A Groups: structure and notation

In high energy theory one has plenty of opportunity to use results from group theory, for which Ref. [488] is one of the most often used sources. We will be interested in linear *representations* of groups, i.e., the applications of abstract groups in the form of linear transformations of a vector space, V. By specifying this vector space together with a basis, the group representation is specified in the form of matrices that map vectors from V linearly into vectors that are also in V. A telegraphically brief and cursory review of some of the useful results in group theory provided here cannot possibly compete with the serious sources such as Refs. [565, 258, 287, 581, 201, 80, 333, 260, 334, 256, 447].

A.1 Groups: definitions and applications

This cluster of appendices describes the general algebraic structure of groups and in particular of Lie groups, and then discusses the general properties of the application of groups in physics. This is important for understanding the content of scientific models and their relation with Nature, for the description of which these models were invented.

A.1.1 Axioms and a rough classification

We will need several group-theoretical and algebraic structures and their concrete applications, and they are briefly described here.

Groups

A group G consists of a set of elements $\{a, b, c, \dots\}$ equipped with a binary operation $*$ that satisfies the following axioms (given here with a textual "translation" of the formal symbolism):

1. $\forall a, b \in G,\ a * b \in G;$ (A.1a)
   *For each ($\forall$) two elements a, b from the group G, the result of the binary operation $a * b$ is also in ($\in$) the group G, making the operation $*$ **closed**;*
2. $\forall a, b, c \in G,\ a * (b * c) = (a * b) * c;$ (A.1b)
   *The binary operation $*$ is **associative**, i.e., the result of a repeated application of the binary operation $*$ is independent from the order in which the two operations are computed;*
3. $\exists e \in G,\ \forall a \in G: \quad a * e = e * a = a;$ (A.1c)
   *There exists ($\exists$) a **neutral element** (e) of the group G, such that the results of the binary operations $a * e$ and $e * a$ equal the original element a, for each ($\forall$) a of the group G.*

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

4. $\forall a \in G,\ \exists a^{-1} \in G: \quad a * a^{-1} = a^{-1} * a = e.$ (A.1d)

*For each ($\forall$) element a of the group G, there exists ($\forall$) an **inverse element** a^{-1} in the group, such that the results of the binary operations $a * a^{-1}$ and $a^{-1} * a$ equal the neutral element, e.*

Pedantically, it is not necessary to require that the neutral and the inverse elements are *both-sided*: it suffices to require that there exist, say, the left-neutral element ($\mathbb{1}_L * a = a$) and the left-inverse element ($a_L^{-1} * a = \mathbb{1}$); the existence of the right-neutral element ($a * \mathbb{1}_D = a$) and the right-inverse element ($a * a_D^{-1}$), as well as the equalities ($\mathbb{1}_L = \mathbb{1}_D$ and $a_L^{-1} = a_D^{-1}$) then follow [331, 332].

A group is called abelian (commutative) if the binary operation *commutes*: $(a * b) = (b * a)$, for each two $a, b \in G$; otherwise, the group is called non-abelian (non-commutative). A group G is called *additive* if $*$ is an addition, and *multiplicative* if $*$ is a multiplication.

According to the number of their elements, groups are classified as:

1. Finite, with a finite number of elements. For example, $\mathbb{Z}_2 = \{1, -1; \cdot\}$ is the multiplicative group that consists of two elements, 1 and -1.
2. Countably infinite, with countably infinitely many elements. For example, $\{\mathbb{Z}; +\}$ is the additive group of all (countably many) integers.
3. Continuous, with a continuum of elements, which are further subdivided as:
   (a) Finite-dimensional. For example, $U(1)$ is the multiplicative group of (complex) unitary numbers,[1] i.e., numbers of the form $e^{i\varphi}$, where $\varphi \simeq \varphi + 2\pi$. The number of group elements is continuously infinite, since there is one element for each of the continuously many angles $\varphi \in [0, 2\pi]$. These angles evidently form a subset of the 1-dimensional real axis, $\mathbb{R}^1$, and $U(1)$ is a 1-dimensional group.
   (b) Infinite-dimensional.[2] For example, *Diff*(S^1) is the multiplicative group of all diffeomorphisms (continuous reparametrizations) of the circle, which is a concrete example of the group of general coordinate transformations [☞ Definition 9.1 on p. 319], useful within the theoretical system of strings.

Coset

Besides groups, we also need the concept of a ***coset***: For any group G and its subgroup H, the (right) coset G/H consists of the elements

$$\textbf{coset}: \quad G/H := \{g \simeq g * h \ : \quad g \in G,\ h \in H\}, \tag{A.2}$$

where $*$ is the binary operation in the group G and in the subgroup $H \subset G$. In other words, the coset elements are defined as equivalence classes "up to right 'multiplication' by elements from H." The left coset is defined similarly, and if the group G is abelian, the left and the right coset are identical, of course.

This formal definition describes some very familiar examples:

Days of the week Consider the additive group of integers $\mathbb{Z}_+$ (which is abelian, i.e., commutative), and its subgroup $7\mathbb{Z}_+$, the additive group of integers that are divisible by 7. The coset $\mathbb{Z}_7 := \mathbb{Z}_+/7\mathbb{Z}_+$ is then defined as the additive group of equivalence classes of integers $\mathbb{Z}_+$, where numbers $n \in \mathbb{Z}$ and $n + k$ (for each $k \in 7\mathbb{Z}$) are regarded as equivalent ($\simeq$). The coset $\mathbb{Z}_7$ therefore consists of elements

$$[0 \simeq 7 \simeq 14 \simeq \ldots],\ [1 \simeq 8 \simeq 15 \simeq \ldots],\ [2 \simeq 9 \simeq 16 \simeq \ldots],\ \ldots \tag{A.3}$$

which may be ***represented***:

$$\Big\{[0], [1], [2], [3], [4], [5], [6]\Big\} = \mathbb{Z}_7, \tag{A.4}$$

[1] It follows that their modulus, i.e., absolute value is 1: $z^{-1} = z^* \Rightarrow 1 = z^* z = |z|^2 \Rightarrow |z| = 1$, as $|z| \geqslant 0$.
[2] These are further subdivided into several classes, but this will not concern us here.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

and where the classes $[n]$ may be identified with the days of the week, $[0]$=Sunday, $[1]$=Monday, etc. Indeed, seven days from Monday is again Monday, twenty-one days before Saturday was again Saturday, $7n$ days from Tuesday is again Tuesday, etc.

Circle Consider the additive group of real numbers $\mathbb{R}_+$ and its subgroup of additive numbers $2\pi\mathbb{Z}_+$, the elements of which are integral multiples of 2π. The coset $\mathbb{R}_+/2\pi\mathbb{Z}_+$ then may be identified with the circle S^1, as the coset $\mathbb{R}_+/2\pi\mathbb{Z}_+$ is parametrized by the equivalence classes of real number $[\phi \simeq \phi + 2n\pi]$, for each $n \in \mathbb{Z}$, known as *angles*. Thus, $\mathbb{R}_+/2\pi\mathbb{Z}_+ \cong S^1$.

It is useful to know that all n-dimensional spheres may be identified with the coset

$$S^n := \left\{ \mathrm{x} \in \mathbb{R}^{n+1} : \sum_{i=0}^{n} x_i{}^2 = r^2 \right\} \cong SO(n+1)/SO(n), \tag{A.5}$$

where $SO(n)$ is the group of real and orthogonal $n\times n$ matrices of determinant $+1$. For the details of the isomorphism ($\cong$), the Reader is directed to the literature on Lie groups [565, 258, 581, 256, 80, 260, 333, 447].

Quotient space

The following generalization of the coset turns out to be very useful. Let V be a vector space over the field $\Bbbk$, and $\mu : V \to V$ some mapping of that vector space into itself. One then says that

$$V/\mu := \{[\vec{v} \simeq \mu(\vec{v})] : v \in V\} \tag{A.6}$$

is a quotient space of the vector space V by the action of the mapping μ. The coset is then the special case of the quotient space, where V is regarded as an additive group,[3] and μ is a mapping that preserves this structure, e.g.:

1. Adding integral linear combinations of a specified collection of vectors $\vec{w}_i \in V$, $i = 1, 2, 3, \dots$; indeed, the subset $\{n\vec{v}_0 : n \in \mathbb{Z}\}$ evidently forms a subgroup of the additive group V.
   Example: The 2-dimensional torus $T^2 = \mathbb{R}^2/\Lambda$, where $\Lambda = \{nL_1\hat{\mathrm{e}}_1 + mL_2\hat{\mathrm{e}}_2\}$ is a Cartesian lattice with spacings L_1 and L_2, which are then the circumferences of one and the other circle in the torus.
2. (An)isotropic homothety: rescaling of the (basis) vectors

$$\mu : (\hat{\mathrm{e}}^1, \hat{\mathrm{e}}^2, \dots) \to (\lambda^{a_1}\hat{\mathrm{e}}^1, \lambda^{a_2}\hat{\mathrm{e}}^2, \dots) \in V \tag{A.7}$$

   where $0 \neq \lambda \in \Bbbk$, and since $\vec{a} = a_i\hat{\mathrm{e}}^i$ is an invariantly defined vector, the definition (A.7) is in fact independent of the choice of a basis $\{\hat{\mathrm{e}}^1, \hat{\mathrm{e}}^2, \dots\} \in V$.
   Example: The n-dimensional sphere S^n may be identified also with the quotient space $\mathbb{R}^{n+1}/\mathbb{R}^*_{>0}$, where $\mathbb{R}^*_{>0}$ is the multiplicative group of positive real numbers and the particular action on $\mathbb{R}^{n+1}$ is isotropic:

$$\mu : (x^0, x^1, \dots) \mapsto (\lambda x^0, \lambda x^1, \dots), \quad \lambda > 0. \tag{A.8}$$

   Every element $\mathbb{R}^{n+1}/\mathbb{R}^*_{>0}$ then looks like a ray in the $(n+1)$-dimensional space, starting at the coordinate origin (not including the origin itself) to infinity (not including infinity). Taking one point to represent each ray, e.g., at a *same*, fixed distance from the coordinate origin, then gives the familiar image of the n-dimensional sphere.

[3] The sum of any two vectors is again a vector; adding vectors is associative; $\vec{0}$ is the neutral element with respect to addition; $-\vec{v}$ is the "inverse" vector with respect to addition.

Besides, the physical degrees of freedom in all gauge fields and potentials (including also gravitation) always have the structure of a quotient space [☞ Examples 11.1–11.4, p. 416–417]: the number of physical polarizations of a gauge particle is always smaller than the number of components of the mathematical object (gauge 4-vector, metric tensor, etc.) that must be used to represent the particle.

A.1.2 Lie groups

Of the finite-dimensional continuous groups, of special interest are the so-called Lie groups, G, the elements of which may be written as $g(\mathrm{a}) := \exp\{i\,a^j\,T_j\}$, where summing over j is understood, $\mathrm{a} := (a^1, \dots, a^n)$ is an n-tuple of *parameters*, n the dimension of the group, and T_j are the group generators. Conversely, the group generators, T_j, are obtained by linearizing:

$$T_j := -i\frac{\partial g(\mathrm{a})}{\partial a^j}\Big|_{a^k=0}. \tag{A.9}$$

This means that the space of elements of every Lie group has a well-defined tangent plane in every point, whereupon this group space is a smooth manifold, which locally looks like a Euclidean n-dimensional space. The non-abelian structure of a group G reflects in the difference

$$\mathbb{1} - g(\mathrm{a})\,g(\mathrm{b})\,g(\mathrm{a})^{-1}\,g(\mathrm{b})^{-1} = a^i b^j [T_i, T_j] + \cdots \tag{A.10}$$

where "…" denotes contributions of higher order in parameters a, b.[4] Since a product of group elements must again be a group element, the product $g(\mathrm{a})\,g(\mathrm{b})\,g(\mathrm{a})^{-1}\,g(\mathrm{b})^{-1}$ must be expressible as $g(\mathrm{c}) = \mathbb{1} + ic^j T_j + \cdots$ for some c, from which it follows that the generators T_i must satisfy the relations

$$\big[\,T_j\,,\,T_k\,\big] = i\,f_{jk}{}^m\,T_m, \tag{A.11a}$$

where the coefficients $f_{ij}{}^k = -f_{ji}{}^k$ are the group *structure constants*, and the binary operation [,] is called the *commutator*, or the *Lie bracket*.

Definition A.1 *Formally, the n-dimensional vector space $\mathfrak{A}$, the elements of which are of the form $a^j T_j$, and for which the multiplicative operation*

$$(a^j T_j) * (b^k T_k) := a^j b^k [T_j, T_k] = (i\,a^j b^k\,f_{jk}{}^m)\,T_m \in \mathfrak{A} \tag{A.11b}$$

*is defined is called the **algebra** of the group G.*

Comment A.1 *Since both the Lie groups and the Lie algebras have **continuously** many elements, omitting (the action of) finitely many elements does not change the formal relation between a group and its algebra, but it is important to account for such elements.*

Example A.1 For example, the Pauli matrices

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{A.12}$$

may be used as generators of the group $SU(2)$, the elements of which are of the form $\exp\{i\,a_j \sigma^j\}$, and also as a basis for the $\mathfrak{su}(2)$ algebra, the elements of which are of the

[4] Recall that, in this book, n-vectors as a whole are denoted by upright letters, so a and b are n-vectors with components a^i and b^i, $i = 1, 2, \dots, n$.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

form $a_j\sigma^j$.[5] On the other hand, the $SU(2)$ group elements are defined (in its fundamental representation) as 2×2 unitary matrices with unit determinant. That certainly includes both the 2×2 identity matrix $\mathbb{1} = \exp\{i\,\mathbb{O}\}$ that corresponds to the coordinate origin in the a-space, $\mathrm{a} = (0,0,0)$. However, the $SU(2)$ group also includes the element $-\mathbb{1} = \exp\{i\pi\mathbb{1}\}$, which is omitted in the relation between the $SU(2)$ group and the $\mathfrak{su}(2)$ algebra, since $\pi\mathbb{1} \neq a_j\boldsymbol{\sigma}^j$, and $\pi\mathbb{1} \not\subset \mathfrak{su}(2)$. Thus, although $\exp\{i\,a_j\sigma^j\}$ differs from $SU(2)$ by continuously infinitely many elements of the form $-\mathbb{1}\exp\{i\,a_j\sigma^j\}$, all the omitted elements may be recovered by multiplying (from left or from right) $\exp\{i\,a_j\sigma^j\}$ by $-\mathbb{1}$, the action of which then is the one (and so finite) difference between $SU(2)$ and $\exp\{i\,a_j\sigma^j\}$.

Together, $\mathbb{1} \subset \exp\{i\,a_j\sigma^j\}$ and this omitted element, $-\mathbb{1}$, form a multiplicative finite subgroup of $SU(2)$, denoted $\mathbb{Z}_2 = \{\mathbb{1}, -\mathbb{1}\} \subset SU(2)$. The representations of the group $SU(2)$ that are eigenspaces of the $\exp\{i\pi\mathbb{1}\}$ element of this subgroup $\mathbb{Z}_2 \subset SU(2)$ and have the eigenvalue $+1$ are called tensorial, while the ones with the eigenspace -1 are spinors. Notice that the choice $\mathrm{a} = (0,0,\phi)$ represents the rotation about the third axis; by writing the standard generator as $\frac{1}{2}\boldsymbol{\sigma}^3$, we find this to represent a rotation by the angle $\frac{1}{2}\phi$ – as befits, e.g., a 2-component spin-$\frac{1}{2}$ wave-function, and which is why it changes sign upon a 2π-rotation.

Whereas every algebra $\mathfrak{A}$ gives rise to a group G by means of "exponentiating," i.e., by defining that $g := \exp\{a\} \in G$ for every $a \in \mathfrak{A}$, not infrequently the algebra $\mathfrak{A}$ also *contains* a multiplicative group $\mathfrak{A}^\times$ with the algebra "multiplication" as the binary operation in $\mathfrak{A}^\times$. We thus have the *formal* relation of these three structures $\mathfrak{A}^\times \subset \mathfrak{A} \xrightarrow{\exp} G$.

Example A.2 Note that the $\{1, i, -1, -i\}$-multiples of the 2×2 identity matrix and the Pauli matrices also form a multiplicative group of 16 elements:

$$\begin{aligned}\{&\mathbb{1},\ \sigma^1, \sigma^2,\ \sigma^3,\ i\mathbb{1},\ i\sigma^1,\ i\sigma^2,\ i\sigma^3,\\ &-\mathbb{1},\ -\sigma^1,\ -\sigma^2, -\sigma^3,\ -i\mathbb{1},\ -i\sigma^1,\ -i\sigma^2,\ -i\sigma^3\}.\end{aligned} \tag{A.13}$$

Indeed, the Pauli matrices satisfy *two* relations:

$$\big[\,\sigma^j\,,\,\sigma^k\,\big] = 2i\,\varepsilon^{jk}{}_\ell\,\sigma^\ell, \qquad \text{as well as} \qquad \big\{\,\sigma^j\,,\,\sigma^k\,\big\} = 2\,\delta^{jk}\,\mathbb{1}, \tag{A.14}$$

where $\{A, B\} := A\,B + B\,A$ is the *anticommutator*. Thus, the formula

$$\sigma^j\,\sigma^k = \delta^{jk}\,\mathbb{1} + i\varepsilon^{jk}{}_\ell\,\sigma^\ell \tag{A.15}$$

has, for each $j, k = 1,2,3$, precisely one element on the right-hand side. Thus, multiplying Pauli matrices, one produces the identity matrix and i-multiples of the Pauli matrices, and these must be added to the list of elements of the group. Multiplying in that extended collection, one obtains the (-1)- and $(-i)$-multiples of the Pauli matrices as well as

[5] The standard choice of using the *halves* of the Pauli matrices makes the structure constants equal to i-fold multiples of the Levi-Civita symbol, the same algebra as the rotation generators, $L^x := i(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y})$, etc. cyclically, so that $[L^j, L^k] = i\varepsilon^{jk}{}_\ell L^\ell$. Using the Pauli matrices instead, we have that $[\sigma^j, \sigma^k] = 2i\,\varepsilon^{jk}{}_\ell\,\sigma^\ell$.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

$(-\mathbb{1})$ and $(i\mathbb{1})$, which also must be added to the list of elements. Multiplication within that again-extended collection also yields the $(-i\mathbb{1})$, and this completes the procedure of closing the set: The 16 elements (A.13) form a group with respect to the familiar matrix-multiplication.

For all semisimple Lie algebras,[6] the Killing form

$$g_{jl} := -f_{jk}{}^{m} f_{lm}{}^{k} \tag{A.16}$$

is positive-definite, and serves as a metric tensor, and defines

$$f_{jkl} := f_{jk}{}^{m} g_{ml}, \tag{A.17}$$

which may be shown to be a totally antisymmetric tensor.

Digression A.1 It is worth noting that the relation (A.11a) determines only the antisymmetric product of the generators. The symmetric product, the so-called *anticommutator*, remains free to be specified separately:

$$\{ T_j\,, T_k \} = N\,\delta_{jk}\mathbb{1} + \tfrac{1}{2}\,d_{jk}{}^{m}\,T_m, \qquad \text{if the set } \{\mathbb{1}, T_1, \dots, T_n\} \text{ is complete.} \tag{A.18a}$$

In any given representation, the vector space (representation) V of dimension $r := \dim(V)$ is given, upon which the operators T_j act as $r \times r$ matrices, and the normalization constant N depends on r. Also, these $r \times r$ matrices that play the role of the generators T_j typically satisfy certain additional conditions: they may be symmetric, Hermitian, traceless, etc. If the collection of matrix representatives $\{\mathbb{1}, T_1, \dots, T_n\}$ is complete for the specified type of matrices, the relation (A.18a) follows automatically. Otherwise, one expects that the anticommutators $\{T_j, T_k\}$ include matrices that cannot be represented as the linear combination $\mathbb{1}$ and T_j. Thus, both the existence of the relation (A.18a) and then also the constants $d_{jk}{}^{m}$ strongly depend on the representation of the generators T_j.

If the additional relation (A.18a) exists, its combination with the relation (A.11a) reduces

$$T_j\,T_k = N\,\delta_{jk}\mathbb{1} + (i f_{jk}{}^{m} + \tfrac{1}{2}\,d_{jk}{}^{m})\,T_m \tag{A.18b}$$

to a linear combination of the identity $\mathbb{1}$ and algebra generators T_j – which provides more information than the abstract defining requirement of the Lie algebra (A.11a). Thus, abstract Lie algebras include less structure than what their applications in physics not infrequently have [☞ Section A.1.4].

A.1.3 Groups in (fundamental) physics

Every model of every physical system uses some collection of variables[7] that quantify the system, and imposes relations between those variables in the form of systems of equations and conditions for those variables, appropriate for the physical system being described. That system of equations, together with all conditions on the domain of variables and the operators used to write

[6] A Lie algebra $\mathfrak{A}$ is semisimple if it has no abelian (commutative) direct summand, i.e., no abelian subalgebra that commutes with the whole algebra $\mathfrak{A}$; for the precise statement, see Refs. [581, 256].

[7] In this general description, "variables" includes every mathematical symbol that may have a value, thus, variables include both arguments of some functions, as well as those functions, and various additional parameters.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

out the specified equations forms the *mathematical model*, ***M***, of the physical system. In lieu of experimental results against the model, one regards the model as adequately representing the considered physical system, and one often identifies experimental results in routine conversation. However, it is very important not to confuse in principle the components in this description of the physical system [☞ Figure A.1].

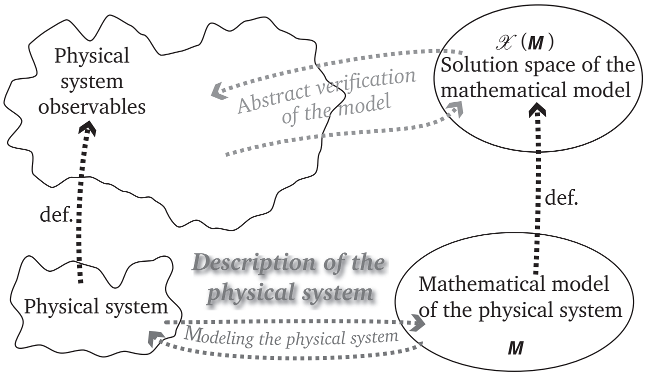

Figure A.1 Relations between the physical system and its observables, as well as the mathematical model and its space of solutions. The smoothness of the mathematical side of this image indicates the fundamental idealizations.

Symmetries of physics systems and symmetry breaking

The situation is actually more complicated than shown in Figure A.1. Namely, the observables in realistic physical systems are usually not specified "once and for all," and their improved definition is an iterative process. In turn, in realistic cases, only some of the theoretically definable observables can be measured in practice, and this subset must be marked. Besides, real physical systems often contain details that are either included in the mathematical model or neglected from it in an iterative or layered fashion. The situation in realistic cases then looks more like the diagram in Figure A.2.

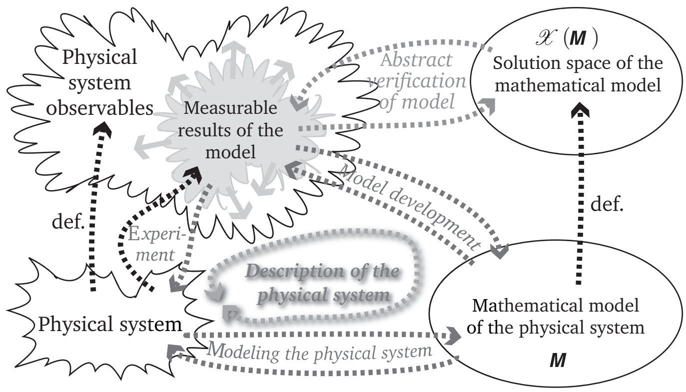

Figure A.2 Relations between the physical system and its observables, as well as the mathematical model, its space of solutions, and their comparisons via experiments.

The collection and domain of variables and operators needed for the description of the physical system usually permits certain changes of those variables and operators, without such re-definitions affecting any concrete, measurable result for the physical system, and obtained by means of this model. Alternatively, the model may be viewed as a mathematical system of equations

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

and conditions, which defines the space of solutions of the system, i.e., the space of solutions of the model, $\mathscr{X}(\boldsymbol{M})$ – regardless whether those solutions can be computed.

The procedure of changing those variables in a way that changes no measurable aspect of the model is called a *symmetry transformation* of the model of the physical system, and the property of the system that permits such a change is called a *symmetry* of the model, i.e., of the system represented by the model. Similarly, instead of the physical system one may consider any system of equations where the "aspect of measurability" need not have a specific meaning. A symmetry is then by definition a transformation that does not change the space of solutions of the specified mathematical model, i.e., the system of equations that represents a physical system. It is important to conceptually distinguish the physical criterion for symmetries (the non-changing of the collection of all measurable results) from the mathematical one (the non-changing of the space of solutions to the given system). It is also important to note that both criteria are *hard*:

1. One cannot a priori know which physically measurable results may possibly exist in a given model, even when these "observables" are "well defined" in general; for example, in classical physics these are "all real C^k-functions over the phase space."[8]
2. Most mathematical systems are insoluble. Indeed, for a randomly chosen system of (differential and algebraic equation) one knows neither how to find or determine the exact solution ("in closed form"), nor of an algorithm of an iterative method for obtaining such a solution, and sometimes even all the known approximations do not suffice for a concrete application. Moreover, it may well be the case that many mathematical systems are not soluble even in principle.

In spite of that – in practice, and so in models that have so far been considered – it is not infrequently possible to definitively determine if a particular transformation is a symmetry of the system or not. In addition, the models used in practice of course form an "infinitesimally" teeny subset of all possible models, and they are chosen precisely so that – besides adequately representing the *interesting physical systems* – they are "sufficiently soluble" so as to be of practical use.

> ***Comment A.2*** *In addition, note that the concrete solutions often do not possess all the symmetries of the system that they solve. In that case, however, the symmetry of the system transforms one concrete solution into another.*

Symmetry transformations evidently satisfy the group axioms (A.1) when the binary operation of two transformations implies their successive application, and one thus speaks of *symmetry groups*, in this mathematical sense. Also, because of this nature of application of group theory, groups are always regarded as groups of concrete transformations within a concrete model, and not as an abstract structure.

That also implies that by the "symmetry of a physical system" one in fact understands the symmetry of the model of that system, conditioned also by the approximations that have been applied in the model by way of neglecting details of the physical system, and mathematical idealizations in the model. As improvements to the model often add details that lessen the number and domain of symmetries, improvements to the model reduce the symmetry group to a subgroup G_1 of the original group G_0. One says that the additional details break the original group into its subgroup, $G_1 \subset G_0$. Although only G_1 is then the "real" symmetry group, the extended structure $G_1 \subset G_0$ provides useful additional information about the model. Not infrequently, the improvements to the

[8] The choice of k and the type of functions (C^0-functions are continuous, C^1-functions are smooth, etc.) depends on the requirements in a concrete application. In classical physics, one *usually* restricts to C^2-functions, as the equations of motion are differential equations of second order, and at least the second derivatives need to be well defined. However, more detailed requirements in the analysis of deformations require higher derivatives, so the required function type must be adapted.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

model may be organized iteratively, corresponding to a *chain* $G_2 \subset G_1 \subset G_0$ of subgroups, which may have an alternative $G_2 \subset G_1' \subset G_0$. The entire *web* of such chains of subgroups provides a hierarchy of model improvements, which corresponds to a hierarchy of physical phenomena and corresponding corrections to measurable results of the model, such as energy.

A simple example

As an illustration of the ideas and concepts depicted in Figure A.2 on p. 457, consider the very familiar example:

$$F = m\,a = m\,\frac{d^2x}{dt^2}, \qquad \text{with the conditions} \qquad x|_{t=t_0} = x_0, \quad \frac{dx}{dt}\Big|_{t=t_0} = v_0. \tag{A.19}$$

In the familiar application of these equations, F and m are parameters in the problem; respectively, the force that acts upon a given body and the mass (measure of inertia) of that body. The function of time, $x = x(t)$, is the position of the body, and x_0, v_0 are boundary (initial) conditions.

The physical system of all bodies of mass m under the action of a force F is thus represented by the model $\boldsymbol{M}$, which is the *abstraction and simplification* of the physical system and which consists of the differential equation (A.19) together with the conditions x_0, v_0 that specify the concrete conditions of a concrete body in a concrete situation to which the model may be applied.[9]

The mathematical solution of this model (assuming that F and m are independent of time) is the function

$$x = x(t) = x_0 + v_0\,(t-t_0) + \frac{F}{2m}(t-t_0)^2, \tag{A.20}$$

so the *space of mathematical solutions* is the abstract space $\mathscr{X}(\boldsymbol{M})$, of the *four-parameter family of functions* $x = x(t; F, m, x_0, v_0)$. Since the t-dependence is determined by the equation (A.20), this space of mathematical solutions has four dimensions, with the coordinates F, m, x_0, v_0. The **phase diagram** is the partitioning of this 4-dimensional space into regions where the model behaves uniformly, and where the passage from one region into another – through some interface region – represents a phase transition in the system.

Similarly named, but something entirely different, is the **phase space**, Φ. For this system, this is the 2-dimensional space parametrized by the values of the pair of functions $\big(x(t), p(t)\big)$, where $p(t) := m\frac{dx}{dt}$. The motion of the body sweeps a path in Φ, parametrized by time. The space of physical observables is then the infinite-dimensional space of all (continuous, and if desired perhaps also analytic and/or square-integrable, etc.) real functions $\mathscr{M}(\Phi)$ over the 2-dimensional phase space Φ.

Finally, the space of measurable model results, $\mathscr{R}(\boldsymbol{M})$, is – *in principle* – a subspace of the space $\mathscr{M}(\Phi)$; see Figure A.3. In this simple model, however, every element of the space $\mathscr{M}(\Phi)$ in fact may be represented as a model result, the question is only whether it is *experimentally possible* to directly measure that result: Namely, it may be that the result must be "factored" into factors and/or summands, which one measures directly and which are then "put together" into the *indirectly* "measured" complex result – but our goal here is not to delve into the details of *experimental* methods.[10] For the model (A.19), in fact, $\mathscr{R}(\boldsymbol{M}) = \mathscr{M}(\Phi)$, i.e., every observable of this physical system may in fact be represented by an (in principle measurable) result of the model (A.19).

Some symmetries: Assuming that the mass of the body is an absolute constant, the differential equation (A.19) has (among others, also) two independent symmetries:

$$P: \begin{cases} x \to -x, \\ F \to -F; \end{cases} \tag{A.21a}$$

[9] Model (A.19) neglects whatever friction may exist, the resistance of the medium through which the body may be moving, etc.

[10] For example, even a relatively simple observable such as speed is *usually* not measured directly, but one measures independently the observables of "traversed distance" and "elapsed time," and speed is then *computed* as their ratio.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

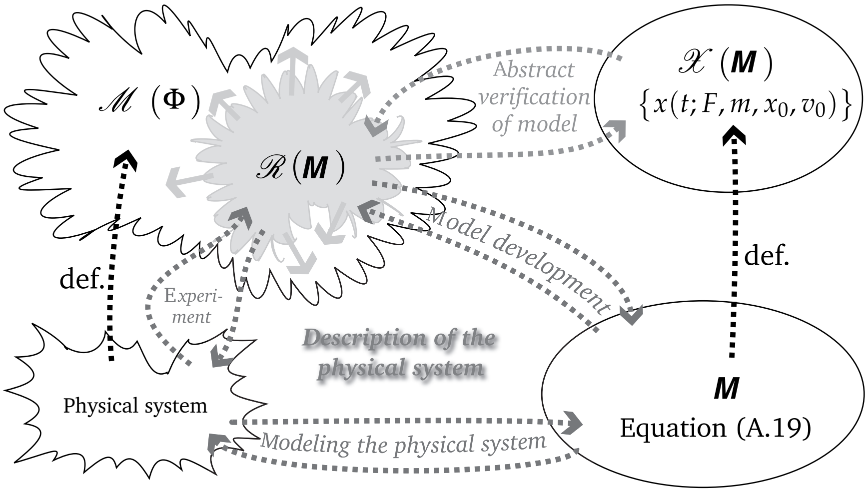

Figure A.3 Relations between the physical system (a body of mass m under the influence of the force F) and its observables, its mathematical model, the space of solutions thereof and the measurable results of this model, as well as their comparisons via experiments. The reason for the relation $\mathscr{R}(\boldsymbol{M}) \subsetneq \mathscr{M}(\Phi)$ is evident: there exist real functions over the phase space $\Phi = \{x, p_x\}$ which are therefore *observables* in the formal sense, but for which no one knows how such a function in fact might be measured, whereupon they do not belong to $\mathscr{R}(\boldsymbol{M})$.

$$T_\tau: \quad t \to t + \tau, \qquad \tau \in \mathbb{R}. \tag{A.21b}$$

The operation P is the mirror reflection of one of the spatial coordinates, which is in 3-dimensional space equivalent to the reflection of all three coordinates through the coordinate original. Its physical meaning is that one is free to pick the direction of measuring the position x either to the right or to the left, from some initially specified point identified as the coordinate origin. Of course, a change of this convention requires the sign of the force to also be changed simultaneously and correspondingly. The other symmetry, T_τ, is the time-translation. The solution (A.20) is also invariant with respect to the first of these two symmetries if the parameters (the integration constants) x_0, v_0 simultaneously satisfy

$$P: \begin{cases} x_0 & \to & -x_0, \\ v_0 & \to & -v_0, \end{cases} \tag{A.22}$$

which is in agreement with the definition x_0, v_0 as the position and speed in the $t = t_0$ moment: if the convention of measuring positions is changed from right-ward to left-ward, all quantities $x(t), x_0, v_0, F$ evidently change signs. With respect to the simultaneous action of the operation P, specified by the relations (A.21)–(A.22), both the system (A.19) and its solutions (A.20) – each a solution by itself! – are invariant with respect to P, i.e., this transformation is a symmetry in the most direct sense.

On the other hand, the physical interpretation of the time-translation is that the behavior of the system does not depend on when we begin to measure time, and this is a symmetry in a slightly indirect sense. Namely, from the fact that the function (A.20) remains a solution of the system (A.19) although it is not invariant under the action of T,[11] it follows that the solutions of a

[11] The constant t_0 does not change under the action of the operation T – indeed, t_0 is chosen so as to be an absolute constant that specifies the beginning of time measurements for the purposes of the applications of the model to a concrete physical model.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

system need not possess all the symmetries of the system [☞ Comment A.2 on p. 458]. However, with t_0 as a fixed constant, we have

$$T_\tau : \qquad t_0 \to t'_0 := t_0 + \tau, \tag{A.23}$$

$$x(t;t_0) = x_0 + v_0(t-t_0) + \frac{F}{2m}(t-t_0)^2$$
$$\to \; x(t;t'_0) = \Big(\underbrace{x_0 - v_0\tau + \frac{F}{2m}\tau^2}_{x'_0}\Big) + \Big(\underbrace{v_0 - \frac{F}{m}\tau}_{v'_0}\Big)(t-t_0) + \frac{F}{2m}(t-t_0)^2. \tag{A.24}$$

Indeed, the symmetry transformation of the system (A.19) changes the integration constants and turns the solution where the time measurement began at t_0, into the solution where the time measurement began at t'_0. Thereby, the symmetry T_τ of the system (A.19) is not a symmetry of a concrete solution (A.20), but transforms one concrete solution into another concrete solution. Thus, the transformation (A.19) is a symmetry of the *entire space* of solutions, $\mathscr{X}(\boldsymbol{M})$.

Symmetries and conservation laws In classical physics, the implications of such symmetries are the content of the (Amalie Emmy) Noether theorem, whereby in classical physics and briefly:

> ***Theorem A.1 (Amalie Emmy Noether)*** *Every continuous symmetry has a corresponding current 4-vector, j^μ, which satisfies the continuity equation, $\partial_\mu j^\mu = 0$, and $\int \mathrm{d}^3\vec{r}\, j^0$ is the corresponding "charge," conserved in time.*

Generally, *additive symmetries* (such as T_τ) have additive conserved charges, and for T_τ, this is the total energy of the system:

$$\frac{\mathrm{d}x}{\mathrm{d}t}\cdot(\text{A.19}) \;\Rightarrow\; m\frac{\mathrm{d}x}{\mathrm{d}t}\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - \frac{\mathrm{d}x}{\mathrm{d}t}F = 0, \tag{A.25}$$

$$\Rightarrow\; \frac{\mathrm{d}E}{\mathrm{d}t} = 0, \quad \text{where} \quad E := \frac{m}{2}\Big(\frac{\mathrm{d}x}{\mathrm{d}t}\Big)^2 - xF. \tag{A.26}$$

The energy E therefore does not change in time, and the Noether theorem connects this property to the symmetry $T_\tau : t \to t+\tau$ of the differential equation (A.19), owing to the fact that $\frac{\mathrm{d}}{\mathrm{d}(t+\tau)} = \frac{\mathrm{d}}{\mathrm{d}t}$. The additivity of energy means that the energy of a combined system is the sum of energies of the individual sub-systems.

Similarly, the multiplicative symmetries (such as P), have multiplicative conserved "charges"; for P, this is the *parity* of the system.[12] $P : f(x) = f(-x) = p f(x)$, $p = \pm 1$. The multiplicativity of parity means that the parity of a combined system is the product of the parities of the individual sub-systems.

— ❦ —

In quantum physics, the relation between symmetries and conserved quantities is even more direct:

> ***Conclusion A.1*** *Let P and Q be two canonically conjugate variables in the sense of the classical description of a system, and P and Q the respectively corresponding operators so $[\mathsf{P},\mathsf{Q}] = i\hbar$. Then the operator $\frac{1}{\hbar}\mathsf{P}$ **generates** translations of the eigenvalues of the operator Q and **vice versa**:*
>
> $$e^{iq_0\mathsf{P}/\hbar}\,\mathsf{Q}\,e^{-iq_0\mathsf{P}/\hbar} = \mathsf{Q} + q_0. \tag{A.27}$$

[12] Of course, the eigenvalue of parity may be written $e^{i\pi\tilde{p}}$, with $\tilde{p} = 0$ or 1, and this $\tilde{p}$ would then be a conserved mod-2 additive quantity. The "multiplicative" practice followed herein is, however, the generally accepted one.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

If the translation of eigenvalues of Q is a symmetry of the system and H the Hamiltonian of the quantum description of this system, then P is a conserved quantity, and ***vice versa***. *More precisely, if $Q \simeq Q + q_0$ then*

$$\frac{dP}{dt} = \frac{1}{i\hbar}\big[H\,,\,P\big] + \frac{\partial P}{\partial t} = 0, \tag{A.28}$$

and conversely: if $[H,P] = 0$ and P does not explicitly depend on time, the eigenvalues of P are conserved quantities and $Q \to Q + q_0$ is a symmetry of the system; the unitary operators $U_{q_0} := \exp\{\frac{i}{\hbar} q_0 P\}$ ***realize*** *this symmetry.*

For a proof and a detailed discussion, see standard textbooks of quantum mechanics, such as [407, 471, 328, 480, 472, 242, 360, 29, 339, 324].

The best known example of this relation is provided by the canonically conjugate pair (position, momentum). In coordinate representation, the operator $\frac{1}{\hbar} p_x = -i\frac{\mathrm{d}}{\mathrm{d}x}$ is indeed the generator of translation in the x-coordinate:

$$e^{iap_x/\hbar} f(x) = e^{a\frac{\mathrm{d}}{\mathrm{d}x}} f(x) = \sum_{k=0}^{\infty} \frac{a^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k} f(x) = \sum_{k=0}^{\infty} \frac{a^k}{k!} f^{(k)}(x) = f(x+a). \tag{A.29}$$

Since p does not explicitly depend on time, the condition that p is a conserved quantity reduces to the condition that p commutes with the Hamiltonian. Since evidently $[p, \frac{1}{2m}p^2] = 0$ this condition becomes $[p, V(x)] = 0 = \frac{\hbar}{i}[\frac{\mathrm{d}}{\mathrm{d}x}, V(x)]$, i.e., that the potential is a constant. Indeed, for a constant potential, $x \to x + x_0$ is a manifest symmetry. Table A.1 lists several examples of often used symmetries and the corresponding conserved quantities. Absolutely essential is the fact that conserved quantities are eigenvalues of operators that generate corresponding transformations. So, for example, the unitary operator $U_{\vec{a}} = \exp\{i\vec{a}\cdot\vec{p}\}$ produces translation in space $\vec{r} \to \vec{r} + \vec{a}$, and the operator $U_{\vec{\xi}} = \exp\{i\vec{\xi}\cdot\vec{r}\}$ produces translation in the momentum space: $\vec{p} \to \vec{p} + \vec{\xi}$.

Table A.1 Some examples of continuous symmetries and corresponding conserved quantities. For various transformations, "charge" denotes various physical quantities; for translation of the phase of complex wave-functions, "charge" is indeed the electric charge.

Symmetry			**Conserved quantity**	
Time translation	$t \to t + t_0$	$\leftrightarrow$	Energy	E
Space translation	$\vec{r} \to \vec{r} + \vec{r}_0$	$\leftrightarrow$	Linear momentum	$\vec{p}$
Rotation (about the z-axis)	$\phi \to \phi + \phi_0$	$\leftrightarrow$	Angular momentum	L_z
Gauge transformation (general)	Phase shift	$\leftrightarrow$	Charge (general)	q
Reflection (through coordinate origin)	$\vec{r} \to -\vec{r}$	$\leftrightarrow$	Parity	P

A.1.4 *Matrix groups and bilinear invariants*

Application of group theory in physics always implies concrete action of the group elements upon concrete physical objects, i.e., upon the mathematical variables that represent those objects in the particular model of the physical system. The linear group action is then always in the form of a linear transformation of a vector space that those mathematical variables span, so these are always matrix groups.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

The most often used matrix groups for $n, p, q \in \mathbb{N}$ are defined as follows [581, 260]:

***GL*(*n*;$\mathbb{k}$)** is the group of invertible $n\times n$ matrices with $\mathbb{k}$-elements, where $\mathbb{k} = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{H}$ denotes the *base field* of rational, real, complex and quaternion numbers, respectively.

***SL*(*n*;$\mathbb{k}$)** is the subgroup of $GL(n;\mathbb{k})$, the elements, $\mathbb{A}$, of which have unit determinant, and so preserve the volume element: $\mathrm{d}^n(\mathbb{A}\,\mathrm{x}) = \mathrm{d}^n\mathrm{x}$ for $\mathrm{x} \in \mathbb{k}^n$.

***O*(*p*, *q*;$\mathbb{k}$)** is the subgroup $GL(p{+}q;\mathbb{k})$ the elements of which are $\eta_{(p,q)}$-orthogonal,

$$\mathbf{L}^T\,\eta_{(p,q)}\,\mathbf{L} = \eta_{(p,q)}, \qquad \eta_{(p,q)} := \mathrm{diag}(\underbrace{1,\ldots,1}_{p\ \text{times}},\underbrace{-1,\ldots,-1}_{q\ \text{times}}) \tag{A.30}$$

and preserve the pseudo-Riemannian scalar product:

$$(\mathrm{x},\mathrm{y})_{(p,q)} := x^\mu\,\eta_{\mu\nu}\,y^\nu = \sum_{\mu=1}^{p} x^\mu\,y^\mu - \sum_{\mu=p+1}^{p+q} x^\mu\,y^\mu, \qquad \mathrm{x},\mathrm{y} \in \mathbb{k}^{p,q}. \tag{A.31}$$

***SO*(*p*, *q*;$\mathbb{k}$)** is the subgroup of $O(p,q;\mathbb{k})$, the elements of which have unit determinant.

***Sp*(2*n*;$\mathbb{k}$)**, for $\mathbb{k} = \mathbb{R}$ or $\mathbb{C}$, is the subgroup of $SL(2n;\mathbb{k})$, the elements of which preserve the symplectic quadratic form

$$\mathrm{x}\wedge\mathrm{y} := 2\sum_{\mu=1}^{2n} x^\mu \wedge x^{\mu+n} = x^\mu\,\Omega_{\mu\nu}\,x^\nu, \qquad [\Omega_{\mu\nu}] = \begin{bmatrix} \mathbb{0} & \mathbb{1} \\ -\mathbb{1} & \mathbb{0} \end{bmatrix}, \tag{A.32}$$

and where the $2n\times 2n$ matrix $[\Omega_{\mu\nu}]$ is called the "symplectic identity."

***U*(*p*, *q*)** is the subgroup of $GL(p{+}q;\mathbb{C})$, the elements of which are unitary and preserve the Hermitian scalar product

$$\langle \mathrm{x}|\mathrm{y}\rangle_{(p,q)} := \sum_{\mu=1}^{p} (x^\mu)^*\,y^\mu - \sum_{\mu=p+1}^{p+q} (x^\mu)^*\,y^\mu, \qquad \mathrm{x},\mathrm{y} \in \mathbb{C}^{p,q}. \tag{A.33}$$

***Sp*(*p*, *q*) = *U*(*p*, *q*;$\mathbb{H}$)** is the subgroup of $GL(2p{+}2q;\mathbb{H})$, the elements of which are quaternion-unitary and preserve the quaternion–Hermitian scalar product

$$\langle \mathrm{z}|\mathrm{w}\rangle_{(p,q)} := \bar{z}^\mu\,\eta_{\mu\nu}\,w^\nu = \sum_{i=\mu}^{p} \bar{z}^\mu\,w^\mu - \sum_{\mu=p+1}^{p+q} \bar{z}^\mu\,w^\mu, \qquad \mathrm{z},\mathrm{w} \in \mathbb{H}^{p,q}, \tag{A.34}$$

where $\bar{x}^\mu$ denotes the quaternion-conjugate of x^μ. This group in fact is not *symplectic*, in the sense that it does not preserve any symplectic quadratic form. Because of this, for the previous group, $Sp(2n;\mathbb{k})$, the base field is always denoted, and for $Sp(p,q)$ never, and by convention $Sp(n) \equiv Sp(n,0)$.

***SU*(*p*, *q*)** is the subgroup of $U(p,q)$, the elements of which have unit determinant.

Quaternion (also known as hyper-complex) numbers and algebra were invented by William Rowan Hamilton, in 1843. Quaternion numbers may be defined as the formal sum $\mathrm{q} = x^0 + ix^1 + jx^2 + kx^3$, where i,j,k are the formal quaternion units that satisfy $i^2 = j^2 = k^2 = ik = -1$. The quaternion-conjugate number is then equal to $\bar{\mathrm{z}} = z^0 - iz^1 - jz^2 - kz^3$. Quaternions do not commute, and $ij = k$ but $ji = -k$, etc. Quaternion "units" may be represented by the complex matrices

$$1 \to \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right], \quad i \to \left[\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right], \quad j \to \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right], \quad k \to \left[\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right], \tag{A.35}$$

so the definitions that use quaternions may be rewritten as complex-matrix definitions.

Owing to frequent use, the base field $\mathbb{R}$ is not written for orthogonal groups so that "$SO(1,3)$" means $SO(1,3;\mathbb{R})$, the base field $\mathbb{C}$ is not written for unitary groups so that "$SU(3)$" means $SU(3;\mathbb{C})$.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

A.1.5 Exercises for Section A.1

✎ ***A.1.1*** *Prove relation (A.10) by explicit expansion of exponential functions.*

✎ ***A.1.2*** *Prove that the collection* $\{\mathbb{1}, (i\boldsymbol{\sigma}^1), (i\boldsymbol{\sigma}^2), (i\boldsymbol{\sigma}^3), -\mathbb{1}, (-i\boldsymbol{\sigma}^1), (-i\boldsymbol{\sigma}^2), (-i\boldsymbol{\sigma}^3)\}$ *forms a group, which is a subgroup of the group (A.13).*

✎ ***A.1.3*** *Show that scaling operations* $\{\mathsf{R}_\rho : x \to \rho x,\ \rho \in (-\infty, +\infty)\}$ *form a group if the binary operation is consecutive application. For* R_ρ *to be a symmetry of the system (A.19), one must require that simultaneously* $\mathsf{R}_\rho : F \to \rho F$*. Determine the action that is consistent with (a) the group structure, and (b) physical meaning of all symbols in the expressions (A.19)–(A.20).*

✎ ***A.1.4*** *Show that* $\{\mathbb{1}, \mathsf{P}\}$ *forms a subgroup of* $\{\mathsf{R}_\rho : x \to \rho x,\ \rho \in (-\infty, +\infty)\}$.

✎ ***A.1.5*** *Show that* $\{\mathbb{1}, \mathsf{T} : t \to -t\}$ *forms a group.*

✎ ***A.1.6*** *Show that* $\{\mathbb{1}, \mathsf{T}, \mathsf{P}, (\mathsf{PT})\}$ *forms a group. Show that* $[\mathsf{T}, \mathsf{P}] = 0$.

A.2 The $U(1)$ group

The multiplicative group of unitary complex numbers, $U(1) = \{e^{i\varphi}, \cdot\}$ where $\varphi \in \mathbb{R}^1$ and $\varphi \simeq \varphi + 2\pi$, is one of the best known in theoretical physics. Representations of the group $U(1)$ are complex functions upon which the group acts by phase transformation: $f \to e^{iq_f\varphi} f$, $f \in \mathbb{C}$. In the general case, the real number q_f is called the "charge" of the particle represented by the function f, and the representation is unambiguously specified by the charge. In the case of the application in electromagnetism, the charge is the electric charge. The $U(1)$ charges are simply additive:

$$U(1) : f \to e^{iq_f\varphi} f,\ g \to e^{iq_g\varphi} g, \qquad \Rightarrow \qquad (fg) \to e^{i(q_f+q_g)\varphi}(fg). \tag{A.36}$$

It is possible – although it is rarely so denoted – to define the $U(1)$ group elements as

$$U(1) = \{\, e^{i\varphi \mathsf{Q}},\ \varphi \simeq \varphi + 2\pi \,\}, \tag{A.37}$$

where the operator Q is the *generator* of the group, and q_f the eigenvalue of the eigenfunction: $\mathsf{Q}f = q_f f$. Thus, $e^{i\varphi\mathsf{Q}} f = e^{i\varphi q_f} f = e^{iq_f\varphi} f$. In complex analysis, q_f is called the winding number of the complex function f, and in the physical application of complex analysis the product $(q_f\varphi)$ is called the *phase* of the function f. This relationship between complex analysis and its application in gauge models is of special importance in string models from the worldsheet perspective [☞ Section 11.2.3], where one easily switches to the complex coordinate system $(\tau, \sigma) \to z = \sigma + i\tau$, and so also to complex analysis.

A.2.1 Exercises for Section A.2

✎ ***A.2.1*** *Given two mutually commuting* $U(1)$ *groups, generated respectively by the mutually commuting Hermitian operators* A *and* B*, show that the two-parameter family of elements* $g_{a,b} := \exp\{i(a\mathsf{A} + b\mathsf{B})\}$ *form the abelian group* $U(1)_A \times U(1)_B$ *for* $a, b \in \mathbb{R}$ *with respect to the usual multiplication.*

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

✎ ***A.2.2*** *Show that any two linearly independent linear combinations of A and B from the previous exercise can serve as generators for the group $U(1)_A \times U(1)_B$. The particular choice $C_+ := A+B$ generates the* **diagonal** *subgroup $U(1)_+ \subset U(1)_A \times U(1)_B$, while the combination $C_- := A-B$ generates the complementary $U(1)_- \subset U(1)_A \times U(1)_B$.*

✎ ***A.2.3*** *Show that $U(1)_+$ and $U(1)_-$ as defined in the previous problem commute with each other and that, as groups, $U(1)_+ \times U(1)_- = U(1)_A \times U(1)_B$.*

A.3 The $SU(2)$ group

This group is familiar from the quantum-mechanical formalism of spin and orbital angular momentum. The group is *generated* by any three operators that satisfy the relations

$$[J_j, J_k] = i\,\varepsilon_{jkl}\,\delta^{lm}\,J_m := i\,\varepsilon_{jk}{}^m\,J_m. \tag{A.38a}$$

The $SU(2)$ group elements are then operators of the form $U_{\vec{a}} := \exp\{ia^j J_j\}$. Conversely, the generators may be formally defined by the relation

$$J_k := \frac{1}{i}\left[\frac{\partial g(\vec{a}\,)}{\partial\, a^k}\right]_{\vec{a}=\vec{0}}. \tag{A.38b}$$

It follows that the quadratic J^2 operator commutes with all three J_j:

$$[J^2, J_j] = 0, \quad j = 1,2,3, \qquad \text{where} \qquad J^2 := J_1^2 + J_2^2 + J_3^2, \tag{A.38c}$$

so the operators[13] J^2 and J_3 have a simultaneous (common) basis of eigenfunctions $|j,m\rangle$:

$$J^2|j,m\rangle = j(j{+}1)|j,m\rangle, \qquad J_3|j,m\rangle = m|j,m\rangle, \tag{A.38d}$$

where

$$\triangle m \in \mathbb{Z}, \quad j := \max(m) \;\Rightarrow\; -j \leqslant m \leqslant +j. \tag{A.38e}$$

It follows that j and m are both either integral (tensorial) or half-integral (spinorial), and that

$$J_\pm := (J_1 \pm iJ_2), \quad J_\pm|j,m\rangle = \sqrt{j(j{+}1) - m(m{\pm}1)}\,|j,m{\pm}1\rangle. \tag{A.38f}$$

Note that

$$J_+|j,j\rangle \equiv 0, \qquad \text{as well as} \qquad J_-|j,-j\rangle \equiv 0 \tag{A.39}$$

by virtue of relations (A.38f), as derived in Digression A.2.

Digression A.2 (Proof of equation (A.38), following Ref. [18]) As no two operators from the collection $\{J_1, J_2, J_3\}$ commute, there is no subsystem of mutually commuting operators that would have a simultaneous eigenbasis. One thus chooses one, usually J_3, to find its eigenstates. Then, one proves by direct computation that

$$[J_i, J^2] = 0, \qquad J^2 := J_1^2 + J_2^2 + J_3^2, \quad i = 1,2,3. \tag{A.40a}$$

[13] The choice of J_3 is arbitrary, and is called the quantization axis choice for angular momentum in quantum mechanics.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Since J_3 and J^2 commute, they have a simultaneous (common) eigenbasis:

$$J^2|\lambda,m\rangle = \lambda|\lambda,m\rangle, \qquad J_3|\lambda,m\rangle = m|\lambda,m\rangle, \tag{A.40b}$$

which may always be ortho-normalized via the Gram–Schmidt procedure:

$$\langle\lambda',m'|\lambda,m\rangle = \delta_{\lambda',\lambda}\,\delta_{m',m}. \tag{A.40c}$$

Using the remaining two operators, J_1, J_2, we define

$$J_\pm := J_1 \pm iJ_2, \qquad (J_\pm)^\dagger = J_\mp, \tag{A.40d}$$

so that

$$J_\pm J_\mp = J_1^2 + J_2^2 \pm J_3, \qquad J^2 = J_+J_- + J_-J_+ + J_3^2, \tag{A.40e}$$

$$[J_3, J_\pm] = \pm J_\pm \quad \text{and} \quad [J_\pm, J_\mp] = \pm 2J_3. \tag{A.40f}$$

Next, check how $J_\pm$ act upon $|\lambda,m\rangle$:

$$J_3\big(J_\pm|\lambda,m\rangle\big) = \big(J_\pm J_3 \pm J_\pm\big)|\lambda,m\rangle = (m\pm1)\big(J_\pm|\lambda,m\rangle\big), \tag{A.40g}$$

so it must be that

$$J_\pm|\lambda,m\rangle = N_\pm(m)\,|\lambda,m{\pm}1\rangle. \tag{A.40h}$$

Thus, the operators $J_\pm$ raise/lower the second eigenvalue, m, but do not change the first, λ.

Since J_3 and J^2 are Hermitian operators, λ, m must be real numbers. Also,

$$\begin{aligned}\lambda = \langle J^2\rangle &= \langle J_+J_-\rangle + \langle J_-J_+\rangle + \langle J_3^2\rangle\\ &= \big\|J_-|\lambda,m\rangle\big\|^2 + \big\|J_+|\lambda,m\rangle\big\|^2 + m^2 \geqslant m^2.\end{aligned} \tag{A.40i}$$

Thus m^2, and so also m, has a maximum; let $j := \max(m)$. Then $J_+|\lambda,+j\rangle$ would have to be proportional to $|\lambda,j{+}1\rangle$. However, since $(j{+}1) > j = \max(m)$, $|\lambda,j{+}1\rangle$ cannot exist. It follows that

$$J_+|\lambda,j\rangle = 0. \tag{A.40j}$$

Applying $\langle\lambda,m|J_-$ to this result, we have that

$$\begin{aligned}0 = \langle\lambda,j|J_-J_+|\lambda,j\rangle &= \langle\lambda,j|(J_1^2+J_2^2-J_3)|\lambda,j\rangle = \langle\lambda,j|(J^{\,2}-J_3^2-J_3)|\lambda,j\rangle\\ &= \lambda - j(j{+}1), \qquad \Rightarrow \qquad \lambda = j(j{+}1).\end{aligned} \tag{A.40k}$$

Following the analogous reasoning for $J_+ \leftrightarrow J_-$, we obtain that

$$\min(m) = -j, \qquad J_-|\lambda,-j\rangle = 0. \tag{A.40l}$$

Renaming the basis $|\lambda,m\rangle \mapsto |j,m\rangle$, we have that

$$J^2|j,m\rangle = j(j{+}1)|j,m\rangle, \qquad J_3|j,m\rangle = m|j,m\rangle. \tag{A.40m}$$

In addition, the operators $J_i, J^2, J_\pm$ can change m only in unit increments. It follows that

$$\triangle m \in \mathbb{Z}, \quad |m| \leqslant j := \max(m), \qquad \Rightarrow \qquad \begin{cases} j \in \mathbb{Z}_{\geqslant 0}, & \text{tensors;}\\ j \in \mathbb{Z}_{\geqslant 0}+\frac{1}{2}, & \text{spinors.}\end{cases} \tag{A.40n}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Similarly, we have that

$$|N_\pm(m)|^2 = \langle j,m|J_\pm J_\mp|j,m\rangle = j(j{+}1) - m(m{\pm}1), \tag{A.40o}$$

so

$$N_\pm(m) = \sqrt{j(j{+}1) - m(m{\pm}1)}. \tag{A.40p}$$

A.3.1 Representations of $SU(2)$

The relations (A.38d)–(A.38f) imply that

$$\begin{aligned} U_{\vec\varphi}|j,m\rangle &= e^{i\varphi^k J_k}|j,m\rangle = \exp\big\{i\big(\varphi^+ J_+ + \varphi^- J_- + \varphi^3 J_3\big)\big\}|j,m\rangle \\ &= \sum_{|m'|\leqslant j} c_{m,m'}|j,m'\rangle. \end{aligned} \tag{A.41}$$

That is, the action of the unitary operator $U_{\vec\varphi}$ does not change j in $|j,m\rangle$ upon which it acts, but – for a general choice of $\vec\varphi$ – transforms any one $|j,m\rangle$ into a linear combination of all $|j,m'\rangle$ with all the permitted values of m'. The abstract vector space

$$V_j := \left\{ \sum_{|m|\leqslant j} c_m|j,m\rangle,\ \ (c_{-j},\dots,c_j) \in \Bbbk^{2j+1} \ \text{ and equation (A.41)} \right\} \cong \Bbbk^{2j+1}, \tag{A.42a}$$

$$U_{\vec\varphi}: V_j \to V_j, \qquad V_j \ \text{ is a } \left\{ \begin{smallmatrix}\text{tensorial}\\ \text{spinorial}\end{smallmatrix} \right\} \text{ representation if } j \left\{ \begin{smallmatrix}\text{integral}\\ \text{half-integral}\end{smallmatrix} \right\} \tag{A.42b}$$

is a $(2j{+}1)$-dimensional (unitary) representation of the $SU(2)$ group, i.e., the $SU(2)$ group maps the vector space V_j into itself, and $SU(2)$ is a group of symmetries of the vector space V_j, for every $2j \in \mathbb{Z}_{\geqslant 0}$. Correspondingly, the same partitioning of representations into these two subclasses is also obtained by partitioning into the eigen-representations of the element $\exp\{i\pi\mathbb{1}\} \in \mathbb{Z}_2 \subset SU(2)$ [☞ Example A.1 on p. 454].

Table A.2 on p. 469 lists the first several such representations. It is important to *keep* in mind that the spaces V_j are not simply copies of $\Bbbk^{2j+1}$ (where $\Bbbk = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$, as required), but imply the $SU(2)$ action (A.41). It follows that no $SU(2)$ representation V_j contains a strictly smaller representation $V_{j'}$, with $j' < j$. One says that every representation V_j is *irreducible*.

Digression A.3 The results (A.40m)–(A.40n) give a *complete list* of irreducible representations of the $SU(2)$ group and its algebra (A.38a):

1. tensorial representations, of which the most familiar are:
   (a) scalars, i.e., invariants, represented by $|0,0\rangle$;
   (b) 3-vectors, represented by the basis $\{|1,-1\rangle, |1,0\rangle, |1,+1\rangle\}$;
   (c) (spin-2) quadrupoles, represented by $\{|2,-2\rangle, |2,-1\rangle, |2,0\rangle, |2,+1\rangle, |2,+2\rangle\}$;
   etc.
2. spinorial representations, of which the most familiar are:
   (a) spin-$\frac12$ systems, represented by the basis $\{|\frac12,-\frac12\rangle, |\frac12,+\frac12\rangle\}$;
   (b) spin-$\frac32$ systems, represented by the basis $\{|\frac32,-\frac32\rangle, |\frac32,-\frac12\rangle, |\frac32,+\frac12\rangle, |\frac32,+\frac32\rangle\}$;
   etc.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Note that the bases of formal vectors $\{|j,m\rangle,\ |m|\leqslant j\}$ are just a formal notation for bases of spherical harmonics $\{Y_j^m(\theta,\phi),\ |m|\leqslant j\}$, which are the coordinate representation of the formal $|j,m\rangle$. For example,

$$|1,+1\rangle \leftrightarrow Y_1^{+1}(\theta,\phi) = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{+i\phi}, \tag{A.43a}$$

$$|1,0\rangle \leftrightarrow \quad Y_1^{0}(\theta,\phi) = +\sqrt{\frac{3}{4\pi}}\cos\theta, \tag{A.43b}$$

$$|1,-1\rangle \leftrightarrow Y_1^{-1}(\theta,\phi) = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{-i\phi}, \tag{A.43c}$$

from which it follows that Cartesian coordinates may be expressed as

$$x = r\,\sin\theta\,\cos\phi = -r\sqrt{\frac{2\pi}{3}}\Big(Y_1^1(\theta,\phi)+Y_1^{-1}(\theta,\phi)\Big) \quad\leftrightarrow\quad |1,+1\rangle+|1,-1\rangle, \tag{A.43d}$$

$$y = r\,\sin\theta\,\sin\phi = i\,r\sqrt{\frac{2\pi}{3}}\Big(Y_1^1(\theta,\phi)-Y_1^{-1}(\theta,\phi)\Big) \quad\leftrightarrow\quad |1,+1\rangle-|1,-1\rangle, \tag{A.43e}$$

$$z = r\,\sin\phi \qquad = r\sqrt{\frac{4\pi}{3}}\,Y_1^0(\theta,\phi) \quad\leftrightarrow\quad |1,0\rangle. \tag{A.43f}$$

Similar relations exist for all bases $\{|j,m\rangle,\ |m|\leqslant j\}$ for $j\in\mathbb{Z}$. The other half of representations, the spinors $\{|j,m\rangle,\ |m|\leqslant j\}$ for $(j+\frac{1}{2})\in\mathbb{Z}$ also have an analogous representation in terms of spherical and Cartesian coordinates, but are less well known, and are double-valued and so are not determined unambiguously.

Table A.2 lists several well-known irreducible representations of the $SU(2)$ group, denoted in several alternative and oft-used forms, and Table A.3 on p. 470 lists the first few spherical harmonics $Y_j^m(\theta,\phi)$, which are the functional representation[14] (in spherical coordinates) of the abstract elements $|j,m\rangle$. Of course, the abstract operators $J_\pm, J_3$ and J^2 also have a corresponding functional representation:

$$J_\pm = \pm e^{\pm i\phi}\Big[\frac{\partial}{\partial\theta}\pm i\cot(\theta)\frac{\partial}{\partial\phi}\Big], \tag{A.44a}$$

$$J_3 = -i\frac{\partial}{\partial\phi}, \tag{A.44b}$$

$$J^2 = -\Big[\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\Big(\sin(\theta)\frac{\partial}{\partial\theta}\Big)+\frac{1}{\sin^2(\theta)}\frac{\partial^2}{\partial\phi^2}\Big]. \tag{A.44c}$$

Evidently and except for J_3, computations with the abstract operators and eigenstates of the $SU(2)$ group are simpler than with the functional representation of these.

— ❦ —

Other than the formal ($|j,m\rangle$) and the functional ($Y_j^m(\theta,\phi)$) notation, the matrix notation is also widely used. It is well known that halves of the Pauli matrices (A.147)

$$\mathbb{J}_1^{(1/2)} = \tfrac{1}{2}\begin{bmatrix}0&1\\1&0\end{bmatrix}, \qquad \mathbb{J}_2^{(1/2)} = \tfrac{1}{2}\begin{bmatrix}0&-i\\i&0\end{bmatrix}, \qquad \mathbb{J}_3^{(1/2)} = \tfrac{1}{2}\begin{bmatrix}1&0\\0&-1\end{bmatrix}, \tag{A.45a}$$

[14] Unfortunately, the word "representation" is used in two slightly different senses: here, it is in the sense of "realization," in distinction from the technical sense of an "$SU(2)$ group representation," according to the definition (A.42).

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Table A.2 Several smallest representations of the $SU(2)$ group; formal ket-notation precisely corresponds to spherical harmonics $|j,m\rangle \leftrightarrow Y_j^m(\theta,\phi)$ when $j \in \mathbb{Z}$.

	Dim.	Formal ket-notation	Index[a]	Matrix
V_0	**1**	$\{\|0,0\rangle\}$	t	$[x]$
$V_{\frac{1}{2}}$	**2**	$\{\|\tfrac{1}{2},-\tfrac{1}{2}\rangle,\|\tfrac{1}{2},+\tfrac{1}{2}\rangle\}$	t^a	$\begin{bmatrix}x^1\\x^2\end{bmatrix}$
V_1	**3**	$\{\|1,-1\rangle,\|1,0\rangle,\|1,+1\rangle\}$	$t^{(ab)}$	$\begin{bmatrix}x=t^{(11)}\\y=t^{(12)}\\z=t^{(22)}\end{bmatrix}$
$V_{\frac{3}{2}}$	**4**	$\{\|\tfrac{3}{2},-\tfrac{3}{2}\rangle,\|\tfrac{3}{2},-\tfrac{1}{2}\rangle,\|\tfrac{3}{2},+\tfrac{1}{2}\rangle,\|\tfrac{3}{2},+\tfrac{3}{2}\rangle\}$	$t^{(abc)}$	$\begin{bmatrix}x^1=t^{(111)}\\ \vdots \\x^4=t^{(222)}\end{bmatrix}$
V_2	**5**	$\{\|2,-2\rangle,\|2,-1\rangle,\|2,0\rangle,\|2,+1\rangle,\|2,+2\rangle\}$	$t^{(abcd)}$	$\begin{bmatrix}x^1=t^{(1111)}\\ \vdots \\x^5=t^{(2222)}\end{bmatrix}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
V_j	**2j+1**	$\{\|j,-j\rangle,\|j,1-j\rangle\cdots,\|j,j-1\rangle,\|j,+j\rangle\}$	$t^{(a_1\cdots a_{2j})}$	$\begin{bmatrix}x^1=t^{(1\cdots1)}\\ \vdots \\x^{2j+1}=t^{(2\cdots2)}\end{bmatrix}$

[a] The indices are $a,b,c\ldots\in\{1,2\}$; round parentheses denote symmetrization: $t^{(ab)}=+t^{(ba)}$.

satisfy the relations (A.38a), which identifies the eigenvectors of the $\mathbb{J}_3^{(1/2)}$-matrix with the eigenvectors of the abstract operator J_3:

$$\begin{bmatrix}1\\0\end{bmatrix} \leftrightarrow |\tfrac{1}{2},+\tfrac{1}{2}\rangle \quad\text{and}\quad \begin{bmatrix}0\\1\end{bmatrix} \leftrightarrow |\tfrac{1}{2},-\tfrac{1}{2}\rangle. \tag{A.45b}$$

In a fully identical fashion, the matrices

$$\mathbb{J}_1^{(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix},\qquad \mathbb{J}_2^{(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix}0&-i&0\\i&0&-i\\0&i&0\end{bmatrix},\qquad \mathbb{J}_3^{(1)} = \begin{bmatrix}1&0&0\\0&0&0\\0&0&-1\end{bmatrix}, \tag{A.46a}$$

also satisfy the relations (A.38a), which identifies the eigenvectors of the $\mathbb{J}_3^{(1)}$-matrix with the eigenvectors of the abstract operator J_3:

$$\begin{bmatrix}1\\0\\0\end{bmatrix} \leftrightarrow |1,+1\rangle,\qquad \begin{bmatrix}0\\1\\0\end{bmatrix} \leftrightarrow |1,0\rangle \quad\text{and}\quad \begin{bmatrix}0\\0\\1\end{bmatrix} \leftrightarrow |1,-1\rangle, \tag{A.46b}$$

and,

$$\mathbb{J}_1^{(3/2)} = \frac{1}{2}\begin{bmatrix}0&\sqrt{3}&0&0\\\sqrt{3}&0&2&0\\0&2&0&\sqrt{3}\\0&0&\sqrt{3}&0\end{bmatrix},\quad \mathbb{J}_2^{(3/2)} = \frac{1}{2}\begin{bmatrix}0&-\sqrt{3}i&0&0\\\sqrt{3}i&0&-2i&0\\0&2i&0&-\sqrt{3}i\\0&0&\sqrt{3}i&0\end{bmatrix},\quad \mathbb{J}_3^{(3/2)} = \frac{1}{2}\begin{bmatrix}3&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-3\end{bmatrix}, \tag{A.46c}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Table A.3 The first few spherical harmonics

$Y_0^0 = \frac{1}{\sqrt{4\pi}}$	$= \frac{1}{\sqrt{4\pi}}$	$Y_2^0 = \sqrt{\frac{15}{16\pi}}(3\cos^2\theta - 1)$	$= \sqrt{\frac{5}{16\pi}}\,\frac{3z^2-r^2}{r^2}$
$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}$	$= -\sqrt{\frac{3}{8\pi}}\,\frac{x+iy}{r}$	$Y_3^3 = -\sqrt{\frac{35}{64\pi}}\sin^3\theta\, e^{3i\phi}$	$= -\sqrt{\frac{35}{64\pi}}\,\frac{(x+iy)^3}{r^3}$
$Y_1^0 = +\sqrt{\frac{3}{4\pi}}\cos\theta$	$= \sqrt{\frac{3}{4\pi}}\,\frac{z}{r}$	$Y_3^2 = \sqrt{\frac{105}{32\pi}}\sin^2\theta\,\cos\theta\, e^{2i\phi}$	$= \sqrt{\frac{105}{32\pi}}\,\frac{(x+iy)^2 z}{r^3}$
$Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta\, e^{2i\phi}$	$= \sqrt{\frac{3}{8\pi}}\,\frac{(x+iy)^2}{r}$	$Y_3^1 = \sqrt{\frac{21}{64\pi}}\sin\theta(1-5\cos^2\theta)e^{i\phi}$	$= -\sqrt{\frac{21}{64\pi}}\,\frac{(x+iy)(5z^2-r^2)}{r^3}$
$Y_2^1 = \sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\phi}$	$= \sqrt{\frac{3}{4\pi}}\,\frac{(x+iy)z}{r}$	$Y_3^0 = \sqrt{\frac{7}{16\pi}}(5\cos^2\theta - 3)\cos\theta$	$= \sqrt{\frac{7}{16\pi}}\,\frac{z(5z^2-3r^2)}{r^3}$

similarly provide a 4-dimensional realization for spin-$\frac{3}{2}$ systems. An analogous matrix realization of the operators $\vec{J}$ and eigenvectors $|j,m\rangle$ is of course possible for all j.

Finally, in the tensor notation, we have

$$t^1 \;\leftrightarrow\; |\tfrac{1}{2},+\tfrac{1}{2}\rangle \qquad \text{and} \qquad t^2 \;\leftrightarrow\; |\tfrac{1}{2},-\tfrac{1}{2}\rangle, \tag{A.47a}$$

which, with the definition $(u,v) := (t^1,t^2)$, implies the definitions

$$\mathfrak{J}_1^{(1/2)} := \tfrac{1}{2}\Big(v\frac{\partial}{\partial u} + u\frac{\partial}{\partial v}\Big), \quad \mathfrak{J}_2^{(1/2)} := \tfrac{i}{2}\Big(v\frac{\partial}{\partial u} - u\frac{\partial}{\partial v}\Big), \quad \mathfrak{J}_3^{(1/2)} := \tfrac{1}{2}\Big(u\frac{\partial}{\partial u} - v\frac{\partial}{\partial v}\Big). \tag{A.47b}$$

For $j=1$, one typically identifies the formal tensor variables $t^{(11)}, t^{(12)}, t^{(22)}$ with the Cartesian x,y,z, respectively, and we have the well-known

$$\mathfrak{J}_1^{(1)} := i\Big(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\Big), \quad \mathfrak{J}_2^{(1)} := i\Big(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\Big), \quad \mathfrak{J}_3^{(1)} := i\Big(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\Big). \tag{A.47c}$$

As each of these notations and representations is convenient in some but not all computations, it behooves the Reader to practice "translating" from any one of these representations into any other one.

It is useful to note that the Levi-Civita symbol,

$$\varepsilon_{ab}: \quad \varepsilon_{12} = 1 = -\varepsilon_{21}, \;\; \varepsilon_{11} = 0 = \varepsilon_{22}\,, \tag{A.48}$$

is invariant with respect to $SU(2)$ transformations, since, using relations (B.38) and after the computation (B.37), it follows that the change of basis $t^a \to \tau^a$ produces

$$\mathrm{d}^2 t := \tfrac{1}{2}\varepsilon_{ab}\,\mathrm{d}t^a \mathrm{d}t^b = \det\Big[\frac{\partial(t^1,t^2)}{\partial(\tau^1,\tau^2)}\Big]\,\tfrac{1}{2}\varepsilon_{ab}\,\mathrm{d}\tau^a \mathrm{d}\tau^b = \mathrm{d}^2\tau, \tag{A.49}$$

since the determinant of $SU(2)$ transformations equals $\det\big[\frac{\partial(t^1,t^2)}{\partial(\tau^1,\tau^2)}\big] = 1$ by definition. The analogous situation holds by definition for all $SU(n)$ groups, but for $SU(2)$, exceptionally, ε_{ab} is a rank-2 tensor, and so may also serve as an (antisymmetric!) metric tensor, which is appropriate for anticommuting variables that are used in supersymmetry [☞ Chapter 10].

A.3.2 The SU(2) and SO(3) groups

Rotations in real, 3-dimensional space may be represented as real, orthogonal 3×3 matrices with unit determinant. Their successive application may be identified with matrix multiplication, which does not commute, and this multiplicative group is denoted $SO(3)$. Its algebra, $\mathfrak{so}(3)$, is identical

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

to the $\mathfrak{su}(2)$ algebra. However, although $\mathfrak{so}(3) = \mathfrak{su}(2)$, the groups $SO(3)$ and $SU(2)$ differ: the $SU(2)$ action upon all representations $V_j = \{|j,m\rangle,\ |m| \leqslant j\}$ is single-valued.

In distinction, the group $SO(3)$ action is single-valued upon integral (tensorial) representations $V_j = \{|j,m\rangle,\ |m| \leqslant j \in \mathbb{Z}\}$, but not upon half-integral (spinorial) representations, $V_j = \{|j,m\rangle,\ |m| \leqslant j,\ (j+\frac{1}{2}) \in \mathbb{Z}\}$. Since a φ-rotation about the x^3-axis acts by $\exp\{i\,\varphi\,J_3\}$ and eigenvalues of J_3 on elements of spinorial representations V_j are half-integral, spinors are "double-valued functions" under $SO(3)$ rotations. By an appropriate change of basis, it is easy to show that the eigenvalues of any one component $\vec{J}$, in any direction, are equal to their J_3 eigenvalues. Thus, the conclusion about double-valuedness of the elements of spinorial representations V_j holds for rotations about any axis. Thus, spinors change their sign upon any 360°-rotation; only 720°-rotations act upon them as the identity.

Since the algebras are identical, $\mathfrak{so}(3) = \mathfrak{su}(2)$, the elements of both algebras – and so also both the $SU(2)$ and the $SO(3)$ generators – are rightfully called angular momenta. Understanding this 2–1 relationship between these groups, $SU(2)$ is the two-fold covering of the $SO(3)$ group, and the elements of the $SU(2)$ group are also frequently called rotations. Pedantically, the $SU(2)$ group is the double covering of the $SO(3)$ group of rotations.

A.3.3 Addition of angular momenta

In the concrete application of the $SU(2)$ group in elementary particle physics, it is important to keep in mind that angular momentum is not a *directly* measurable quantity.

This is partly true also in classical physics of macroscopic bodies: for an ice-skater in a pirouette or a spinning top, the angular momentum cannot be measured directly. Instead, usually, one identifies a "marking" on the spinning object (the ice-skater's face or a pattern on the top), and the angular velocity is determined by following the motion of this marking. Independently, one determines the moment of inertia for the same object in some way,[15] and then *computes* the angular momentum from the so-obtained values of the moment of inertia and the angular velocity. That is, there's no such thing as an "angularmomentumometer."

With elementary particles, the situation is even more indirect: by definition, elementary particles cannot have a "marking" the motion of which one could follow even in principle, so as to measure the angular velocity, compute the moment of inertia, etc. Instead, the angular momentum is even *defined* indirectly. For example, the intrinsic angular moment of an electron – the so-called spin – is in fact a fictive rotation [☞ Digression 4.1 on p. 132] which one *computes*, by way of relation (4.24a), from the *measured* magnetic dipole momentum.

In the situation when we have several magnetic fields, it is perfectly logical to compute their vectorial sum. Conversely, since the dipole momenta of these magnetic fields define spins and orbital angular momenta,[16] to the sum of magnetic fields then corresponds a sum of angular momenta, both intrinsic ("spins") and relative ("orbital").

— ❦ —

The technique of adding angular momenta in quantum theory differs from "ordinary vectorial addition" which is expected in classical physics, and this is discussed in great detail in standard textbooks of quantum mechanics. We recall here the basic relations.

---

[15] In principle, this is possible by approximating the geometry of the object and its mass distribution, whereupon one *computes* the moment of inertia by integrating, or by physically applying a force, and the moment of inertia is *computed* as the ratio of the applied torque and the produced change in its angular velocity.

[16] Although Bohr's model of the atom depicts the electron as a point-particle that rotates about a point-like proton, so that the rotation of the electron's charge forms a current that produces an "orbiting magnetic field," experiments actually only measure this magnetic field, from which then – in turn – one *concludes* about the rotating of the *mental image* of the point-like electron in Bohr's atom.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Let $\{L_1, L_2, L_3\}$ and $\{S_1, S_2, S_3\}$ be two triples of operators, of which each independently satisfies the relations (A.38a) – ***regardless of their physical meaning*** – and let

$$[L_j, S_k] = 0 \quad \text{for every pair of indices } j,k = 1,2,3. \tag{A.50}$$

These two triples then generate two separate copies of the $SU(2)$ group, where elements of one commute with the elements of the other, and we have $SU(2)_L \times SU(2)_S$. One then defines

$$J_j := L_j + S_j \qquad \Rightarrow \qquad [J_j, J_k] = i\varepsilon_{jk}{}^m J_m, \tag{A.51}$$

and the triple J_i generates the *diagonal subgroup* $SU(2)_J \subset SU(2)_L \times SU(2)_S$. For each triple, one defines operators such as J^2 and $J_\pm$, yielding results akin to (A.38), repeating the computations in Digression A.2 on p. 465:

$$L^2|\ell, m_\ell\rangle = \ell(\ell{+}1)|\ell, m_\ell\rangle, \qquad L_3|\ell, m_\ell\rangle = m_\ell|\ell, m_\ell\rangle; \tag{A.52a}$$

$$S^2|s, m_s\rangle = s(s{+}1)|s, m_s\rangle, \qquad S_3|s, m_s\rangle = m_s|s, m_s\rangle; \tag{A.52b}$$

$$J^2|j, m_j\rangle = j(j{+}1)|j, m_j\rangle, \qquad J_3|j, m_j\rangle = m_j|j, m_j\rangle. \tag{A.52c}$$

The relation (A.50) implies that L^2, L_3, S^2, S_3 all mutually commute, so that the tensor product of the eigenbases (A.52a) and (A.52b),

$$|\ell, s; m_\ell, m_s\rangle := |\ell, m_\ell\rangle \otimes |s, m_s\rangle, \tag{A.53a}$$

is a simultaneous eigenbasis of all four operators:

$$L^2|\ell, s; m_\ell, m_s\rangle = \ell(\ell{+}1)|\ell, s; m_\ell, m_s\rangle, \qquad S^2|\ell, s; m_\ell, m_s\rangle = s(s{+}1)|\ell, s; m_\ell, m_s\rangle, \tag{A.53b}$$

$$L_3|\ell, s; m_\ell, m_s\rangle = m_\ell|\ell, s; m_\ell, m_s\rangle, \qquad S_3|\ell, s; m_\ell, m_s\rangle = m_s|\ell, s; m_\ell, m_s\rangle. \tag{A.53c}$$

The operator J_3 commutes with L^2, L_3, S^2, S_3, but is of course not linearly independent since equation (A.51) implies that $J_3 = L_3 + S_3$. We thus also have that

$$J_3|\ell, s; m_\ell, m_s\rangle = (m_\ell{+}m_s)|\ell, s; m_\ell, m_s\rangle. \tag{A.53d}$$

In turn,

$$[J^2, L_3] = 2i\,\varepsilon^{jk}{}_3 L_j S_k = 2i(L_1S_2 - L_2S_1) = -[J^2, S_3], \tag{A.54}$$

and J^2 does not commute with every operator from the collection $\{L^2, L_3, S^2, S_3\}$. Thus, the 4-plet $\{L^2, L_3, S^2, S_3\}$ is a maximal collection of linearly independent mutually commuting operators.

In turn, the operators $\{J^2, L^2, S^2, J_3\}$ also all mutually commute, and since L_3 and S_3 do not commute with J^2, this second operator quartet is also a maximal collection of linearly independent mutually commuting operators. Thus, they too have a simultaneous eigenbasis:

$$J^2|j, \ell, s; m_j\rangle = j(j{+}1)|j, \ell, s; m_j\rangle, \qquad L^2|j, \ell, s; m_j\rangle = \ell(\ell{+}1)|j, \ell, s; m_j\rangle, \tag{A.55a}$$

$$J_3|j, \ell, s; m_j\rangle = m_j|j, \ell, s; m_j\rangle, \qquad S^2|j, \ell, s; m_j\rangle = s(s{+}1)|j, \ell, s; m_j\rangle. \tag{A.55b}$$

In textbooks of quantum mechanics, L are identified with the orbital angular momentum, S with the spin and J with the "total" angular momentum, (e.g., of an electron in a hydrogen atom). Ignoring the fact that J does not include the nuclear spin, and so in reality is *not* the total angular momentum, there exist many situations where there are more than two triples of operators each of which satisfies the relations such as do L and S, and where at least some of such operators have no relation with rotations, even if fictitious. For example, there is no obstruction to add – akin to

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

equations (A.51) – the angular momentum of a nucleon in one nucleus, say, with the isospin of that or any other nucleon.

Thus, L and S as well as their eigenbasis (A.53a) will be referred to as "constituent," and J and the eigenbasis (A.55) will be referred to as "composite."

Of course, since both bases are complete, it follows that every element of one may be expressed in terms of the elements of the other:

$$|\ell, s; m_\ell, m_s\rangle = \sum_{j=|\ell-s|}^{\ell+s} C^{j,m_j}_{\ell,s;m_\ell,m_s} |j, \ell, s; m_j\rangle, \tag{A.56a}$$

$$|j, \ell, s; m_j\rangle = \sum_{\substack{m_\ell=-\ell \\ |m_s|=|m_j-m_\ell|\leqslant s}}^{\ell} (C^{j,m_j}_{\ell,s;m_\ell,m_s})^* |\ell, s; m_\ell, m_s\rangle, \tag{A.56b}$$

where

$$C^{j,m_j}_{\ell,s;m_\ell,m_s} := \langle j, \ell, s; m_j | \ell, s; m_\ell, m_s\rangle \equiv \langle j, m_j | \ell, s; m_\ell, m_s\rangle \tag{A.56c}$$

are the Clebsch–Gordan coefficients, which by standard convention all have real values. In addition, we have:

Theorem A.2 *For the sum of two triples of operators, $L_i + S_i = J_i$, each of which satisfies relations (A.38) and (A.50), the relations (A.52) follow, as well as:*

$$|\ell - s| \leqslant j \leqslant (\ell + s), \qquad |j - \ell| \leqslant s \leqslant (j + \ell), \qquad |j - s| \leqslant \ell \leqslant (j + s), \tag{A.57}$$

$$m_j = m_\ell + m_s, \qquad |m_j| \leqslant j, \qquad |m_\ell| \leqslant \ell, \qquad |m_s| \leqslant s, \tag{A.58}$$

where j, ℓ and s assume precisely once all the integrally separated values within the indicated limits.

Thus, using the notation from the left-most two columns of Table A.2 on p. 469, we have that

$$V_\ell \otimes V_s = \oplus_{j=|\ell-s|}^{(\ell+s)} V_j \qquad \Leftrightarrow \qquad (\mathbf{2\ell{+}1}) \otimes (\mathbf{2s{+}1}) = \oplus_{j=|\ell-s|}^{(\ell+s)} (\mathbf{2j{+}1}). \tag{A.59}$$

For example:

$V_\ell \otimes V_s = V_j$	$\Leftrightarrow$	$(\mathbf{2\ell{+}1}) \otimes (\mathbf{2s{+}1}) = (\mathbf{2j{+}1})$
$V_{1/2} \otimes V_{1/2} = V_1 \oplus V_0$	$\Leftrightarrow$	$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1}$
$V_1 \otimes V_{1/2} = V_{3/2} \oplus V_{1/2}$	$\Leftrightarrow$	$\mathbf{3} \otimes \mathbf{2} = \mathbf{4} \oplus \mathbf{2}$
$V_1 \otimes V_1 = V_2 \oplus V_1 \oplus V_0$	$\Leftrightarrow$	$\mathbf{3} \otimes \mathbf{3} = \mathbf{5} \oplus \mathbf{3} \oplus \mathbf{1}$
$V_2 \otimes V_1 = V_3 \oplus V_2 \oplus V_2$	$\Leftrightarrow$	$\mathbf{5} \otimes \mathbf{3} = \mathbf{7} \oplus \mathbf{5} \oplus \mathbf{3}$

(A.60)

and so on. The first row here corresponds to the detailed relations

$$V_{1/2} = \big\{c_+|\tfrac{1}{2}, +\tfrac{1}{2}\rangle + c_-|\tfrac{1}{2}, -\tfrac{1}{2}\rangle\big\}, \tag{A.61}$$

$$\begin{aligned}\big\{c_+|\tfrac{1}{2}, +\tfrac{1}{2}\rangle + c_-|\tfrac{1}{2}, -\tfrac{1}{2}\rangle\big\} &\otimes \big\{c'_+|\tfrac{1}{2}, +\tfrac{1}{2}\rangle' + c'_-|\tfrac{1}{2}, -\tfrac{1}{2}\rangle'\big\} \\ &= \big\{c_1|1, +1\rangle + c_0|1, 0\rangle + c_{-1}|1, -1\rangle\big\} \oplus \big\{c'_0|0, 0\rangle\big\}\end{aligned} \tag{A.62}$$

where $\{c_+, c_-\}$, $\{c'_+ c'_-\}$ and $\{c_1, c_0, c_{-1}; c'_0\}$ are coefficients in the linear combinations appropriate for the vector spaces $V_{1/2}$, $V'_{1/2}$, V_1 and V_0, and where

$$V_1 : \begin{cases} |1, +1\rangle = |\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle', & \text{(A.63a)} \\ |1, 0\rangle = \tfrac{1}{\sqrt{2}}\Big(|\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle' + |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle'\Big), & \text{(A.63b)} \\ |1, -1\rangle = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle', & \text{(A.63c)} \end{cases}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

$$V_0: \qquad |0,0\rangle = \tfrac{1}{\sqrt{2}}\Big(|\tfrac12,+\tfrac12\rangle|\tfrac12,-\tfrac12\rangle' - |\tfrac12,-\tfrac12\rangle|\tfrac12,+\tfrac12\rangle'\Big). \tag{A.63d}$$

For bigger groups this detailed representation is also possible, but the notation becomes more complicated, so statements expressed in the "dimensional" notation, in the right-hand side of tabulation (A.60), are more often found in the physics literature.

Corollary A.1 *Every representation V_j may be assigned a* ***parity****, $\pi(V_j) := 2j \pmod 2$, so $\pi(V_j) = 0$ for tensors, and $\pi(V_j) = 1$ for spinors [☞ definition (A.42)]. Then it follows that parity is mod-2 additive: $\pi(V_\ell \otimes V_s) \equiv 2(\ell+s) \mod 2$.*

Finally, the tensor/index-notation is also used, especially for larger groups, and in that notation the relations (A.63) become

$$\mathbf{2}\otimes\mathbf{2} = \mathbf{3}\oplus\mathbf{1} \quad \leftrightarrow \quad t^\alpha \otimes u^\beta = \underbrace{v^{(\alpha\beta)}}_{\text{3 comps.}} \oplus \underbrace{\Big(v^{[\alpha\beta]} = v\,\varepsilon^{\alpha\beta}\Big)}_{\text{1 component}}, \tag{A.64}$$

where

$$V_0 = \{c_0|0,0\rangle\} = \{b_0\,t\}, \tag{A.65a}$$
$$V_{1/2} = \{c_+|\tfrac12,+\tfrac12\rangle + c_-|\tfrac12,-\tfrac12\rangle\} = \{b_1\,t^1 + b_2\,t^2\}, \tag{A.65b}$$
$$V_1 = \{c_1|1,+1\rangle + c_0|1,0\rangle + c_{-1}|1,-1\rangle\} = \{b_{11}\,t^{(11)} + b_{12}\,t^{(12)} + b_{22}\,t^{(22)}\}, \tag{A.65c}$$

and so on. The formal variables t for V_0, $\{t^1, t^2\}$ for $V_{1/2}$, $\{t^{(11)}, t^{(12)}, t^{(22)}\}$ for V_1, etc., play the role of basis vectors in the tensor notation. Also, the Levi-Civita symbol $\varepsilon_{\alpha\beta}$ is an $SU(2)$-invariant antisymmetric 2-form, so the antisymmetric rank-2 tensor may be identified with the invariant: $v^{[\alpha\beta]} \mapsto v = (\frac12\varepsilon_{\alpha\beta}v^{[\alpha\beta]})$. Similarly, we have the projections

$$V_1\otimes V_{1/2} \supset V_{1/2} \quad \Leftrightarrow \quad t^{(\alpha\beta)}u^\gamma \mapsto v^\alpha := (\varepsilon_{\beta\gamma}t^{(\alpha\beta)}u^\gamma), \tag{A.66a}$$
$$V_{3/2}\otimes V_{1/2} \supset V_1 \quad \Leftrightarrow \quad t^{(\alpha\beta\gamma)}u^\delta \mapsto v^{(\alpha\beta)} := (\varepsilon_{\gamma\delta}t^{(\alpha\beta\gamma)}u^\delta), \tag{A.66b}$$
$$V_1\otimes V_1 \supset V_1 \quad \Leftrightarrow \quad t^{(\alpha\beta)}u^{(\gamma\delta)} \mapsto v^{(\alpha\gamma)} := (\varepsilon_{\beta\delta}t^{(\alpha\beta)}u^{(\gamma\delta)}), \tag{A.66c}$$
$$V_1\otimes V_1 \supset V_0 \quad \Leftrightarrow \quad t^{(\alpha\beta)}u^{(\gamma\delta)} \mapsto v := (\varepsilon_{\alpha\gamma}\varepsilon_{\beta\delta}t^{(\alpha\beta)}u^{(\gamma\delta)}), \tag{A.66d}$$

and so on.

A.3.4 $SU(2)$-covariant operators and the Wigner–Eckart theorem

Relations (A.38d) and (A.38f) have a very simple generalization from eigen-*vectors* to covariant/eigen-*operators*: If a $(2r+1)$-tuple of operators $\{T^{(r)}_\rho,\ |\rho| \leqslant r\}$ satisfies the relations

$$[\,J^2, T^{(r)}_\rho\,] = r(r+1)T^{(r)}_\rho, \qquad [\,J_3, T^{(r)}_\rho\,] = \rho\,T^{(r)}_\rho, \tag{A.67a}$$

then also

$$[\,J_\pm, T^{(r)}_\rho\,] = \sqrt{r(r+1) - \rho(\rho+1)}\,T^{(r)}_{\rho\pm1}, \tag{A.67b}$$

and the formal vector space $\{\sum_{\rho=-r}^{r} c_\rho T^{(r)}_\rho, c_\rho \in \mathbb{R}\} \cong \mathbb{R}^{2r+1}$ is also an $SU(2)$ group representation. Then we have [362, 363, 471, 328, 480, 134, 391, 407, 472, 360, 29, 339, 242, 3, 110, 324, for example]:

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Theorem A.3 (Wigner–Eckart) *For the $(2r{+}1)$-tuple of operators $\{T^{(r)}_\rho,\ |\rho| \leqslant r\}$ that satisfy relations (A.67), for vectors $|j, m_j; \alpha\rangle$ that satisfy relations (A.38d) and if α represents additional eigenvalues of operators independent of J, we have*

$$\langle j' m'_j; \alpha' | T^{(r)}_\rho | j, m_j; \alpha\rangle = \langle j', m'_j | r, j; \rho, m_j\rangle \, \langle j'; \alpha' \| T^{(r)} \| j; \alpha\rangle, \tag{A.68}$$

where $\langle j'; \alpha' \| T^{(r)} \| j; \alpha\rangle$ is the so-called reduced matrix element (amplitude) that does not depend on m_j, ρ, m'_j, and $\langle j', m'_j | r, j; \rho, m_j\rangle$ is a Clebsch–Gordan coefficient.

This theorem is most often used when ratios of matrix elements are needed where the reduced matrix elements are equal and cancel in the ratio.

— ❦ —

For a practical use of relations (A.56) and the Wigner–Eckart theorem A.3, one needs the numerical values of the Clebsch–Gordan coefficients. To this end one most often uses tables [242, 105] [☞ also [294]], although there is a "closed formula" [328]:

$$C^{j,m_j}_{\ell,s;m_\ell,m_s} = \delta_{m_j,m_\ell+m_s}\, A^j_{\ell,s}\, B^{j,m_j}_{\ell,s;m_\ell,m_s}\, D^{j,m_j}_{\ell,s;m_\ell,m_s}\,, \tag{A.69a}$$

$$\delta_{m_j,m_\ell+m_s} = \begin{cases} 1 & \text{if } m_j = m_\ell + m_s, \\ 0 & \text{if } m_j \neq m_\ell + m_s; \end{cases} \tag{A.69b}$$

$$A^j_{\ell,s} := \sqrt{\frac{(\ell{+}s{-}j)!\,(j{+}\ell{-}s)!\,(s{+}j{-}\ell)!\,(2j{+}1)}{(\ell{+}s{+}j{+}1)!}}\,, \tag{A.69c}$$

$$B^{j,m_j}_{\ell,s;m_\ell,m_s} := \sqrt{(j{+}m_j)!\,(j{-}m_j)!\,(\ell{+}m_\ell)!\,(\ell{-}m_\ell)!\,(s{+}m_s)!\,(s{-}m_s)!}\,, \tag{A.69d}$$

$$D^{j,m_j}_{\ell,s;m_\ell,m_s} := \sum_r \frac{(-1)^r}{(\ell{-}m_\ell{-}r)!\,(s{+}m_s{-}r)!\,(j{-}s{+}m_\ell{+}r)!\,(j{-}\ell{-}m_s{+}r)!\,(\ell{+}s{-}j{-}r)!\,r!}\,, \tag{A.69e}$$

where the sum over r is limited by the facts that division by factors in the denominator produces a zero when

$$r > (\ell{-}m_\ell),\ (s{+}m_s),\ (\ell{+}s{-}j), \qquad r < 0,\ (s{-}j{-}m_\ell),\ (\ell{+}m_s{-}j), \tag{A.69f}$$

which makes the sum finite. Evidently, formula (A.69) is not best suited for quick computations "by heart," but is appropriate for machine computation.

— ❦ —

In view of these well-known results for the $SU(2)$ group and its algebra, we have:

Conclusion A.2 *For applications of any group in physics, it is desired to have, in order of importance (and technical demand):*

1. *the complete list of finite-dimensional unitary representations, such as (A.42),*
2. *the complete list of decompositions of products, such as (A.59),*
3. *the complete list of Clebsch–Gordan coefficients, such as (A.69), or at least a method/algorithm for their computation.*

It is fascinating that for off-shell representations of supersymmetry not even the first task is solved☝, not even in N-extended supersymmetric quantum mechanics [☞ Section 10.4].

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

A.3.5 Exercises for Section A.3

✎ ***A.3.1*** *Using the differential representation (A.44) of* J^2, J_3 *and* $J_\pm$ *as well as the functional representations (A.43a)–(A.43c), verify the general results (A.40m), and (A.40h) with (A.40p) for the cases* $j = 1$, $m = \pm 1, 0$.

✎ ***A.3.2*** *Verify by explicit computation that the matrices (A.46a) satisfy the* $\mathfrak{su}(2)$ *algebra relations (A.38a). Construct the* 3×3 *matrix representative of* $(\vec{\mathbb{J}}^{(1)})^2$.

✎ ***A.3.3*** *Given two separate triples of Hermitian operators,* $\vec{L}$ *and* $\vec{S}$*, satisfying the* $\mathfrak{su}(2)$ *algebra (A.38a) and commuting mutually (A.50), prove that equation (A.51) defines the one and only nontrivial linear combination that also satisfies the* $\mathfrak{su}(2)$ *algebra (A.38a).*

A.4 The $SU(3)$ group

The $SU(3)$ group is defined as the group of 3×3 unitary matrices with unit determinant.

Digression A.4 Corollary A.1 on p. 474 defines parity for representations of the $SU(2)$ group, which is additive for products of representations. Similarly, the $SU(3)$ group has a *triality*: representations are either real with triality 0, or a conjugate pair of complex representations with triality 1 and $-1 \cong 2$. The triality of a product of two representations with trialities t_1 and t_2, respectively, is $(t_1 + t_2) \pmod 3$. Similarly, one defines a mod-n additive "n-ality" of representations of the $SU(n)$ group for every n.

A.4.1 The $\mathfrak{su}(3)$ algebra

As a generalization of the relations (A.38a) for the $SU(2)$ group generators and a special case of the general relation for all Lie algebras (A.11a), the $SU(3)$ group is generated by *eight* operators Q_a that satisfy the relations

$$\big[\, Q_a\,,\, Q_b \,\big] = i\, f_{ab}{}^{c}\, Q_c. \tag{A.70}$$

It is useful to note the $SU(3)$ analogue of the generator matrices (A.45a), i.e., the standard choice among the matrix realizations (of the *doubles*[17]) of the $SU(3)$ generators in the smallest, 3-dimensional and *fundamental* representation are the so-called Gell-Mann matrices:

$$\boldsymbol{\lambda}_1 = \begin{bmatrix} 0&1&0\\ 1&0&0\\ 0&0&0 \end{bmatrix},\quad \boldsymbol{\lambda}_2 = \begin{bmatrix} 0&-i&0\\ i&0&0\\ 0&0&0 \end{bmatrix},\quad \boldsymbol{\lambda}_3 = \begin{bmatrix} 1&0&0\\ 0&-1&0\\ 0&0&0 \end{bmatrix},\quad \boldsymbol{\lambda}_4 = \begin{bmatrix} 0&0&1\\ 0&0&0\\ 1&0&0 \end{bmatrix}, \tag{A.71a}$$

$$\boldsymbol{\lambda}_5 = \begin{bmatrix} 0&0&-i\\ 0&0&0\\ i&0&0 \end{bmatrix},\quad \boldsymbol{\lambda}_6 = \begin{bmatrix} 0&0&0\\ 0&0&1\\ 0&1&0 \end{bmatrix},\quad \boldsymbol{\lambda}_7 = \begin{bmatrix} 0&0&0\\ 0&0&-i\\ 0&i&0 \end{bmatrix},\quad \boldsymbol{\lambda}_8 = \tfrac{1}{\sqrt{3}} \begin{bmatrix} 1&0&0\\ 0&1&0\\ 0&0&-2 \end{bmatrix}. \tag{A.71b}$$

The first three matrices evidently generate one of continuously many $SU(2) \subset SU(3)$ subgroups. This choice of matrix representations of generators shows that the structure constants $f_{abc} = f_{ab}{}^{d}\, g_{dc}$ are totally antisymmetric, $f_{abc} = -f_{bac} = -f_{acb} = -f_{cba}$, and we have

$$f_{123} = 1, \quad f_{458} = f_{678} = \tfrac{\sqrt{3}}{2}, \quad f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = \tfrac{1}{2}. \tag{A.71c}$$

It is useful to know that

$$\mathrm{Tr}(\boldsymbol{\lambda}^a \boldsymbol{\lambda}^b) = 2\delta^{ab}. \tag{A.72}$$

[17] Just as *halves* of Pauli matrices close the $\mathfrak{su}(2)$ algebra, *halves* of the Gell-Mann matrices close the $\mathfrak{su}(3)$ algebra.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

A.4.2 Representations of SU(3)

One may define the ket-notation as well as the matrix notation for every Lie group, but only the dimensional and the tensor/index notation are shown here:

$$\mathbf{1} \simeq t, \qquad \mathbf{3} \simeq t^{\alpha}, \qquad \mathbf{3}^* \simeq t_{\alpha} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma}t^{[\beta\gamma]}, \qquad \alpha,\beta,\gamma,\ldots = 1,2,3, \tag{A.73}$$

$$\mathbf{6} \simeq t^{(\alpha\beta)}, \qquad \mathbf{6}^* \simeq t_{(\alpha\beta)}, \qquad \mathbf{8} \simeq t^{\alpha}{}_{\beta},\ t^{\alpha}{}_{\alpha} \equiv 0, \qquad \mathbf{10} \simeq t^{(\alpha\beta\gamma)}, \qquad \text{etc.} \tag{A.74}$$

Here, e.g., $t^{(\alpha\beta)}$ is the symmetric 3×3 matrix, $t^{[\alpha\beta]}$ is the antisymmetric 3×3 matrix, $t^{\alpha}{}_{\beta}$ is the Hermitian 3×3 matrix the trace of which vanishes, etc. It is important to recall the identities:

$$\varepsilon_{\alpha\beta\gamma}\varepsilon^{\delta\epsilon\phi} = \delta^{\delta}_{\alpha}\delta^{\epsilon}_{\beta}\delta^{\phi}_{\gamma} - \delta^{\delta}_{\alpha}\delta^{\phi}_{\beta}\delta^{\epsilon}_{\gamma} + \delta^{\phi}_{\alpha}\delta^{\delta}_{\beta}\delta^{\epsilon}_{\gamma} - \delta^{\phi}_{\alpha}\delta^{\epsilon}_{\beta}\delta^{\delta}_{\gamma} + \delta^{\epsilon}_{\alpha}\delta^{\phi}_{\beta}\delta^{\delta}_{\gamma} - \delta^{\epsilon}_{\alpha}\delta^{\delta}_{\beta}\delta^{\phi}_{\gamma}, \tag{A.75a}$$

$$\Rightarrow \qquad \varepsilon_{\alpha\beta\gamma}\varepsilon^{\delta\epsilon\gamma} = \delta^{\delta}_{\alpha}\delta^{\epsilon}_{\beta} - \delta^{\epsilon}_{\alpha}\delta^{\delta}_{\beta}, \qquad \varepsilon_{\alpha\beta\gamma}\varepsilon^{\delta\beta\gamma} = 2\delta^{\delta}_{\alpha}, \qquad \varepsilon_{\alpha\beta\gamma}\varepsilon^{\alpha\beta\gamma} = 6. \tag{A.75b}$$

Then,

$$\mathbf{3}\otimes\mathbf{3} = \mathbf{6}_S \oplus \mathbf{3}^*_A \quad\Leftrightarrow\quad t^{\alpha}\,s^{\beta} = t^{(\alpha}s^{\beta)} + t^{[\alpha}s^{\beta]}, \qquad \begin{cases} t^{(\alpha}s^{\beta)} &:= \tfrac{1}{2}\big(t^{\alpha}s^{\beta} + t^{\beta}s^{\alpha}\big), \\ t^{[\alpha}s^{\beta]} &:= \tfrac{1}{2}\big(t^{\alpha}s^{\beta} - t^{\beta}s^{\alpha}\big); \end{cases} \tag{A.76a}$$

where subscripts S and A, respectively, denote the symmetric and antisymmetric parts of a product. Next,

$$\begin{aligned} \mathbf{6}\otimes\mathbf{3} = \mathbf{10}\oplus\mathbf{8} \quad\Leftrightarrow\quad t^{(\alpha\beta)}\,s^{\gamma} &= t^{(\alpha\beta}s^{\gamma)} + \tfrac{4}{3}\,t^{(\alpha[b)}s^{\gamma]}, \\ t^{(\alpha\beta}s^{\gamma)} &:= \tfrac{1}{3}\big(t^{(\alpha\beta)}s^{\gamma} + t^{(\beta\gamma)}s^{\alpha} + t^{(\gamma\alpha)}s^{\beta}\big), \\ t^{(\alpha[\beta)}s^{\gamma]} &:= \tfrac{1}{4}\big(\big(t^{(\alpha\beta)}s^{\gamma} - t^{(\alpha\gamma)}s^{\beta}\big) + \big(t^{(\beta\alpha)}s^{\gamma} - t^{(\beta\gamma)}s^{\alpha}\big)\big) \\ &= \tfrac{1}{4}\big(2t^{(\alpha\beta)}s^{\gamma} - t^{(\alpha\gamma)}s^{\beta} - t^{(\beta\gamma)}s^{\alpha}\big) \end{aligned} \tag{A.76b}$$

where it follows that $t^{(\alpha[\beta)}s^{\gamma]}\,\varepsilon_{\alpha\beta\gamma} \equiv 0$;

$$\mathbf{3}^*\otimes\mathbf{3} = \mathbf{8}\oplus\mathbf{1} \quad\Leftrightarrow\quad t_{\alpha}\,s^{\beta} = \Big(t_{\alpha}s^{\beta} - \tfrac{1}{3}\delta^{\beta}_{\alpha}(t_{\gamma}s^{\gamma})\Big) + \tfrac{1}{3}\delta^{\beta}_{\alpha}(t_{\gamma}s^{\gamma}). \tag{A.76c}$$

Besides, we also have that

$$\tfrac{4}{3}\,t^{(\alpha[\beta)}s^{\gamma]}\,\varepsilon_{\beta\gamma\delta} = t^{(\alpha\beta)}s^{\gamma}\,\varepsilon_{\beta\gamma\delta} =: (t^{(\cdot\cdot)}s^{\cdot})^{\alpha}{}_{\delta} \qquad : \qquad \delta_{\alpha}{}^{\delta}\,(t^{(\cdot\cdot)}s^{\cdot})^{\alpha}{}_{\delta} \equiv 0, \tag{A.76d}$$

so that $(t^{(\cdot\cdot)}s^{\cdot})^{\alpha}{}_{\delta}$ is a Hermitian matrix with vanishing trace. Since $\varepsilon_{\beta\gamma\delta}\varepsilon^{\epsilon\phi\delta} = \delta^{\beta}_{\epsilon}\delta^{\gamma}_{\phi} - \delta^{\gamma}_{\epsilon}\delta^{\beta}_{\phi}$, we have also the “converse” relations:

$$(t^{(\cdot\cdot)}s^{\cdot})^{\alpha}{}_{\delta}\,\varepsilon^{\beta\gamma\delta} = t^{(\alpha\beta)}s^{\gamma} - t^{(\alpha\gamma)}s^{\beta}, \qquad \tfrac{2}{3}(t^{(\cdot\cdot)}s^{\cdot})^{(\alpha}{}_{\delta}\,\varepsilon^{\beta)\gamma\delta} = \tfrac{4}{3}\,t^{(\alpha[\beta)}s^{\gamma]}. \tag{A.76e}$$

Finally, we also need the combination (A.76a), (A.76b) and (A.76c):

$$\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} = \big(\mathbf{6}_S\oplus\mathbf{3}^*_A\big)\otimes\mathbf{3} = \big(\mathbf{10}_S\oplus\mathbf{8}\big)\oplus\big(\mathbf{8}\oplus\mathbf{1}_A\big), \tag{A.76f}$$

where the subscripts S and A, respectively, denote the totally symmetric and totally antisymmetric product, and two 8-plets have a mixed symmetry:

$$(\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3})_S \quad\leftrightarrow\quad t^{(\alpha}u^{\beta}v^{\gamma)}, \qquad (\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3})_A \quad\leftrightarrow\quad t^{[\alpha}u^{\beta}v^{\gamma]}. \tag{A.76g}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

For the cubic expressions with mixed symmetry, there exist many possible choices, one of which follows from the iterative procedure (A.76f):

$$(\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3})_{\mathbf{8}_{(1)}} = \Big(\underbrace{(\mathbf{3}\otimes\mathbf{3})_S}_{\mathbf{6}}\otimes\mathbf{3}\Big)_{\mathbf{8}_{(1)}} \leftrightarrow (t^{(\alpha}u^{\beta)})v^\gamma\varepsilon_{\beta\gamma\delta}, \tag{A.76h}$$

$$(\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3})_{\mathbf{8}_{(2)}} = \Big(\underbrace{(\mathbf{3}\otimes\mathbf{3})_A}_{\mathbf{3}^*}\otimes\mathbf{3}\Big)_{\mathbf{8}_{(2)}} \leftrightarrow (t^{\alpha}u^{\beta}\varepsilon_{\alpha\beta\delta})v^\gamma\big(\delta^\delta_\epsilon\delta^\phi_\gamma - \tfrac{1}{3}\delta^\delta_\gamma\delta^\phi_\epsilon\big). \tag{A.76i}$$

These two expressions provide two linearly independent 3×3 Hermitian matrices with a vanishing trace.

These results indicate that the answer to Exercises A.3.1 and A.3.2 on 476, is given by Weyl's general construction:

> ***Construction A.1 (Weyl)*** *All finitely dimensional unitary representations of every Lie group may be constructed projecting n-fold tensor products of the **fundamental** (**spinorial** for Spin groups) representation, $V^{\otimes n}$, by means of the so-called Young symmetrizer.*

The computations (A.76a)–(A.76f) provide concrete examples of this construction:

0. For the $SU(3)$ group, the fundamental (defining) representation is the complex 3-dimensional, denoted $\mathbf{3}$, also denoted in the tensor representation as $\mathbf{3}=\{t^\alpha,\ \alpha=1,2,3\}$.
1. The product $\mathbf{3}\otimes\mathbf{3}$ may be projected to:
   (a) the symmetric part of the product, $\mathbf{6}$: $t^{(\alpha}u^{\beta)}=+t^{(\beta}u^{\alpha)}$, and
   (b) the antisymmetric part of the product, $\mathbf{3}^*$: $t^{[\alpha}u^{\beta]}=-t^{[\beta}u^{\alpha]}$, which is isomorphic to the conjugate 3-dimensional representation: $t^{[\alpha}u^{\beta]}\varepsilon_{\alpha\beta\gamma}=(t{\cdot}u{\cdot})_\gamma$.
2. The product $\mathbf{6}\otimes\mathbf{3}$ may be projected to:
   (a) the totally symmetric part of the product, $\mathbf{10}$: $t^{(\alpha\beta}u^{\gamma)}=+t^{(\beta\alpha}u^{\gamma)}=+t^{(\gamma\beta}u^{\alpha)}=+t^{(\alpha\gamma}u^{\beta)}$, and
   (b) the part of the product with mixed symmetry, $\mathbf{8}$: $t^{(\alpha[\beta)}u^{\gamma]}=+t^{(\beta[\alpha)}u^{\gamma]}$, but the $\beta\leftrightarrow\gamma$ antisymmetrization in $t^{(\alpha[\beta)}u^{\gamma]}$ is broken by imposing the $\alpha\leftrightarrow\beta$ symmetrization.

Projecting may be understood also as a linear mapping of vector spaces:

$$\mathrm{Sym}:\mathbf{3}\otimes\mathbf{3}\to\mathbf{6}\qquad\text{and}\qquad \mathbf{3}^*=\ker\big(\mathrm{Sym}(\mathbf{3}\otimes\mathbf{3})\big), \tag{A.77}$$

that is, $\mathbf{3}^*$ is the part of the product $\mathbf{3}\otimes\mathbf{3}$ that is annihilated by symmetrization. A consistent and iterative application of this procedure is called "Young symmetrization."

To decompose the triple tensor product, we may use the table of coefficients:

		$\alpha\beta\gamma$	$\alpha\gamma\beta$	$\gamma\alpha\beta$	$\gamma\beta\alpha$	$\beta\gamma\alpha$	$\beta\alpha\gamma$
10	$t^{(\alpha\beta\gamma)}$	+1	+1	+1	+1	+1	+1
8	$t^{(\alpha[\beta)\gamma]}$	+2	−1	−1	−1	−1	+2
8	$t^{[\alpha(\beta]\gamma)}$	+2	+1	−1	+1	−1	−2
1	$t^{[\alpha\beta\gamma]}$	+1	−1	+1	−1	+1	−1

(A.78)

so that, e.g.:

$$\begin{aligned} t^{(\alpha[\beta)\gamma]} &\propto \big(+2t^{\alpha\beta\gamma}-t^{\alpha\gamma\beta}-t^{\gamma\alpha\beta}-t^{\gamma\beta\alpha}-t^{\beta\gamma\alpha}+2t^{\beta\alpha\gamma}\big)\propto\big(2t^{(\alpha\beta)\gamma}-t^{(\alpha\gamma)\beta}-t^{(\gamma\beta)\alpha}\big)\\ &\propto\big[\big(t^{(\alpha\beta)\gamma}-t^{(\alpha\gamma)\beta}\big)+\big(t^{(\alpha\beta)\gamma}-t^{(\beta\gamma)\alpha}\big)\big]\propto\big(t^{\alpha[\beta\gamma]}+t^{\beta[\alpha\gamma]}\big), \end{aligned} \tag{A.79}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

which agrees with the result (A.76b). However, table (A.78) also provides the identity

$$t^{\alpha\beta\gamma} = t^{(\alpha\beta\gamma)} + t^{\big(\alpha[\beta)\gamma]} + t^{\big[\alpha(\beta]\gamma)} + t^{[\alpha\beta\gamma]}, \tag{A.80}$$

which reproduces the decomposition (A.76f).

To simplify decompositions such as (A.76), we use (Alfred) Young tableaux, which provide yet another alternative notation for representations of Lie groups [581, 168] and for the Young symmetrization mentioned in Construction A.1.

Construction A.2 (Young) *The fundamental, complex n-dimensional $SU(n)$ group representation is depicted by a box, $\square$. A symmetric product $Sym(\mathbf{n} \otimes \mathbf{n})$ is depicted by placing two boxes next to each other: $\yng(2)$. An antisymmetric product $\ker(Sym(\mathbf{n} \otimes \mathbf{n}))$ is depicted by placing two boxes one under the other: $\yng(1,1)$. Therefore,*

$$\square \otimes \square = \yng(2) \oplus \yng(1,1). \tag{A.81}$$

*A **Young tableau** is no more than n vertically stacked horizontal series of boxes, where:*

1. *all horizontal series being from the same position on the left,*
2. *no horizontal series has more boxes than the one above it;*
3. *a column of n boxes depicts the $SU(n)$-invariant tensor $\varepsilon^{\alpha_1 \cdots \alpha_n}$,*
   and may be deleted from the tableau.

Example A.3 Decomposition (A.76f), i.e., (A.80) is then depicted as

$$\square \otimes \square \otimes \square = \Big(\yng(2) \oplus \yng(1,1)\Big) \otimes \square = \Big(\yng(3) \oplus \yng(2,1)\Big) \oplus \Big(\yng(2,1) \oplus \yng(1,1,1)\Big), \tag{A.82}$$

where multiplication and decomposition are performed iteratively, by attaching the right-hand box to the left-hand tableau in all possible and permitted ways.

For complete rules for multiplying arbitrary tableaux – and for all Lie groups – the interested Reader is directed to the literature [581, 168].

Example A.4 May it suffice here to list the following four examples:

$$\yng(3) \otimes \square = \yng(4) \oplus \yng(3,1), \qquad \yng(2,1) \otimes \square = \yng(3,1) \oplus \yng(2,2) \oplus \yng(2,1,1), \tag{A.83a}$$

$$\yng(1,1) \otimes \square = \yng(2,1) \oplus \yng(1,1,1), \qquad \yng(2,2) \otimes \square = \yng(3,2) \oplus \yng(2,2,1), \tag{A.83b}$$

where it is understood that tableaux that have more than n vertically stacked boxes are discarded for the $SU(n)$ group. Thus, in the third example, only the first tableau remains for $SU(3)$, but both summands remain for $SU(n)$, $n > 3$.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

When the $SU(3)$ group structure is applied to the "flavor" of hadrons, the 3-dimensional representation, **3**, which is spanned by u-, d- and s-quarks, in the Young tableau notation, the quarks are depicted by a box and antiquarks with a column of two boxes. Then, it is clear that:

1. Mesons (bound states of a quark and an antiquark) are depicted by Young tableaux from the product $\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array} \otimes \square$, i.e., $\mathbf{3}^* \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8}$.
2. Baryons (bound states of tri quarks) are depicted by Young tableaux (A.82).
3. Other $SU(3)_f$ group representations can appear only in "exotic" bound states such as di-mesons $(\bar{q}\,\bar{q}\,q\,q)$, di-baryons $(q\,q\,q\,q\,q\,q)$, etc.

There exist two useful combinatorial formulae, for which we first need a function that associates to every box one more than the total number of boxes to the right and below a given box. Because of the geometric shape of the union of the counted boxes, this function is called the "hook number." In Young tableaux (A.84) the values of the "hook numbers" are inscribed into the boxes:

$$\begin{array}{|c|c|}\hline 3&1\\\hline 1\\\cline{1-1}\end{array}\,, \qquad \begin{array}{|c|c|c|}\hline 4&2&1\\\hline 1\\\cline{1-1}\end{array}\,, \qquad \begin{array}{|c|c|c|}\hline 5&3&1\\\hline 3&1\\\cline{1-2} 1\\\cline{1-1}\end{array}\,, \qquad \text{etc.} \tag{A.84}$$

Then, the representation depicted by the tableau YT appears

$$n_{YT} := \frac{N!}{\text{product of "hook numbers"}} \tag{A.85}$$

times in the tensor product $V^{\otimes N}$.

Example A.5 For the examples in the series (A.84) this formula yields:

$$\frac{3!}{1{\cdot}3{\cdot}1} = 2, \qquad \frac{4!}{1{\cdot}2{\cdot}4{\cdot}1} = 3, \qquad \frac{6!}{1{\cdot}3{\cdot}5{\cdot}1{\cdot}3{\cdot}1} = 16. \tag{A.86}$$

This number is also the dimension of the representation of the permutation group S_N represented by this Young tableau, so that $N! = \sum (n_{YT})^2$, with the sum extending over all the tableaux with N boxes.

Example A.6 For baryons represented as 3-quark bound states, we cited the fact that $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$; see the discussion around equation (2.40) and also in Section 4.4. The formula (A.85) then proves that there are two separately counted, 2-dimensional representations of permutation symmetries S_3:

$$\square \otimes \square \otimes \square = \left(\tfrac{3!}{3\cdot2\cdot1}\right)\begin{array}{|c|}\hline 3\\\hline 2\\\hline 1\\\hline\end{array} \oplus \left(\tfrac{3!}{3\cdot1\cdot1}\right)\begin{array}{|c|c|}\hline 3&1\\\hline 1\\\cline{1-1}\end{array} \oplus \left(\tfrac{3!}{3\cdot2\cdot1}\right)\begin{array}{|c|c|c|}\hline 3&2&1\\\hline\end{array} \quad = 1\cdot\begin{array}{|c|}\hline 3\\\hline 2\\\hline 1\\\hline\end{array} \oplus 2\cdot\begin{array}{|c|c|}\hline 3&1\\\hline 1\\\cline{1-1}\end{array} \oplus 1\cdot\begin{array}{|c|c|c|}\hline 3&2&1\\\hline\end{array}\,, \tag{A.87a}$$

$$3! = \left(\tfrac{3!}{3\cdot2\cdot1} = 1\right)^2 + \left(\tfrac{3!}{3\cdot1\cdot1} = 2\right)^2 + \left(\tfrac{3!}{3\cdot2\cdot1} = 1\right)^2 = 1 + 4 + 1, \tag{A.87b}$$

where the "hook numbers" are inserted into the respective boxes on the right, to aid the computation. These two separately counted identical representations in the middle of the expansion, depicted as $\begin{array}{|c|c|}\hline \ &\ \\\hline \ \\\cline{1-1}\end{array}$, have mixed symmetry and correspond to the baryon octets, **8**. In turn, the totally antisymmetric and the totally symmetric representations occur at the beginning and the end of the expansion, respectively; these are both 1-dimensional (unique) and occur once in the expansion.

For the second formula, for the $SU(n)$ group, inscribe into every row of boxes the ascending series of integers, starting with n in the top row, with $(n-1)$ in the second row and so on; these are called the "box numbers." The dimension of the $SU(n)$ representation depicted by the tableau YT is then given by the formula

$$d_{YT} = \frac{\text{product of "box numbers"}}{\text{product of "hook numbers"}}. \tag{A.88}$$

Example A.7 For the examples (A.84) and the $SU(4)$ group, we have the dimensions

$$\frac{(4{\cdot}5)(3)}{1{\cdot}3{\cdot}1} = 20, \qquad \frac{(4{\cdot}5{\cdot}6)(3)}{1{\cdot}2{\cdot}4{\cdot}1} = 45, \qquad \frac{(4{\cdot}5{\cdot}6)(3{\cdot}4)(2)}{1{\cdot}3{\cdot}5{\cdot}1{\cdot}3{\cdot}1} = 64, \tag{A.89}$$

while for the $SU(3)$ group, the same tableaux have the dimensions

$$\frac{(3{\cdot}4)(2)}{1{\cdot}3{\cdot}1} = 8, \qquad \frac{(3{\cdot}4{\cdot}5)(2)}{1{\cdot}2{\cdot}4{\cdot}1} = 15, \qquad \frac{(3{\cdot}4{\cdot}5)(2{\cdot}3)(1)}{1{\cdot}3{\cdot}5{\cdot}1{\cdot}3{\cdot}1} = 8. \tag{A.90}$$

Note that the formula for dimensions of $SU(n)$ tableaux (A.88) automatically returns zero if the tableau contains a column of more than n boxes: for $SU(2)$, the third tableau in the sequence (A.84) yields $\frac{(2\cdot3\cdot4)(1\cdot2)(0)}{1\cdot3\cdot5\cdot1\cdot3\cdot1} = 0$.

— ❦ —

The notational systems presented have their advantages but also their shortcomings:

1. The dimensional notation is unambiguous only for the $SU(2)$ group, and one must use additional "decorations" to distinguish the distinct representations that happen to have the same dimension.
2. The ket-notation is unambiguous, but requires specifying some complete collection of mutually commuting (Casimir) operators – such as J^2 and J_3 for $SU(2)$ – and their eigenvalues [☞ [488] for a list of Casimir operators].
3. The tensor/index notation is unambiguous, but the specification of the various symmetrization patterns using round parentheses and square brackets quickly becomes unwieldy and confusing.
4. The matrix notation requires ever bigger matrices.
5. Young tableaux are unambiguous and very compact, but products of arbitrary representations for some of the Lie groups may well require very complex rules [168].

Thus, in practice, one typically uses a combination of at least two notational systems, and so it is very important to know all the notational systems and how to successfully "translate" from any one into any other one of them.

A.4.3 Exercises for Section A.4

✎ ***A.4.1*** *Using the formula (A.88), compute the dimensions of the representations depicted by all the Young tableaux in the decomposition (A.81) and verify agreement for* $n = 3, 4, 5$.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

✎ ***A.4.2*** *Using the formula (A.88), compute the dimensions of the representations depicted by all the Young tableaux in the decomposition (A.82) and verify agreement for* $n = 3, 4, 5$.

✎ ***A.4.3*** *Using the formula (A.88), compute the dimensions of the representations depicted by all the Young tableaux in the decomposition (A.83) and verify agreement for* $n = 3, 4, 5$.

A.5 Orthogonal and *Spin* groups

We have already encountered the rotation group $SO(3)$, and the Lorentz group $SO(1,3)$. In the general case, the group $SO(p,q)$ is the group of linear transformations of real $(p+q)$-dimensional vectors $(x^1, \dots, x^{p+q})$, which preserve the bilinear scalar product [581, 260, 334]:

$$(\mathrm{x}\cdot\mathrm{y})_{p,q} := x^1y^1 + \cdots x^py^p - x^{p+1}y^{p+1} - \cdots x^{p+q}y^{p+q}. \tag{A.91}$$

This definition is equivalent to the statement that elements of the $SO(p,q)$ group may be represented as $(p+q)\times(p+q)$ matrices $\mathbf{L}$, which satisfy the requirement of the generalized orthogonality

$$\mathbf{L}^T\boldsymbol{\eta}_{(p,q)}\mathbf{L} \stackrel{!}{=} \boldsymbol{\eta}_{(p,q)} \qquad \Leftrightarrow \qquad \boldsymbol{\eta}_{(p,q)}\mathbf{L}^T\boldsymbol{\eta}_{(p,q)} \stackrel{!}{=} \mathbf{L}^{-1}, \tag{A.92}$$

where $\boldsymbol{\eta}_{(p,q)}$ is the diagonal matrix with the first p diagonal elements equal to $+1$, and the remaining q elements equal to -1. In the usual case $(p,q) = (1,3)$ and $\boldsymbol{\eta} := \boldsymbol{\eta}_{(1,3)}$.

A.5.1 Spinors

Just as Dirac constructed the spinorial representation $\{\hat{\mathrm{e}}_a\Psi^a\}$, starting from the 4-vector $\mathrm{p} = \hat{\mathrm{e}}^\mu p_\mu$ of the Lorentz group $SO(1,3)$, this can also be done for every $SO(p,q)$, and the $SO(p,q)$ transformation of those spinors is just as two-valued. Analogously to the double covering of the $SO(3)$ group one also defines the double covering of every $SO(p,q)$ group, denoted $Spin(p,q)$. As representations of the $Spin(p,q)$ group, both tensors and spinors are single-valued functions. The algebra of the $Spin(p,q)$ group is denoted $\mathfrak{spin}(p,q)$, and it is worth knowing that $\mathfrak{spin}(p,q) = \mathfrak{spin}(p+q,0) = \mathfrak{spin}(p+q)$. In other words, for a fixed $p+q$, different $Spin(p,q)$ groups differ only in the "finite part" [☞ Definition A.1 on p. 454, and Comment A.1 on p. 454] and their algebras are identical. For $p+q \leqslant 6$, there exist additional identities among algebras:

$$\mathfrak{spin}(3) = \mathfrak{su}(2), \quad \mathfrak{spin}(4) = \mathfrak{su}(2)\oplus\mathfrak{su}(2), \quad \mathfrak{spin}(5) = \mathfrak{sp}(4), \quad \mathfrak{spin}(6) = \mathfrak{su}(4), \tag{A.93}$$

Table A.4 Some low-dimensional ($p+q \leqslant 6$) spin groups; $Spin(p,q) = Spin(q,p)$ [☞ Section A.1.4]

$Spin(1) \cong O(1) \cong \mathbb{Z}_2$	$Spin(2,2) \cong SU(1,1)\times SU(1,1)$
$Spin(2) \cong U(1) \cong SO(2)$	$Spin(5) \cong Sp(2)$
$Spin(1,1) \cong GL(1;\mathbb{R})$	$Spin(1,4) \cong Sp(1,1)$
$Spin(3) \cong SU(2) \cong Sp(1) \cong SL(1;\mathbb{H})$	$Spin(2,3) \cong Sp(2;\mathbb{R})$
$Spin(1,2) \cong SU(1,1)$	$Spin(6) \cong SU(4)$
$Spin(4) \cong SU(2)\times SU(2)$	$Spin(1,5) \cong SL(2;\mathbb{H})$
$Spin(1,3) \cong SL(2;\mathbb{C}) \cong Spin(3;\mathbb{C})$	$Spin(2,4) \cong SU(2,2)$
$SO^{\uparrow}(1,3) \cong SO(3;\mathbb{C})$	$Spin(3,3) \cong SL(4;\mathbb{R})$

$Spin(p,q)$ groups for $p+q > 6$ are not isomorphic to other Lie groups.

which imply identities between the corresponding groups. Spin groups are defined as double coverings of orthogonal groups, i.e., the general relation

$$SO(p,q) = Spin(p,q)/\mathbb{Z}_2. \tag{A.94}$$

For physics applications, the practical meaning of this relation is that the multiplicative group $\mathbb{Z}_2 = \{\mathbb{1}, -\mathbb{1}\}$ is a subgroup of $Spin(p,q)$. Tensorial representations do not transform under this $\mathbb{Z}_2$-action, while spinorial ones change their sign under the action of $-\mathbb{1} \in \mathbb{Z}_2$. This sign equals $(-1)^{\mathrm{F}}$, where F is the so-called "fermion number" defined in the text leading to equations (10.44): $\mathrm{F} = 0$ for bosons and $\mathrm{F} = 1$ for fermions. Thus, spinorial representations of the $Spin(p,q)$ group are double-valued with respect to the $SO(p,q)$-action, and are not "true" functions; tensorial representations are single-valued under both the $Spin(p,q)$- as well as the $SO(p,q)$-action. May it suffice here to quote without proof [565, 258, 581, 256, 80, 260, 333, 447]:

> ***Theorem A.4*** *If for two groups, G_1 and G_2, it is true that $G_1 = G_2/H$, then $H \subset G_2$ is a subgroup of G_2, and elements of G_1 are obtained by identifying those elements from G_2 that differ only by the action of the subgroup $H \subset G_2$. Besides, the representations of G_1 are H-invariant representations of G_2.*

The relation $SO(p,q) = Spin(p,q)/\mathbb{Z}_2$ then implies that the $SO(p,q)$ representations are $\mathbb{Z}_2$-invariant $Spin(p,q)$ representations – and those are the tensors, the fermion number of which is $\mathrm{F} = 0$.

Spinors are, however, the $Spin(p,q)$ representations that are not invariant with respect to the action of the subgroup $\mathbb{Z}_2 \subset Spin(p,q)$ – the spinors' fermion number is $\mathrm{F} = 1$ and they change their sign pod under the action of the nontrivial $\mathbb{Z}_2$ element. Let $g_+, g_- \in Spin(p,q)$ be group elements that differ only by the $\mathbb{Z}_2$-action, so g_+ does not change its sign while g_- does. Since the relation $SO(p,q) = Spin(p,q)/\mathbb{Z}_2$ implies that the group elements $g \in SO(p,q)$ are obtained by identifying $g := [g_+ \simeq g_-]$, it is clear that the $SO(p,q)$-action upon spinors is *double-valued*.

A.5.2 $Spin(1,3)$

In relativistic physics, what is physically relevant is not the Euclidean length in spacetime, but the *interval*, of the form $\sqrt{(x^0)^2 - (x^1)^2 - \cdots}$. Thus, for relativistic physics purposes, we are most often interested in Lorentz groups $SO(1,n)$ and their double coverings, $Spin(1,n)$, where n is the number of spatial dimensions. The algebras of these groups are the same as for their Euclidean counterparts, so the identities (A.93) may be used, but it is important to keep in mind that the group $Spin(1,n)$ differ from $Spin(1{+}n)$; see Table A.4 on p. 482.

From Table A.4 on p. 482, we have that

$$Spin(1,3) = SL(2,\mathbb{C}), \tag{A.95}$$

where $SL(2,\mathbb{C})$ denotes the group of complex 2×2 matrices of unit determinant. This group is generated by

$$\boldsymbol{\tau}_j := \tfrac{1}{2}\boldsymbol{\sigma}^j \quad \text{and} \quad \widetilde{\boldsymbol{\tau}}_j := \tfrac{i}{2}\boldsymbol{\sigma}^j, \qquad j = 1,2,3, \tag{A.96}$$

the nonzero commutation relations of which are

$$[\boldsymbol{\tau}_j, \boldsymbol{\tau}_k] = i\varepsilon_{jk}{}^m\,\boldsymbol{\tau}_m, \qquad [\boldsymbol{\tau}_j, \widetilde{\boldsymbol{\tau}}_k] = i\varepsilon_{jk}{}^m\,\widetilde{\boldsymbol{\tau}}_m, \qquad [\widetilde{\boldsymbol{\tau}}_j, \widetilde{\boldsymbol{\tau}}_k] = -i\varepsilon_{jk}{}^m\,\boldsymbol{\tau}_m. \tag{A.97}$$

On the other hand, starting from relation (5.45):

$$[\,\boldsymbol{\gamma}^{\mu\nu}\,,\,\boldsymbol{\gamma}^{\rho\sigma}\,] = \eta^{\mu\rho}\boldsymbol{\gamma}^{\nu\sigma} - \eta^{\mu\sigma}\boldsymbol{\gamma}^{\nu\rho} + \eta^{\nu\sigma}\boldsymbol{\gamma}^{\mu\rho} - \eta^{\nu\rho}\boldsymbol{\gamma}^{\mu\sigma}. \tag{A.98}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

This shows that $J_j := \frac{1}{2i}\varepsilon_{jkl}\boldsymbol{\gamma}^{kl}$ with $j,k,l = 1,2,3$ satisfy the $\mathfrak{su}(2) = \mathfrak{so}(3)$ subalgebra:

$$\begin{aligned}\big[\,J_1\,,\,J_2\,\big] &= \big[\,(-i\boldsymbol{\gamma}^{23})\,,\,(-i\boldsymbol{\gamma}^{31})\,\big] = -\eta^{23}\boldsymbol{\gamma}^{31} + \eta^{21}\boldsymbol{\gamma}^{33} - \eta^{31}\boldsymbol{\gamma}^{23} + \eta^{33}\boldsymbol{\gamma}^{21} = (-1)(-iJ_3)\\ &= iJ_3,\end{aligned} \tag{A.99}$$

and so forth, for the remaining two permutations, $[J_2, J_3]$ and $[J_3, J_1]$. Denote the remaining elements $K_j := i\boldsymbol{\gamma}^{0j}$, and find

$$\begin{aligned}\big[\,K_1\,,\,K_2\,\big] &= \big[\,i\boldsymbol{\gamma}^{01}\,,\,i\boldsymbol{\gamma}^{02}\,\big] = -\eta^{00}\boldsymbol{\gamma}^{12} + \eta^{02}\boldsymbol{\gamma}^{10} - \eta^{12}\boldsymbol{\gamma}^{00} + \eta^{10}\boldsymbol{\gamma}^{02} = -(+1)(+iJ_3)\\ &= -iJ_3,\end{aligned} \tag{A.100}$$

and so forth, for the remaining two permutations, $[K_2, K_3]$ and $[K_3, K_1]$. Finally, the mixed commutators yield

$$\big[\,J_1\,,\,K_1\,\big] = \big[\,(-i\boldsymbol{\gamma}^{23})\,,\,i\boldsymbol{\gamma}^{01}\,\big] = \eta^{20}\boldsymbol{\gamma}^{31} - \eta^{21}\boldsymbol{\gamma}^{30} + \eta^{31}\boldsymbol{\gamma}^{20} - \eta^{30}\boldsymbol{\gamma}^{21} = 0, \tag{A.101}$$

$$\begin{aligned}\big[\,J_1\,,\,K_2\,\big] &= \big[\,(-i\boldsymbol{\gamma}^{23})\,,\,i\boldsymbol{\gamma}^{02}\,\big] = +\eta^{20}\boldsymbol{\gamma}^{32} - \eta^{22}\boldsymbol{\gamma}^{30} + \eta^{32}\boldsymbol{\gamma}^{20} - \eta^{30}\boldsymbol{\gamma}^{22} = -(-1)(iK_3)\\ &= iK_3,\end{aligned} \tag{A.102}$$

and so forth, for the remaining two permutations, $[K_2, K_3]$ and $[K_3, K_1]$. We thus have the general structure of commutators:

$$\big[\,J_j\,,\,J_k\,\big] = i\varepsilon_{jk}{}^{m}J_m, \qquad \big[\,J_j\,,\,K_k\,\big] = i\varepsilon_{jk}{}^{m}K_m, \qquad \big[\,K_j\,,\,K_k\,\big] = -i\varepsilon_{jk}{}^{m}J_m, \tag{A.103}$$

which are identical in form to the relations (A.97). This shows that the groups $SL(2,\mathbb{C})$ and $Spin(1,3)$, and thus also $SO(1,3) \cong Spin(1,3)/\mathbb{Z}_2$, have identical algebras.

Finally, define

$$M_j := \tfrac{1}{2}J_j + \tfrac{i}{2}K_j \qquad \text{and} \qquad \overline{M}_j := \tfrac{1}{2}J_j - \tfrac{i}{2}K_j, \tag{A.104}$$

and find

$$\big[\,M_j\,,\,M_k\,\big] = i\varepsilon_{jk}{}^{m}M_m, \qquad \big[\,\overline{M}_j\,,\,\overline{M}_k\,\big] = i\varepsilon_{jk}{}^{m}\overline{M}_m, \qquad \big[\,M_j\,,\,\overline{M}_k\,\big] = 0, \tag{A.105}$$

which demonstrates that

$$\begin{aligned}\mathfrak{alg}\big(\,SL(2;\mathbb{C})\big) &= \mathfrak{alg}\big(\,Spin(1,3)\big) = \mathfrak{alg}\big(\,Spin(3;\mathbb{C})\big) = \mathfrak{alg}\big(\,SO(1,3)\big)\\ &= \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R.\end{aligned} \tag{A.106}$$

In the physics literature one sometimes comes across the statement that the Lorentz *group* is isomorphic (or even equals) the product $SU(2)_L \times SU(2)_R$, which is false. Luckily, the precise details of the Lorentz *group* $Spin(1,3)$ and its precise relationship with the groups $SU(2)_L$ and $SU(2)_R$ generated by the operators M_j and $\overline{M}_j$ are usually not relevant, and the relation (A.5.2) for the corresponding algebras suffices.

Note that the discrete operations P and T generate the "finite part" of the $O(1,3)$ group, the group of real 4×4-matrix transformations of spacetime 4-vectors that preserve the relativistic interval. The action of the P and T transformations may then also be represented in the form of 4×4-matrices:[18]

$$P = \begin{bmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix} \qquad \text{and} \qquad T = \begin{bmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}, \qquad \text{so} \qquad PT = -\mathbb{1}. \tag{A.107}$$

[18] Caution: the matrix representation of the operations P and T evidently describes linear operations. However, in quantum theory the operation T is *anti-linear* and its action cannot be represented this way.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

The definition of the $O(1,3)$ group does not include the requirement of a unit determinant, but orthogonality implies that the determinant of $O(1,3)$-matrices equals ± 1. Elements with the determinant -1 do not form a group, as they exclude the identity element, while the elements with determinant $+1$ do form the $SO(1,3)$ group, which then is evidently a subgroup of $O(1,3)$.

The physical meaning of Lorentz transformations requires that the direction of the flow of time remains unchanged. Such transformations form a subgroup of $SO(1,3)$, which is called the orthochronous Lorentz group,[19] denoted $L^{\uparrow} \equiv SO^{\uparrow}(1,3)$. It may be shown that this is a connected group, i.e., every element of the orthochronous Lorentz group may be continuously "shrunk" to the identity element: Every Lorentz transformation may be factorized into the product of three rotations and three Lorentz boosts, akin to the well-known factorization of every rotation into three Euler angle rotations. Each of those six parameters, three angles and three components of velocity, may be continuously shrunk to 0, whereby every orthochronous Lorentz transformation may be continuously shrunk to $\mathbb{1}$.

Denote by $TL^{\uparrow}$ the collection of all products of elements from $L^{\uparrow}$ with the element T; since $L^{\uparrow}$ is continuously connected, so is $TL^{\uparrow}$. The analogy holds for $PL^{\uparrow}$ and for $PTL^{\uparrow}$. It should be evident that the $TL^{\uparrow}$, $PL^{\uparrow}$ and $PTL^{\uparrow}$ components cannot be continuously turned into $\mathbb{1}$, nor can an element of one of these three components be continuously turned into an element of another component. It then follows that the $O(1,3)$ group is a disconnected union of four components: $L^{\uparrow}$, $TL^{\uparrow}$, $PL^{\uparrow}$ and $PTL^{\uparrow}$, and that the disconnected unions $L^{\uparrow}$ and $PTL^{\uparrow}$ form a subgroup $SO(1,3) \subset O(1,3)$.

A.5.3 The Poincaré algebra and group in 1+3-dimensional spacetime

Transformations of the tangent space of 1+3-dimensional spacetime are linear transformations of the space $\mathbb{R}^{1,3}$, of the form

$$x^\mu \to y^\mu = \mathrm{L}^\mu{}_\nu\, x^\nu + \xi^\mu, \tag{A.108}$$

where the matrix $\mathbf{L} = [\mathrm{L}^\mu{}_\nu]$ provides the Lorentz transformations of 4-vectors in (flat) spacetime, and the 4-vector ξ^μ parametrizes translations in spacetime. These transformations have an *induced* action of functions of spacetime, by means of the differential operators

$$x^\mu \to x^\mu + \xi^\mu \quad \Rightarrow \quad f(\mathrm{x}) \to f(\mathrm{x}+\xi) = \exp\left\{\xi^\mu \partial_\mu\right\} f(\mathrm{x}); \tag{A.109}$$

$$x^\mu \to \mathrm{L}^\mu{}_\nu\, x^\nu \quad \Rightarrow \quad f(\mathrm{x}) \to f(\mathbf{L}\mathrm{x}) = \exp\left\{\lambda_\mu{}^\nu \mathrm{L}^\mu{}_\nu\right\} f(\mathrm{x}). \tag{A.110}$$

The translation generators are then differential operators ∂_μ, and the Lorentz transformation generators are

$$\mathrm{L}^\mu{}_\nu = x^\mu \partial_\nu - \eta^{\mu\rho}\eta_{\nu\sigma} x^\sigma \partial_\rho, \tag{A.111a}$$

so that:

boost
$$\mathrm{L}^0{}_i = x^0\partial_i - \eta^{00}\eta_{ij}x^j\partial_0 = ct\tfrac{\partial}{\partial x^i} + \delta_{ij}x^j \tfrac{1}{c}\tfrac{\partial}{\partial t}, \tag{A.111b}$$

$$\mathrm{L}^i{}_0 = x^i\partial_0 - \eta^{ij}\eta_{00}x^0\partial_j = x^i \tfrac{1}{c}\tfrac{\partial}{\partial t} + \delta^{ij} ct \tfrac{\partial}{\partial x^j} = \delta^{ij}\big(ct\tfrac{\partial}{\partial x^j} + \delta_{jk}x^k \tfrac{1}{c}\tfrac{\partial}{\partial t}\big) = \delta^{ij}\,\mathrm{L}^0{}_j, \tag{A.111c}$$

rot.
$$\mathrm{L}^i{}_j = x^i\partial_j - \eta^{ik}\eta_{j\ell}x^\ell\partial_k = \eta^{ik}\big(\eta_{kn}x^n\tfrac{\partial}{\partial x^j} - \eta_{jn}x^n\tfrac{\partial}{\partial x^k}\big)_i = \varepsilon^i{}_{jk}\,\varepsilon^{k\ell}{}_n\, x^n \tfrac{\partial}{\partial x^\ell}. \tag{A.111d}$$

[19] The nomenclature here is not quite standard: some Authors call the full $O(1,3)$ group the Lorentz group while others reserve this name only for the orthochronous component, $SO^{\uparrow}(1,3)$, of the $SO(1,3)$ group.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

For example,

$$\begin{aligned}\exp\{\xi^\mu\partial_\mu\}f(\mathrm{x}) &= \sum_{k=0}^{\infty}\frac{1}{k!}\Big(\Big[\prod_{i=1}^{k}\xi^{\mu_i}\frac{\partial}{\partial y^{\mu_i}}\Big]f(\mathrm{y})\Big)_{\mathrm{y}\to\mathrm{x}}\\ &= \frac{1}{0!}f(\mathrm{x})+\frac{1}{1!}\xi^\mu\Big(\frac{\partial f(\mathrm{y})}{\partial y^\mu}\Big)_{\mathrm{y}\to\mathrm{x}}+\frac{1}{2!}\xi^\mu\xi^\nu\Big(\frac{\partial^2 f(\mathrm{y})}{\partial y^\mu\partial y^\nu}\Big)_{\mathrm{y}\to\mathrm{x}}+\cdots\\ &= f(\mathrm{x}+\xi).\end{aligned} \tag{A.112}$$

Translations in 1+3-dimensional space $\mathbb{R}^{1,3}$ commute and are parametrized by the 4-vector $\xi^\mu \in \mathbb{R}^{1,3}$. As the operators $\frac{\partial}{\partial x^\mu}$ also span the vector space $\mathbb{R}^{1,3}$, we may write that $\mathfrak{tr}(\mathbb{R}^{1,3}_{\mathrm{x}}) \cong \mathbb{R}^{1,3}_{\partial}$, where $\mathfrak{tr}(\mathscr{X})$ denotes "the algebra of translations in space $\mathscr{X}$." It is not hard to verify that

$$[\lambda{\cdot}\mathrm{L}, \lambda'{\cdot}\mathrm{L}] = \lambda''{\cdot}\mathrm{L}, \qquad [\lambda{\cdot}\mathrm{L}, \xi{\cdot}\partial] = \xi'{\cdot}\partial, \qquad [\xi{\cdot}\partial, \xi'{\cdot}\partial] = 0, \tag{A.113a}$$

$$\lambda{\cdot}\mathrm{L} := \lambda_\mu{}^\nu \mathrm{L}^\mu{}_\nu, \qquad \xi{\cdot}\partial := \xi^\mu\partial_\mu, \tag{A.113b}$$

so that the Poincaré algebra is $\mathfrak{po}(1,3) = \mathfrak{spin}(1,3) \dotplus \mathfrak{tr}(\mathbb{R}^{1,3})$, and the Poincaré group is $Po(1,3) = Spin(1,3) \ltimes \mathbb{R}^{1,3}$, where the asymmetric binary symbol $\dotplus$ ($\ltimes$) denotes the semidirect sum (product) and recalls the fact that the left-hand summand (factor) acts upon the right-hand one [☞ the lexicon entry, in Appendix B.1].

A.5.4 Exercises for Section A.5

✎ ***A.5.1*** *Verify equations (A.103) by explicit computation, using however only the definitions (5.45).*

✎ ***A.5.2*** *Verify equations (A.105), using the definitions (A.104) and the previous results.*

✎ ***A.5.3*** *Verify equations (A.113) by explicit computation, using however only the definitions (A.111).*

✎ ***A.5.4*** *Using the definitions (A.111) and your results in the above problems, reconstruct the differential operator representation of the operators $\overline{M}_j$ and $\overline{M}_k$.*

A.6 Spinors and Dirac γ-matrices

$SO(p,q)$ denotes the group of homogeneous and linear transformations of (p,q)-vectors $\vec{v}$ that preserve the bilinear product

$$\vec{v}{\cdot}\vec{u} := \sum_{i=1}^{p} v_i u_i - \sum_{i=p+1}^{p+q} v_i u_i \stackrel{!}{=} \sum_{i=1}^{p+q} (\mathbb{M}_i{}^k u_k){\cdot}\eta^{ij}(\mathbb{M}_j{}^\ell u_\ell), \tag{A.114}$$

where the $(p{+}q)\times(p{+}q)$ matrices $\mathbb{M}$ have a unit determinant and where

$$[\eta^{ij}] = \boldsymbol{\eta}^{(p,q)} = \mathrm{diag}(\underbrace{+1,\dots,+1}_{p},\underbrace{-1,\dots,-1}_{q}). \tag{A.115}$$

The vectors $\vec{v}$ form the defining vector space V_v. One also writes $SO(p,q) = SO(V_v;\boldsymbol{\eta}^{(p,q)})$, where the latter notation quite literally stands for "the group of unimodular orthogonal transformations of the vector space V_v that preserve the bilinear product obtained using the matrix $\boldsymbol{\eta}^{(p,q)}$."

Every (unimodular orthogonal) group $SO(p,q)$ has a double covering (that is also a group), denoted $Spin(p,q)$ [333], the single-valued spinorial representations of which are double-valued representations of the $SO(p,q)$ group. Every $Spin(p,q)$ group has the (Dirac) spinor representation V_Ψ as well as its formally dual representation $(V_\Psi)^* = V_{\overline{\Psi}}$, for which the relation

$$V_\Psi \otimes V_{\overline{\Psi}} \supset V_v, \tag{A.116}$$

holds, where $V_v\ (p,q)$ is the defining vector representation of $SO(p,q)$. For any chosen bases,

$$\hat{\mathrm{e}}_A \in V_\Psi, \qquad \hat{\mathrm{e}}^A \in V_{\overline{\Psi}}, \qquad \text{and} \qquad \hat{\mathrm{e}}^\mu \in V_v, \tag{A.117}$$

and the Dirac $\boldsymbol{\gamma}$-matrices are arrays of the coefficients in the projection (A.116):

$$\hat{\mathrm{e}}_A (\boldsymbol{\gamma}^\mu)^A{}_B\, \hat{\mathrm{e}}^B = \hat{\mathrm{e}}^\mu. \tag{A.118}$$

A.6.1 Dirac matrices in $(3+1)$ dimensional spacetime

The elements $\boldsymbol{\gamma}^\mu$, $\mu = 0,1,2,3$, which satisfy

$$\{\boldsymbol{\gamma}^\mu, \boldsymbol{\gamma}^\nu\} = 2\eta^{\mu\nu}, \qquad \text{with} \qquad [\eta^{\mu\nu}] = \mathrm{diag}(+1,-1,-1,-1), \tag{A.119}$$

form the Clifford algebra $\mathfrak{Cl}(1,3)$. Following Feynman, one defines

$$\not{p} := \boldsymbol{\gamma}^\mu p_\mu \qquad \text{for each 4-vector p.} \tag{A.120}$$

This implies the following definitions and results:

$$\widehat{\boldsymbol{\gamma}} := i\boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^2\boldsymbol{\gamma}^3 := \tfrac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma, \qquad \{\widehat{\boldsymbol{\gamma}}, \boldsymbol{\gamma}^\mu\} = 0, \qquad (\widehat{\boldsymbol{\gamma}})^2 = \mathbb{1}; \tag{A.121a}$$

$$\boldsymbol{\gamma}_\pm := \tfrac{1}{2}[\mathbb{1} \pm \widehat{\boldsymbol{\gamma}}], \qquad [\boldsymbol{\gamma}_+, \boldsymbol{\gamma}_-] = 0, \qquad \boldsymbol{\gamma}_+ + \boldsymbol{\gamma}_- = \mathbb{1}, \qquad (\boldsymbol{\gamma}_\pm)^2 = \boldsymbol{\gamma}_\pm, \tag{A.121b}$$

$$\boldsymbol{\gamma}^{\mu\nu} := \frac{i}{4}[\boldsymbol{\gamma}^\mu, \boldsymbol{\gamma}^\nu], \qquad [\boldsymbol{\gamma}^{\mu\nu}, \boldsymbol{\gamma}^{\rho\sigma}] = \eta^{\mu\rho}\boldsymbol{\gamma}^{\nu\sigma} - \eta^{\mu\sigma}\boldsymbol{\gamma}^{\nu\rho} + \eta^{\nu\sigma}\boldsymbol{\gamma}^{\mu\rho} - \eta^{\nu\rho}\boldsymbol{\gamma}^{\mu\sigma}. \tag{A.121c}$$

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}_\mu = 4\,\mathbb{1}, \qquad \boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}_\mu = 4\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho, \tag{A.122a}$$

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}_\mu = -2\boldsymbol{\gamma}^\nu, \qquad \boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma\boldsymbol{\gamma}_\mu = -2\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma, \tag{A.122b}$$

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho = \eta^{\mu\nu}\boldsymbol{\gamma}^\rho - \eta^{\mu\rho}\boldsymbol{\gamma}^\nu + \eta^{\nu\rho}\boldsymbol{\gamma}^\mu + i\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{\gamma}_\sigma\widehat{\boldsymbol{\gamma}}\ . \tag{A.122c}$$

Theorem A.5 *Owing to the relations (A.119), (A.121a) and (A.122c), it follows that every $\boldsymbol{\gamma}$-matrix polynomial may be reduced to the quadratic polynomial*

$$C_0\mathbb{1} + C_\mu\boldsymbol{\gamma}^\mu + \tfrac{1}{2}C_{\mu\nu}\boldsymbol{\gamma}^{\mu\nu} + \widehat{C}_\mu\boldsymbol{\gamma}^\mu\widehat{\boldsymbol{\gamma}} + \widehat{C}_0\widehat{\boldsymbol{\gamma}}. \tag{A.123}$$

That is, the basis

$$\mathbb{1},\ \boldsymbol{\gamma}^\mu,\ \boldsymbol{\gamma}^{\mu\nu},\ \ \boldsymbol{\gamma}^\mu\widehat{\boldsymbol{\gamma}},\ \widehat{\boldsymbol{\gamma}} \tag{A.124}$$

for the Dirac algebra (A.119) is complete.

We also have

$$\mathrm{Tr}[\boldsymbol{\gamma}^\mu] = 0, \qquad \mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho] = 0, \qquad \mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma\boldsymbol{\gamma}^\lambda] = 0, \qquad \text{etc.} \tag{A.125a}$$

$$\mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu] = 4\eta^{\mu\nu}, \qquad \mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma] = 4(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}), \tag{A.125b}$$

$$\mathrm{Tr}[\widehat{\boldsymbol{\gamma}}] = 0, \qquad \mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\widehat{\boldsymbol{\gamma}}] = 0, \qquad \mathrm{Tr}[\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\rho\boldsymbol{\gamma}^\sigma\widehat{\boldsymbol{\gamma}}] = -4i\varepsilon^{\mu\nu\rho\sigma}. \tag{A.125c}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

These relations imply

$$\not{p}\not{p} = \mathrm{p}^2\,\mathbb{1}, \qquad \not{p}\not{q} = (\mathrm{p}\cdot\mathrm{q} - 2i p_\mu \gamma^{\mu\nu} q_\nu)\mathbb{1}; \tag{A.126a}$$
$$\not{p}\not{q} + \not{q}\not{p} = 2(\mathrm{p}\cdot\mathrm{q})\,\mathbb{1}, \qquad \not{p}\not{q} - \not{q}\not{p} = -4i(p_\mu \gamma^{\mu\nu} q_\nu)\mathbb{1}; \tag{A.126b}$$
$$\mathrm{Tr}[\not{p}\not{q}] = 4\mathrm{p}\cdot\mathrm{q}\,\mathbb{1}, \qquad \mathrm{Tr}[\not{p}\not{q}\not{r}\not{s}] = 4[(\mathrm{p}\cdot\mathrm{q})(\mathrm{r}\cdot\mathrm{s}) - (\mathrm{p}\cdot\mathrm{r})(\mathrm{q}\cdot\mathrm{s}) + (\mathrm{p}\cdot\mathrm{s})(\mathrm{q}\cdot\mathrm{r})]; \tag{A.126c}$$
$$\mathrm{Tr}[\not{p}] = 0 = \mathrm{Tr}[\not{p}\not{q}\not{r}], \quad \mathrm{Tr}[\widehat{\boldsymbol{\gamma}}\not{p}\not{q}\not{r}\not{s}] = 4i\varepsilon^{\mu\nu\rho\sigma} p_\mu\, q_\nu\, r_\rho\, s_\sigma; \tag{A.126d}$$
$$\boldsymbol{\gamma}^\mu \not{p}\not{q}\boldsymbol{\gamma}_\mu = 4\mathrm{p}\cdot\mathrm{q}\,\mathbb{1}, \qquad \boldsymbol{\gamma}^\mu \not{p}\boldsymbol{\gamma}_\mu = -2\not{p}, \qquad \boldsymbol{\gamma}^\mu \not{p}\not{q}\not{r}\boldsymbol{\gamma}_\mu = -2\not{r}\not{q}\not{p}. \tag{A.126e}$$

In physics applications, besides the relations (A.119) that define the Clifford algebra, one *additionally* requires the matrices $\boldsymbol{\gamma}^\mu$ to satisfy

$$(\boldsymbol{\gamma}^0)^\dagger = \boldsymbol{\gamma}^0, \quad \text{and} \quad (\boldsymbol{\gamma}^i)^\dagger = -\boldsymbol{\gamma}^i,\ i = 1,2,3, \qquad \Leftrightarrow \qquad (\boldsymbol{\gamma}^\mu)^\dagger = \boldsymbol{\gamma}^0\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^0. \tag{A.127}$$

This requirement is not an integral part of the definition and structure of Clifford algebras, which one must keep in mind when using mathematical results about Clifford algebras. The use of the algebra (A.119) in the physics literature always assumes the additional conditions (A.127) – as well as their consequences.

Corresponding to Dirac conjugation of spin-$\frac{1}{2}$ fermions (5.49), we have

$$\overline{\Psi} := \Psi^\dagger\boldsymbol{\gamma}^0 \qquad \Leftrightarrow \qquad \overline{\boldsymbol{\gamma}^\mu} := \boldsymbol{\gamma}^0(\boldsymbol{\gamma}^\mu)^\dagger\boldsymbol{\gamma}^0 \overset{(A.127)}{=} \boldsymbol{\gamma}^\mu. \tag{A.128}$$

Therefore,

$$\begin{aligned}\overline{\widehat{\boldsymbol{\gamma}}} &:= \boldsymbol{\gamma}^0(i\boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^2\boldsymbol{\gamma}^3)^\dagger\boldsymbol{\gamma}^0 = -i\boldsymbol{\gamma}^0(\boldsymbol{\gamma}^3)^\dagger\boldsymbol{\gamma}^0\boldsymbol{\gamma}^0(\boldsymbol{\gamma}^2)^\dagger\boldsymbol{\gamma}^0\boldsymbol{\gamma}^0(\boldsymbol{\gamma}^1)^\dagger\boldsymbol{\gamma}^0\boldsymbol{\gamma}^0(\boldsymbol{\gamma}^0)^\dagger\boldsymbol{\gamma}^0 = -i\overline{\boldsymbol{\gamma}^3}\,\overline{\boldsymbol{\gamma}^2}\,\overline{\boldsymbol{\gamma}^1}\,\overline{\boldsymbol{\gamma}^0}\\ &\overset{(A.127)}{=} -i\boldsymbol{\gamma}^3\boldsymbol{\gamma}^2\boldsymbol{\gamma}^1\boldsymbol{\gamma}^0 \overset{(A.119)}{=} -i\boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^2\boldsymbol{\gamma}^3 = -\widehat{\boldsymbol{\gamma}},\end{aligned} \tag{A.129}$$

and so

$$\overline{\boldsymbol{\gamma}_\pm} = \boldsymbol{\gamma}_\mp, \qquad \text{whereby} \qquad \overline{\Psi_\pm} = \overline{\Psi}_\mp. \tag{A.130}$$

Besides the Dirac basis:

$$\boldsymbol{\gamma}^0 = \begin{bmatrix}\mathbb{1} & \mathbb{0}\\ \mathbb{0} & -\mathbb{1}\end{bmatrix}, \quad \boldsymbol{\gamma}^i = \begin{bmatrix}\mathbb{0} & \boldsymbol{\sigma}^i\\ -\boldsymbol{\sigma}^i & \mathbb{0}\end{bmatrix}, \quad \widehat{\boldsymbol{\gamma}} = \begin{bmatrix}\mathbb{0} & \mathbb{1}\\ \mathbb{1} & \mathbb{0}\end{bmatrix}, \tag{A.131}$$

the most often used choices are the Weyl basis:

$$\boldsymbol{\gamma}^0 = \begin{bmatrix}\mathbb{0} & -\mathbb{1}\\ -\mathbb{1} & \mathbb{0}\end{bmatrix}, \quad \boldsymbol{\gamma}^i = \begin{bmatrix}\mathbb{0} & \boldsymbol{\sigma}^i\\ -\boldsymbol{\sigma}^i & \mathbb{0}\end{bmatrix}, \quad \widehat{\boldsymbol{\gamma}} = \begin{bmatrix}\mathbb{1} & \mathbb{0}\\ \mathbb{0} & -\mathbb{1}\end{bmatrix}, \qquad \Psi_{\text{Dirac}} = \begin{bmatrix}\Psi_+\\ \Psi_-\end{bmatrix}; \tag{A.132}$$

and the Majorana basis:

$$\begin{aligned}&\boldsymbol{\gamma}^0 = \begin{bmatrix}\mathbb{0} & \boldsymbol{\sigma}^2\\ \boldsymbol{\sigma}^2 & \mathbb{0}\end{bmatrix}, \quad \boldsymbol{\gamma}^1 = \begin{bmatrix}i\boldsymbol{\sigma}^3 & \mathbb{0}\\ \mathbb{0} & i\boldsymbol{\sigma}^3\end{bmatrix}, \quad \boldsymbol{\gamma}^2 = \begin{bmatrix}\mathbb{0} & -\boldsymbol{\sigma}^2\\ \boldsymbol{\sigma}^2 & \mathbb{0}\end{bmatrix}, \qquad \boldsymbol{\gamma}^3 = \begin{bmatrix}-i\boldsymbol{\sigma}^1 & \mathbb{0}\\ \mathbb{0} & -i\boldsymbol{\sigma}^1\end{bmatrix},\\ &\widehat{\boldsymbol{\gamma}} = \begin{bmatrix}\boldsymbol{\sigma}^2 & \mathbb{0}\\ \mathbb{0} & \boldsymbol{\sigma}^2\end{bmatrix},\end{aligned} \tag{A.133}$$

in which all components of the Dirac spinor Ψ are real, while the Dirac matrices themselves are all imaginary in the Majorana basis.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

A.6.2 Weyl's notation for spinors

The literature about supersymmetry [☞ [189, 562, 560, 129, 76, 308], to list only textbooks] is unfortunately replete with differences in notation and conventions. For consistency, the conventions of Ref. [76] are adopted herein, and the Reader is left to compare with other sources and correctly translate the notation and conventions.

Left and right spinors

The result (5.58) indicates the fact that the Dirac 4-component spinor may, in a Lorentz-invariant way, be separated into a pair of two-component spinors, $\Psi = (\Psi_+, \Psi_-)$, where $\Psi_\pm$ are defined by the projections $\boldsymbol{\gamma}_\pm$ (5.57). This separation reflects the fact that the Lorentz group in 1+3-dimensional spacetime is $Spin(1,3) \cong Spin(3;\mathbb{C}) \cong SL(2;\mathbb{C})$, and that the Lorentz algebra is

$$\mathfrak{spin}(1,3) = \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R. \tag{A.134}$$

That is, Ψ_+ transforms under the $\mathfrak{su}(2)_L$-action and is invariant under the $\mathfrak{su}(2)_R$-action, while Ψ_- transforms the other way around:

$$\Psi_+ \sim (\tfrac{1}{2}, 0), \quad \Psi_- \sim (0, \tfrac{1}{2}) \qquad \text{with respect to } \mathfrak{spin}(1,3) = \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R. \tag{A.135}$$

In physics literature one often encounters the statement "$Spin(1,3) \cong SU(2)_L \times SU(2)_R$," which does not hold for the group. For most all of physics purposes, however, the relation (A.135) suffices, which is true of the algebra; the Reader is directed to the literature [565, 258, 581, 256, 80, 260, 333, 447] for the precise details about these groups, their representations and differences.

For the Weyl spinors (5.58), one uses the 2-component notation:

$$\Psi_+ := \boldsymbol{\gamma}_+ \Psi \mapsto \psi_\alpha, \qquad \psi_\alpha \to \psi'_\alpha = M_\alpha{}^\beta \, \psi_\beta, \tag{A.136a}$$

$$\Psi_- := \boldsymbol{\gamma}_- \Psi \mapsto \overline{\chi}_{\dot\alpha}, \qquad \overline{\chi}_{\dot\alpha} \to \overline{\chi}'_{\dot\alpha} = \overline{\chi}_{\dot\beta} \, (\overline{M}^{-1})^{\dot\beta}{}_{\dot\alpha}. \tag{A.136b}$$

Here, $M_\alpha{}^\beta$ and $\overline{M}^{\dot\beta}{}_{\dot\alpha}$ are matrix elements of $SL(2;\mathbb{C})$-matrices $\mathbb{M} = \exp\{\mathfrak{m}_L\}$ with $\mathfrak{m}_L \in \mathfrak{su}(2)_L$ and $\overline{\mathbb{M}} = \exp\{\mathfrak{m}_R\}$ with $\mathfrak{m}_R \in \mathfrak{su}(2)_R$; the matrices $\mathbb{M}$ and $\overline{\mathbb{M}}$ are independent, and one refers to independent "left" and "right" action.

The spin-$\frac{1}{2}$ wave-functions ψ and χ are used to represent *fermionic* wave-functions, so that the components ψ_α and $\overline{\chi}_{\dot\alpha}$ are *anticommuting* functions.[20] Thus the Levi-Civita symbols $\varepsilon^{\alpha\beta}$ and $\varepsilon^{\dot\alpha\dot\beta}$ serve as (antisymmetric!) metric tensors for "left" and "right" Weyl spinors, ψ, χ and $\overline{\psi}, \overline{\chi}$:

$$(\psi{\cdot}\chi) := \psi_\alpha \varepsilon^{\alpha\beta} \chi_\beta = \psi_1\chi_2 - \psi_2\chi_1 = -\chi_\beta \varepsilon^{\alpha\beta} \psi_\alpha = \chi_\beta \varepsilon^{\beta\alpha} \psi_\alpha = (\chi{\cdot}\psi), \tag{A.137}$$

$$(\overline{\psi}{\cdot}\overline{\chi}) := \overline{\psi}_{\dot\alpha} \varepsilon^{\dot\alpha\dot\beta} \overline{\chi}_{\dot\beta} = \overline{\psi}_1\overline{\chi}_2 - \overline{\psi}_2\overline{\chi}_1 = -\overline{\chi}_{\dot\beta} \varepsilon^{\dot\alpha\dot\beta} \overline{\psi}_{\dot\alpha} = \overline{\chi}_{\dot\beta} \varepsilon^{\dot\beta\dot\alpha} \overline{\psi}_{\dot\alpha} = (\overline{\chi}{\cdot}\overline{\psi}), \tag{A.138}$$

where we must pay attention to detail:

$$\varepsilon^{\alpha\gamma}\varepsilon_{\beta\gamma} = \delta^\alpha_\beta, \quad \text{but} \quad \varepsilon^{\alpha\gamma}\varepsilon_{\gamma\beta} = -\delta^\alpha_\beta; \qquad \varepsilon^{\dot\alpha\dot\gamma}\varepsilon_{\dot\beta\dot\gamma} = \delta^{\dot\alpha}_{\dot\beta}, \quad \text{but} \quad \varepsilon^{\dot\alpha\dot\gamma}\varepsilon_{\dot\gamma\dot\beta} = -\delta^{\dot\alpha}_{\dot\beta}. \tag{A.139}$$

By convention, we set $\varepsilon^{12} = 1 = \varepsilon^{\dot\alpha\dot\beta}$.

[20] To be precise, every component of the field ψ_α and χ_α may be identified with a spacetime-dependent linear combination of anticommuting operators, such as b and $b^\dagger$ in Section 10.1, where creation operators act upon a vacuum state and create states with appropriate fermionic excitations.

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Products of 2-component Weyl spinors satisfy the following identities:

$$\psi_\alpha\chi_\beta = \tfrac{1}{2}\varepsilon_{\alpha\beta}(\psi\cdot\chi) - \tfrac{1}{2}\sigma^{\mu\nu}_{\alpha\beta}(\psi\,\boldsymbol{\sigma}_{\mu\nu}\,\chi), \tag{A.140}$$

$$\overline{\psi}_{\dot\alpha}\overline{\chi}_{\dot\beta} = \tfrac{1}{2}\varepsilon_{\dot\alpha\dot\beta}(\overline{\psi}\cdot\overline{\chi}) - \tfrac{1}{2}\overline{\sigma}^{\mu\nu}_{\dot\alpha\dot\beta}(\overline{\psi}\,\overline{\boldsymbol{\sigma}}_{\mu\nu}\,\overline{\chi}), \tag{A.141}$$

$$\psi^2 := \tfrac{1}{2}\varepsilon^{\alpha\beta}\psi_\alpha\psi_\beta, \qquad \psi_a\overline{\chi}_{\dot\alpha} = -\tfrac{1}{2}\sigma^\mu_{\alpha\dot\alpha}(\psi\,\boldsymbol{\sigma}_\mu\,\overline{\chi}), \tag{A.142}$$

$$(\psi_1\cdot\psi_2)(\psi_3\cdot\psi_4) = -(\psi_1\cdot\psi_3)(\psi_2\cdot\psi_4) - (\psi_1\cdot\psi_4)(\psi_2\cdot\psi_3), \tag{A.143}$$

$$(\psi_1\cdot\psi_2)(\overline{\psi}_3\cdot\overline{\psi}_4) = -\tfrac{1}{2}(\psi_1\boldsymbol{\sigma}^\mu\overline{\psi}_4)(\psi_2\boldsymbol{\sigma}_\mu\overline{\psi}_3). \tag{A.144}$$

Comment A.3 *Since fermionic wave-functions are anticommuting, they must also be nilpotent:*

$$\{\,\psi_\alpha(\mathrm{x})\,,\,\psi_\beta(\mathrm{x})\,\} = 0 \qquad \Rightarrow \qquad \big(\psi_\alpha(\mathrm{x})\big)^2 \equiv 0. \tag{A.145}$$

The notation "ψ^2" is then free for the definition:

$$\psi^2(\mathrm{x}) := \psi_1(\mathrm{x})\psi_2(\mathrm{x}) = \tfrac{1}{2}\varepsilon^{\alpha\beta}\psi_\alpha(\mathrm{x})\psi_\beta(\mathrm{x}). \tag{A.146}$$

4-Vectors and Pauli's matrices

4-vectors such as the spacetime 4-vector x transform as the $(\frac{1}{2},\frac{1}{2})$ representation of the $\mathfrak{spin}(1,3) = \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R$ algebra, i.e., of the $Spin(1,3) \cong SL(2;\mathbb{C})$ group. The $SL(2;\mathbb{C})$ group action on the 4-vector x^μ is easiest represented using Pauli matrices:

$$[\boldsymbol{\sigma}^\mu]_{\alpha\dot\alpha}: \qquad \boldsymbol{\sigma}^0 := \left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right], \quad \boldsymbol{\sigma}^1 := \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right], \quad \boldsymbol{\sigma}^2 := \left[\begin{smallmatrix}0&-i\\i&0\end{smallmatrix}\right], \quad \boldsymbol{\sigma}^3 := \left[\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right], \tag{A.147}$$

which are identified with the index notation $\sigma^\mu_{\alpha\dot\alpha}$ so that, e.g., $\sigma^2_{12} = -i$. Using $\varepsilon^{\alpha\beta}$ and $\varepsilon^{\dot\alpha\dot\beta}$ to "raise" spinor indices and $\eta_{\mu\nu}$ to "lower" the vector index, we have

$$\overline{\sigma}^{\dot\alpha\alpha}_\mu := \varepsilon^{\dot\alpha\dot\beta}\varepsilon^{\alpha\beta}\eta_{\mu\nu}\sigma^\nu_{\beta\dot\beta}: \qquad [\overline{\boldsymbol{\sigma}}_\mu] = \big([\mathbb{1}],[\boldsymbol{\sigma}^1],[\boldsymbol{\sigma}^2],[\boldsymbol{\sigma}^3]\big) = [\boldsymbol{\sigma}^\mu]. \tag{A.148}$$

That is, the matrices $\boldsymbol{\sigma}^\mu$ and $\overline{\boldsymbol{\sigma}}_\mu$ look alike. However, the matrices

$$\boldsymbol{\sigma}_\mu := \eta_{\mu\nu}\boldsymbol{\sigma}^\nu \qquad \text{and} \qquad \overline{\boldsymbol{\sigma}}^\mu := \eta^{\mu\nu}\overline{\boldsymbol{\sigma}}_\nu \tag{A.149}$$

have a differing sign: $[\overline{\boldsymbol{\sigma}}^1] = -[\boldsymbol{\sigma}^1]$, $[\overline{\boldsymbol{\sigma}}^2] = -[\boldsymbol{\sigma}^2]$, $[\overline{\boldsymbol{\sigma}}^3] = -[\boldsymbol{\sigma}^3]$, as well as $[\boldsymbol{\sigma}_1] = -[\overline{\boldsymbol{\sigma}}_1]$, $[\boldsymbol{\sigma}_2] = -[\overline{\boldsymbol{\sigma}}_2]$, $[\boldsymbol{\sigma}_3] = -[\overline{\boldsymbol{\sigma}}_3]$.

One therefore writes

$$\boldsymbol{x} := x^\mu\overline{\boldsymbol{\sigma}}_\mu, \qquad \boldsymbol{x} \to \boldsymbol{x}' = \overline{\mathbb{M}}\,\boldsymbol{x}\,\mathbb{M}^{-1}, \qquad \mathbb{M},\overline{\mathbb{M}} \in SL(2;\mathbb{C}), \tag{A.150}$$

where the matrices $\mathbb{M} = \exp\{i\omega_\mu\boldsymbol{\sigma}^\mu\}$ and $\overline{\mathbb{M}} = \exp\{i\overline{\pi}_\mu\overline{\boldsymbol{\sigma}}^\mu\}$ are independent, and represent the independent "left" and "right" action, so that

$$(\overline{\chi}'\cdot\boldsymbol{x}'\cdot\psi') = (\overline{\chi}\overline{\mathbb{M}}^{-1}\cdot\overline{\mathbb{M}}\,\boldsymbol{x}\,\mathbb{M}^{-1}\cdot\mathbb{M}\psi) = (\overline{\chi}\cdot\boldsymbol{x}\cdot\psi) \tag{A.151}$$

is an $SL(2;\mathbb{C})$-invariant. In the index notation,

$$\begin{aligned}(\overline{\chi}_{\dot\alpha}\,x^{\dot\alpha\alpha}\,\psi_\alpha) \to (\overline{\chi}'_{\dot\alpha}\,x'^{\dot\alpha\alpha}\,\psi'_\alpha) &= \overline{\chi}_{\dot\beta}\,(\overline{M}^{-1})^{\dot\beta}{}_{\dot\alpha}\,\overline{M}^{\dot\alpha}{}_{\dot\gamma}\,x^{\dot\gamma\gamma}\,(M^{-1})_\gamma{}^\alpha\,M_\alpha{}^\beta\,\psi_\beta\\ &= (\overline{\chi}_{\dot\alpha}\,x^{\dot\alpha\alpha}\,\psi_\alpha).\end{aligned} \tag{A.152}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

Finally, notice that

$$\det[\,\boldsymbol{x}\,] = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = x^\mu\, \eta_{\mu\nu}\, x^\nu \overset{(3.17)}{=} \mathbf{x}^2, \tag{A.153}$$

which is also an $SL(2;\mathbb{C})$-invariant:

$$\det[\,\boldsymbol{x}\,] \to \det[\,\boldsymbol{x}'\,] = \det[\,\overline{\mathbb{M}}\,\boldsymbol{x}\,\mathbb{M}^{-1}\,] = \det[\overline{\mathbb{M}}]\det[\,\boldsymbol{x}\,]\det[\mathbb{M}^{-1}] = \det[\,\boldsymbol{x}\,], \tag{A.154}$$

since the $SL(2;\mathbb{C})$ elements are unimodular, $\det[\overline{\mathbb{M}}] = 1 = \det[\mathbb{M}]$.

The Pauli matrices (A.147) and (A.148) satisfy the following useful identities:

$$(\boldsymbol{\sigma}_\mu\overline{\boldsymbol{\sigma}}_\nu + \boldsymbol{\sigma}_\nu\overline{\boldsymbol{\sigma}}_\mu)_\alpha{}^\beta = 2\eta_{\mu\nu}\,\delta_\alpha^\beta, \qquad (\overline{\boldsymbol{\sigma}}_\mu\boldsymbol{\sigma}_\nu + \overline{\boldsymbol{\sigma}}_\nu\boldsymbol{\sigma}_\mu)^{\dot\alpha}{}_{\dot\beta} = 2\eta_{\mu\nu}\,\delta^{\dot\alpha}_{\dot\beta}; \tag{A.155}$$

$$\mathrm{Tr}\,\big[\boldsymbol{\sigma}_\mu\overline{\boldsymbol{\sigma}}_\nu\big] = \mathrm{Tr}\,\big[\overline{\boldsymbol{\sigma}}_\mu\boldsymbol{\sigma}_\nu\big] = 2\eta_{\mu\nu}, \qquad \sigma^\mu_{\alpha\dot\alpha}\,\overline{\sigma}_\mu^{\dot\beta\beta} = 2\delta_\alpha^\beta\,\delta_{\dot\alpha}^{\dot\beta}, \tag{A.156}$$

and are suitable for the conversion of $Spin(1,3)$-tensors into (bi)spinor expressions:

$$V_{\alpha\dot\alpha} := \sigma^\mu_{\alpha\dot\alpha}V_\mu \quad\Leftrightarrow\quad V_\mu = \tfrac{1}{2}\overline{\sigma}_\mu^{\dot\alpha\alpha}V_{\alpha\dot\alpha}. \tag{A.157}$$

It is convenient to also define the matrices

$$(\boldsymbol{\sigma}_{\mu\nu})_\alpha{}^\beta := \tfrac{1}{4}\big(\sigma_{\mu\,\alpha\dot\alpha}\overline{\sigma}_\nu{}^{\dot\alpha\beta} - \sigma_{\nu\,\alpha\dot\alpha}\overline{\sigma}_\mu{}^{\dot\alpha\beta}\big), \qquad (\overline{\boldsymbol{\sigma}}_{\mu\nu})^{\dot\alpha}{}_{\dot\beta} := \tfrac{1}{4}\big(\overline{\sigma}_\mu{}^{\dot\alpha\alpha}\sigma_{\nu\,\alpha\dot\beta} - \overline{\sigma}_\nu{}^{\dot\alpha\alpha}\sigma_{\mu\,\alpha\dot\beta}\big), \tag{A.158a}$$

which, $\boldsymbol{\sigma}_{\mu\nu}$ and $\overline{\boldsymbol{\sigma}}_{\mu\nu}$ independently, close the $\mathfrak{spin}(1,3)$ algebra (A.121c), and for which

$$(\boldsymbol{\sigma}_{\mu\nu})_{\alpha\beta} := (\boldsymbol{\sigma}_{\mu\nu})_\alpha{}^\gamma\varepsilon_{\beta\gamma} \quad\text{and}\quad (\boldsymbol{\sigma}_{\mu\nu})^{\alpha\beta} := \varepsilon^{\alpha\gamma}(\boldsymbol{\sigma}_{\mu\nu})_\gamma{}^\beta, \tag{A.158b}$$

$$(\overline{\boldsymbol{\sigma}}_{\mu\nu})^{\dot\alpha\dot\beta} := (\overline{\boldsymbol{\sigma}}_{\mu\nu})^{\dot\alpha}{}_{\dot\gamma}\varepsilon^{\dot\beta\dot\gamma} \quad\text{and}\quad (\overline{\boldsymbol{\sigma}}_{\mu\nu})_{\dot\alpha\dot\beta} := \varepsilon_{\dot\alpha\dot\gamma}(\overline{\boldsymbol{\sigma}}_{\mu\nu})^{\dot\gamma}{}_{\dot\beta}. \tag{A.158c}$$

For these matrices (with $\varepsilon_{0123} = 1$), it is true that

$$(\boldsymbol{\sigma}_{\mu\nu})_\alpha{}^\beta(\boldsymbol{\sigma}_{\rho\sigma})_\beta{}^\alpha = \tfrac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) + \tfrac{i}{2}\varepsilon_{\mu\nu\rho\sigma}, \tag{A.159}$$

$$(\overline{\boldsymbol{\sigma}}_{\mu\nu})^{\dot\alpha}{}_{\dot\beta}(\overline{\boldsymbol{\sigma}}_{\rho\sigma})^{\dot\beta}{}_{\dot\alpha} = \tfrac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) - \tfrac{i}{2}\varepsilon_{\mu\nu\rho\sigma}. \tag{A.160}$$

Super-derivatives

In supersymmetry research, the so-called "super-derivatives"

$$D_\alpha := \partial_\alpha - i\sigma^\mu_{\alpha\dot\alpha}\overline{\theta}^{\dot\alpha}\partial_\mu \quad\text{and}\quad \overline{D}_{\dot\alpha} := \overline{\partial}_{\dot\alpha} - i\sigma^\mu_{\alpha\dot\alpha}\theta^\alpha\partial_\mu \tag{10.68$'$}$$

are of special importance. They anticommute with the generators of supersymmetry, $Q_\alpha, \overline{Q}_{\dot\alpha}$, and so commute with the operator of the supersymmetry transformation:

$$D_\alpha U_{\epsilon,\overline{\epsilon}} = U_{\epsilon,\overline{\epsilon}}D_\alpha \quad\text{and}\quad \overline{D}_{\dot\alpha}U_{\epsilon,\overline{\epsilon}} = U_{\epsilon,\overline{\epsilon}}\overline{D}_{\dot\alpha}, \qquad U_{\epsilon,\overline{\epsilon}} := \exp\{\,-i(\epsilon^\alpha Q_\alpha + \overline{\epsilon}^{\dot\alpha}\overline{Q}_{\dot\alpha})\,\}. \tag{A.161}$$

The operators $D_\alpha, \overline{D}_{\dot\alpha}$ are then, in fact, literally *invariant* with respect to the supersymmetry action, but their name, "(super)covariant," stuck in the literature; herein, the shorter and more precise term "super-derivative" is used.

The basic property of the super-derivatives,

$$\{\,D_\alpha\,,\,\overline{D}_{\dot\alpha}\,\} = 2\,\hbar^{-1}\sigma^\mu_{\alpha\dot\alpha}P_\mu = -2i\,\sigma^\mu_{\alpha\dot\alpha}\partial_\mu, \tag{10.69$'$}$$

is sometimes called *super-commutativity* and permits simplifying higher-order super-derivatives:

$$D_\alpha D_\beta = \tfrac{1}{2}\epsilon_{\alpha\beta}D^2, \qquad \overline{D}_{\dot\alpha}\overline{D}_{\dot\beta} = \tfrac{1}{2}\epsilon_{\dot\alpha\dot\beta}\overline{D}^2; \tag{A.162}$$

$$D_\alpha D_\beta D_\gamma = 0, \qquad \overline{D}_{\dot\alpha}\overline{D}_{\dot\beta}\overline{D}_{\dot\gamma} = 0; \tag{A.163}$$

$$[D^2, \overline{D}_{\dot\alpha}] = 4i\sigma^\mu_{\alpha\dot\alpha}\varepsilon^{\alpha\beta}\partial_\mu D_\beta, \qquad [\overline{D}^2, D_\alpha] = 4i\sigma^\mu_{\alpha\dot\alpha}\varepsilon^{\dot\alpha\dot\beta}\partial_\mu\overline{D}_{\dot\beta}; \tag{A.164}$$

$$D^2\overline{D}^2 + \overline{D}^2D^2 - 2\varepsilon^{\dot\alpha\dot\beta}\overline{D}_{\dot\alpha}D^2\overline{D}_{\dot\beta} = -16\Box, \quad \Box := \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{A.165}$$

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press

A.6.3 Exercises for Section A.6

✎ **A.6.1** *Prove the relations (A.121) using only the anticommutation relations (A.119).*

✎ **A.6.2** *Prove the relations (A.122) using only the anticommutation relations (A.119).*

✎ **A.6.3** *Prove Theorem A.5 using only the anticommutation relations (A.119).*

✎ **A.6.4** *Prove Theorem A.5 using the Cayley–Hamilton theorem.*

✎ **A.6.5** *Prove the relations (A.125) using only the anticommutation relations (A.119).*

✎ **A.6.6** *Prove the relations (A.126) using only the anticommutation relations (A.119).*

✎ **A.6.7** *Prove the relations (A.162)–(A.165) using only the relations (10.69).*

https://doi.org/10.1017/9781009291507.017 Published online by Cambridge University Press