper, so that it is contracted diagonally, while the original index i is lower, to form again an object with one lower index.

  Note: complex conjugation switches the position of an upper index to a lower one, and vice versa. This applies to both the components of a state (labeled X), as well as the transformations $\mathbf{U}$.

- The states of product representations (or of their irreducible parts) are written as components of tensors with multiple indices. For example, if we have a product $\mathbf{n}\otimes\bar{\mathbf{n}}\otimes\bar{\mathbf{n}}$ of representations (one defining and two conjugate), the states in such a product can be written as $X^i{}_{jk}$, and the components transform as

$$X^i{}_{jk} \mapsto \quad X'^i{}_{jk} = (\mathbf{U})^i{}_{i'} (\mathbf{U}^*)_j{}^{j'} (\mathbf{U}^*)_k{}^{k'} X^{i'}{}_{j'k'}. \tag{68}$$

  Notice that each index gets transformed according to $\mathbf{U}$ or $\mathbf{U}^*$, depending on whether it is upper or lower. We have one upper index for each $\mathbf{n}$ and one lower for each $\bar{\mathbf{n}}$.

  If we expand the unitary transformations via $\mathbf{U} = \mathbb{1} + i\alpha^a \mathbf{T}_a + \ldots$, we can derive the transformation rule of the components $X^i{}_{jk}$ under the action of an infinitesimal generator:

$$(\hat{\mathbf{T}}_a X)^i{}_{jk} = (\mathbf{T}_a)^i{}_m X^m{}_{jk} - (\mathbf{T}_a^*)_j{}^m X^i{}_{mk} - (\mathbf{T}_a^*)_k{}^m X^i{}_{jm}. \tag{69}$$

  The hat on $\hat{\mathbf{T}}$ denotes the complete action on the components, while the unhatted $\mathbf{T}$ denotes the matrix corresponding to the generator in the defining representation. The infinitesimal change of the components is a sum, where one index is infinitesimally transformed at a time; upper indices transform with $\mathbf{T}_a$ (in the defining representation), while lower indices transform with $-\mathbf{T}_a^*$. Again note the placement of indices in the generator matrices, and the the rule of them changing height if complex conjugation is applied.

- Contracting one upper and one lower index can be done in any representation, since the sum is invariant:

$$X^i Y_i \mapsto X'^i Y_i' = (\mathbf{U})^i{}_j X^j \, (\mathbf{U}^*)_i{}^k \, Y_k = (\mathbf{U}^\dagger)^k{}_i (\mathbf{U})^i{}_j \, X^j Y_k = \delta^k{}_j X^j Y_k = X^j Y_j \tag{70}$$

  We used transposition of indices in $\mathbf{U}^*$ to bring them into the correct order for matrix multiplication (sum over last index of the first factor and the first index of the second factor). Transposition transformed $\mathbf{U}^*$ into $\mathbf{U}^\dagger$, after which the Kronecker delta was obtained by using the unitarity of $\mathbf{U}$.

- To form an invariant out of a product of states, each transforming under an irreducible representation, write the product so that all indices are contracted. Beside the field components you are also allowed to use invariant tensors; in SU(n) these are the

Kronecker delta $\delta^i{}_j$ (which simply gives contractions) and the antisymmetric Levi-Civita tensor with n indices: $\varepsilon_{i_1\cdots i_n}$ and $\varepsilon^{i_1\cdots i_n}$. If some indices cannot be contracted and are left over, the result is not an invariant (but instead transforms as a tensor in the uncontracted indices).

Given the rules above, we look at some SU(2) and SU(3) examples:

- SU(2): the defining representation is denoted by $\mathbf{2}$, the conjugate by $\bar{\mathbf{2}}$. The antisymmetric invariant tensors are ε_{ij} and ε^{ij} with 2 indices (with the convention $\varepsilon^{12} = \varepsilon_{21} = 1$). The representations $\mathbf{2}$ and $\bar{\mathbf{2}}$ are equivalent, since we can use the ε to transform a lower index into an upper index and vice versa:

$$X^i \Leftrightarrow X_i = \varepsilon_{ij} X^j. \tag{71}$$

  If $X, Y \sim \mathbf{2}$ (the symbol $\sim$ means "transforms as"), we can form an invariant by

$$X^i Y_i = \varepsilon_{ij} X^i Y^j. \tag{72}$$

  An SU(2) triplet can be written either as a symmetric two index tensor (both indices upper) Z^{ij}, or a traceless Hermitian tensor with one upper and one lower index $Z^i{}_j$. Since it is an adjoint, the upper-lower components can be written as an expansion in generators: $Z^i{}_j = Z^a (\mathbf{T}_a)^i{}_j$, where Z^a are directly the independent degrees of freedom, which are placed into the components of the two index tensor $Z^i{}_j$ in the correct pattern.

- SU(3): the defining representation is labeled by $\mathbf{3}$, its conjugate by $\bar{\mathbf{3}}$; the invariant tensors are ε^{ijk} and ε_{ijk} with $\varepsilon^{123} = \varepsilon_{123} = 1$. A symmetric two index tensor represents for example the representation $\mathbf{6}$ (the reader should convince themselves that the number of degrees of freedom indeed match). If $X, Y \sim \bar{\mathbf{3}}$ and $Z \sim \mathbf{6}$, we can write the invariant by

$$X_i Y_j Z^{ij}. \tag{73}$$

  If instead $X, Y \sim \mathbf{3}$ and $Z \sim \mathbf{6}$, we can write the invariant as

$$(X^*)_i (Y^*)_j Z^{ij}, \tag{74}$$

  noting that conjugation of the components also conjugates the transformation with which they transform. One can also form an invariant out of three triplets: if $X, Y, Z \sim \mathbf{3}$, then the invariant is

$$X^i Y^j Z^k \varepsilon_{ijk}. \tag{75}$$

  This invariant is antisymmetric under the exchange of any two fields (if the fields are scalar), so it can be formed only if X, Y and Z represent three different triplet multiplets (with different degrees of freedom).

  The adjoint is labeled by $\mathbf{8}$, and it can be again written with one upper and one lower index, expanded in the generators (written as matrices in the defining representation): $Z^i{}_j = Z^a (\mathbf{T}_a)^i{}_j$. The components Z^a, where a goes from 1 to 8, are the independent fields; they are placed into the two-index components $Z^i{}_j$ through the stated equation.

3.3 Global and local symmetries

Symmetries (of the Lagrangian density) in field theories can be divided into global and local symmetries. In a global symmetry, the transformation between the fields is the same at all spacetime points x, for example

$$\phi^k(x) \mapsto (e^{i\alpha^a \mathbf{T}_a})^k{}_l \, \phi^l(x). \tag{76}$$

In particular, the parameters α^a specifying the transformation are spacetime constants.

The Lagrangian is invariant under local symmetries if it is invariant under transformations where the parameters are promoted into spacetime functions, i.e. $\alpha^a(x)$. In this case, every space-time point admits its own independent transformation of the fields. Theories with local symmetries are called gauge theories, and the "symmetry" here actually represents a redundancy in the description rather than a transformation between physically distinct states.

Gauge theories are a rich topic, which we shall not cover in these notes in any detail. We shall simply mention a few relevant points regarding them:

- In the "potential" part of the Lagrangian, the global invariants are automatically also local invariants (at a given space-time point, the same transformation is performed on all the fields). In the kinetic part of the Lagrangian, however, local symmetry is imposed only after regular partial derivatives ∂_μ are promoted into covariant derivatives D_μ, defined by

  $$D_\mu := \partial_\mu - igA^a_\mu(x)\hat{\mathbf{T}}_\mathbf{a}, \tag{77}$$

  where g is the gauge coupling, A^a_μ are called the gauge fields, while $\hat{\mathbf{T}}_a$ are the abstract generators of the gauge symmetry (which are, when acting on a specific object, concretely realized in the representation under which that object transforms). The gauge fields A^a_μ transform in such a way that $D_\mu\phi(x)$ transforms the same way as $\phi(x)$: if $\phi(x)$ transforms by

  $$\phi(x)^k \mapsto \phi'^k(x) = (e^{i\alpha^a(x)\mathbf{T}_a})^k{}_l \, \phi^l(x), \tag{78}$$

  then the covariant derivative of that field transforms as

  $$D_\mu\phi^k(x) \mapsto D'_\mu\phi'^k(x) = (e^{i\alpha^a(x)\mathbf{T}_a})^k{}_l(D_\mu\phi^l(x)), \tag{79}$$

  where D'_μ contains the transformed gauge field A'^a_μ.

- Forces in particle physics are implemented with gauge symmetry. The quanta of gauge fields, after quantization, are gauge bosons.

  Example: the Standard Model is a gauge theory with the gauge symmetry $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$, where the factors represent the strong force, the weak force and the hypercharge, respectively. The field content is then specified in terms of irreducible representations of this gauge symmetry group.

3.4 Symmetry breaking

Symmetries in physics can be broken. This can happen in two distinct ways:

1. Explicit breaking occurs when the action (or Lagrangian) is not actually invariant under a certain symmetry. Instead $\mathcal{L}$ contains terms which are not invariant — they explicitly break the symmetry; if such terms are small, the symmetry is approximate.

2. Spontaneous symmetry breaking occurs when the Lagrangian is invariant, but the ground state (minimum of the potential) is not. In such a situation, the laws of physics still have the symmetry, but the symmetry transformations are less straightforward and beside the fields also involve transforming the vacuum state.

We look at an example for each case of symmetry breaking. The observations we make in the examples are general and can be extended to more complicated cases.

3.4.1 Example 1: explicit symmetry breaking

Consider a theory with 2 real scalar fields ϕ_1 and ϕ_2 with the following potential V in the Lagrangian $\mathcal{L} = T - V$:

$$V(\phi_1, \phi_2) = \tfrac{1}{2} m^2 \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}^T \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} + \tfrac{1}{2} \delta m^2 \, \phi_2^2 + \epsilon^2 \phi_1 \phi_2. \tag{80}$$

The parameters m^2, δm^2 and ϵ^2 are real (and such that the minimum of the potential is at $\phi_1 = \phi_2 = 0$), and all have mass dimension 2 (such that m, $\sqrt{\delta m^2}$ and ϵ have mass dimension 1). We can compute the mass-squared matrix with the double derivative to be

$$(\mathbf{M}^2)_{ij} = \left. \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \right|_{\phi_1 = \phi_2 = 0} = \begin{pmatrix} m^2 & \epsilon^2 \\ \epsilon^2 & m^2 + \delta m^2 \end{pmatrix}, \tag{81}$$

and it has the following eigenvalues:

$$m_{1,2}^2 = m^2 + \tfrac{1}{2} \left(\delta m^2 \pm \sqrt{(\delta m^2)^2 + 4\epsilon^4} \right). \tag{82}$$

We make the following observations:

- The m^2 term in Eq. (80) is an O(2) invariant (it is a scalar product of real fields), where $(\phi_1, \phi_2)^T$ transforms as a doublet. The δm^2 and ϵ^2 terms explicitly break this symmetry.

- In the limit $\delta m^2, \epsilon^2 \to 0$, symmetry is restored, since only the first (invariant) term is present. The eigenvalues $m_{1,2}^2$ in this case become degenerate: $m_1^2 = m_2^2$. Components of fields in the same symmetry multiplet have the same mass.

- If the breaking terms $\delta m^2, \epsilon^2 \neq 0$, symmetry is broken. This introduces a splitting of $\pm\sqrt{(\delta m^2)^2 + 4\epsilon^4}$ into the mass-squared eigenvalues. The splitting is controlled entirely by the breaking parameters. If the breaking parameters are small, i.e. $\delta m^2, \epsilon^2 \ll m^2$, then the masses are approximately degenerate: $m_1^2 \approx m_2^2$. Corrections to this relation can then be computed as a power expansion in $\delta m^2/m^2$ and ϵ^2/m^2.

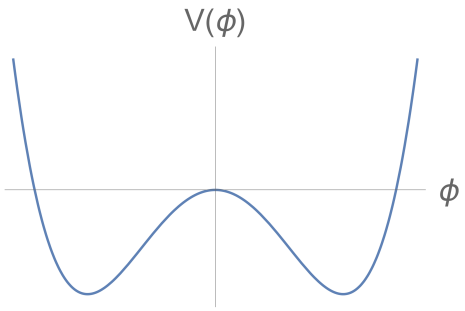
- The above considerations show two possible ways to view this broken symmetry. In the first approach, choosing special values for certain parameters (in our case $\delta m^2 = \epsilon^2 = 0$) can lead to a restored/increased symmetry, even if this were not obvious to us from the beginning. The symmetry manifests itself in the relations between observables, for example degeneracies in the mass spectrum (as in our example). On the other hand, the breaking terms split the degeneracy; if these terms are small, symmetry considerations might still be useful for a better qualitative understanding of approximate degeneracies.

3.4.2 Example 2: spontaneous symmetry breaking

Consider the following potential of 1 real scalar field ϕ:

$$V(\phi) = -\tfrac{1}{2}\mu^2\phi^2 + \tfrac{1}{4}\lambda\phi^4, \tag{83}$$

where μ^2 and λ are real and positive.



The minimization condition $\partial V/\partial\phi = 0$ leads to three candidates for the minimum: $\phi = \pm\sqrt{\mu^2/\lambda}$ and $\phi = 0$. We can see from the schematic picture that there is a local maximum at 0, and two local minima at $\pm\sqrt{\mu^2/\lambda}$; we label $v := \sqrt{\mu^2/\lambda}$.

It is clear that the potential V has a $\mathbb{Z}_2$ symmetry under $\phi \mapsto -\phi$, i.e. $V(\phi) = V(-\phi)$. This is a mirror symmetry. On the other hand, a choice of a local minimum breaks this symmetry, since the two minima are exchanged under the symmetry: $\phi = v \leftrightarrow \phi = -v$.

The quantization of such a theory proceeds by placing oneself into one local minimum, and quantizing the small perturbations around that minimum. Suppose we choose the minimum $\phi = v$. Defining $\tilde{\phi}$ by

$$\tilde{\phi}(x) := \phi(x) - v, \tag{84}$$

we see that $\tilde{\phi}(x)$ represents the kind of perturbations around the minimum that we quantize. Rewriting Eq. (83) with the newly defined degree of freedom $\tilde{\phi}$, we get

$$V = -\tfrac{1}{2}\mu^2(\tilde{\phi} + v)^2 + \tfrac{1}{4}\lambda(\tilde{\phi} + v)^4 \tag{85}$$

$$= \mu^2\,\tilde{\phi}^2 + \sqrt{\mu^2\lambda}\,\tilde{\phi}^3 + \tfrac{1}{4}\lambda\,\tilde{\phi}^4 + \text{const.} \tag{86}$$

where $v = \sqrt{\mu^2/\lambda}$ was used in the second line, and we are not interested in the $\tilde{\phi}$ independent (constant) term. We are indeed at the (local) minimum when $\tilde{\phi} = 0$, since the linear term in $\tilde{\phi}$ disappears and the mass term for $\tilde{\phi}$ is positive.

We now make a few observations:

- It is clear from Eq. (85) that $\tilde{\phi} \mapsto -\tilde{\phi}$ is not a symmetry of the potential V. But the theory still has the $\mathbb{Z}_2$ symmetry $\phi \mapsto -\phi$. This symmetry operation now simultaneously changes both the field $\tilde{\phi}$ and the vacuum, i.e. simultaneously transforming $\tilde{\phi} \mapsto -\tilde{\phi}$ and $v \mapsto -v$ keeps V invariant.

- Once the potential is expanded in powers of the field $\tilde{\phi}$ and the vacuum expectation value v is inserted, as in Eq. (86), the symmetry of the theory is very hard to see from the obtained expression, but is still present. Beside a linear transformation of the fields, it also involves a shift to a new vacuum: $\tilde{\phi} \mapsto -\tilde{\phi} - 2\sqrt{\mu^2/\lambda}$ leaves V invariant.

- Spontaneous breaking can occur for both a discrete or a continuous symmetry. In the continuous case, the chosen minimum is connected to other degenerate minima by continuous transformations, implying flat directions in the potential. These directions correspond to massless degrees of freedom called Goldstone bosons. The Goldstone theorem states that the number of massless particles arising in spontaneous symmetry breaking is equal to the number of broken generators of the continuous symmetry.

- If a continuous local symmetry is spontaneously broken, the Goldstone bosons become unphysical: they are eaten by the gauge bosons, which become massive, and thus acquire a longitudinal degree of freedom. This is called the Higgs mechanism. This occurs for example in the Standard Model, where the local symmetry $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ is spontaneously broken to $\mathrm{SU}(3)_C \times \mathrm{U}(1)_{\mathrm{EM}}$. This was considered in the Standard Model lecture of the summer school; with this remark we conclude these lecture notes.