# The Standard Model

## University of Cambridge Part III Mathematical Tripos


**David Tong**

*Department of Applied Mathematics and Theoretical Physics,*
*Centre for Mathematical Sciences,*
*Wilberforce Road,*
*Cambridge, CB3 OBA, UK*

http://www.damtp.cam.ac.uk/user/tong/standardmodel.html
d.tong@damtp.cam.ac.uk

**Recommended Books and Resources**

For a very elementary introduction to the Standard Model, you could take a look at the lectures on Particle Physics that I wrote for the CERN summer school. They cover the subject in a great deal of detail, but without any real mathematical sophistication. If you're completely new to the wonderful world of subatomic particles, this is a good place to get grounded.

Many undergraduate degrees have courses on particle physics that use quantum mechanics and some elementary group theory, without fully embracing quantum field theory. There are a number of good textbooks catering to these courses. Two that I particularly like are:

- Halzen and Martin, "*Quarks and Leptons*",
- David Griffiths, "*Introduction to Elementary Particles*"

More advanced and really excellent books are

- Cliff Burgess and Guy Moore "*The Standard Model*"
- Mark Thomson, "*Modern Particle Physics*"
- Matt Schwartz, "*Quantum Field Theory and the Standard Model*"

All three have different perspectives. Cliff and Guy's book in particular is closely aligned to the general theme of these lectures. Mark Thomson's book includes many more details about the specifics of particle interactions, while Matt's book is a great all-round QFT book that, as the title suggests, has an increasing focus on the Standard Model as it proceeds.

Finally, if you're serious about particle physics you should acquaint yourself with the all-important Particle Data Group. They have various apps that you can download and, for the more old-fashioned among you, books. Their booklet, available in the download section of the webpage, is particularly useful. They'll even mail you one for free if you ask nicely.

In addition, there are many online lecture notes. You can find links to these on the course webpage.

# Contents

**Acknowledgements**

I'm grateful to Hugh Osborn and Fernando Quevedo who previously lectured a version of this course in Cambridge, and to Wati Taylor for sharing the notes of his MIT course with me. Many thanks to Ben Allanach for explaining various subtle (and less subtle) issues to me.

This course assumes a familiarity with quantum field theory. You will also need to be comfortable with some group theory.

# 0 Introduction

The "Standard Model" is the comically inadequate name that physicists give to the greatest scientific theory of all time.

This theory is the poster child for success in reductionist science. It describes the universe on the most fundamental level and correctly predicts the results of every experiment that we have ever done, sometimes with unprecedented levels of accuracy.

There are parts of the theory that are stunningly beautiful, with different facets sliding together like a perfect jigsaw, locked in place with a mathematical rigidity that means large parts of the world we inhabit could not be any other way. But there are other aspects of the theory that appear much less elegant, with a couple of dozen parameters that cannot be predicted from first principles but only by measuring them in experiment. These parameters don't appear to be completely random; there are patterns within them that surely hint at some structure that lies beyond the Standard Model, a structure that we have yet to uncover.

Boiled down to its essence, the Standard Model describes a bunch of particles, interacting with three forces. These forces are the strong nuclear force, the weak nuclear force, and electromagnetism. The force of gravity is not part of the Standard Mode but it's straightforward to include it by coupling to a dynamical, curved spacetime. (Claims that the Standard Model is incompatible with general relativity are wildly overblown. The two theories work perfectly well together at all energy scales that we can currently probe by experiment. The difficulties only arise when energies approach the Planck scale.)

Each force in the Standard Model is associated to a Lie group. The upshot is that the Standard Model is built around the group

$$G = U(1) \times SU(2) \times SU(3) \ .$$

Why nature chose the numbers, 1,2, and 3 as the building blocks for her most important theory is not known, but you can't help but smile at the decision. Here $SU(3)$ is associated to the strong force and $SU(2)$ is associated to the weak force and $U(1)$ is not associated to electromagnetism but, instead, to an electromagnetic-like force known as *hypercharge*. It too plays a role in the weak force. The theory of electromagnetism that we know and love can be found hiding within the $SU(2) \times U(1)$ factor.

| | | | |
|---|---|---|---|
| electron<br>1 | down quark<br>9 | up quark<br>4 | electron neutrino<br>$\sim 10^{-6}$ |
| muon<br>207 | strange quark<br>186 | charm quark<br>2495 | muon neutrino<br>$\sim 10^{-6}$ |
| tau<br>3483 | bottom quark<br>8180 | top quark<br>340,000 | tau neutrino<br>$\sim 10^{-6}$ |

**Table 1**. The fermions of the Standard Model

Despite the group theoretic similarities of each force, the resulting physics is wildly different. That's because quantum field theory is cool. It does wonderful and unexpected things. Part of the purpose of this course is to learn about these things and why the dynamics of the strong, weak and electromagnetic forces all play very different roles in our world.

These three forces interact with matter which, in the Standard Model, comes in the form of 15 Weyl fermions which, collectively, go by the name of the electron, the up quark, the down quark, and the neutrino. Why we give just four names to 15 fermions is part of the story that we will unravel, but at heart it is to do with representation theory of the group $G$.

At this point, one of the deepest facts about nature rears its head. The subtleties of quantum field theory mean that this quartet of particles – the electron, neutrino, and up and down quarks – have to come together as a collective. You don't have a choice. The theory with just, say, an electron and an up quark and no companions makes no sense. On grounds of mathematical consistency alone, we're obliged to have this quartet of particles with their particular properties. This is where some of the most beautiful aspects of the Standard Model can be found.

But then nature has a surprise, one which we've known about for almost a century and yet we are seemingly no closer to understanding. Nature took that collection of four particles and, for mysterious reasons, chose to replicate it twice over. This means that the matter in our world is not made of 15 fermions with four different names, but instead of 45 fermions with twelve different names. The names of these twelve particles are shown in Table 1 together with their masses, relative to the electron mass which is

$$m_e \approx 0.51 \text{ MeV} .$$

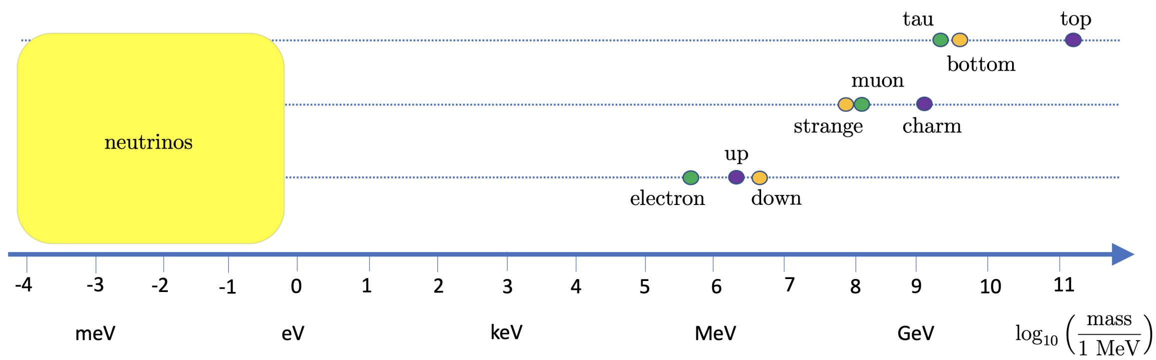


**Figure 1**. Again, the masses of the fermions of the Standard Model. Note that the ordering of particles in each generation is switched.

Each of the three rows in Table 1 is referred to as a different *generation*. The particles in each generation experience identical forces. So, for example, the electron, muon and tau all have electric charge $-1$, the down, strange and bottom quarks all have electric charge $-1/3$ and the up, charm and top quarks all have electric charge $+2/3$. All three neutrinos are neutral.

Similarly, the six quarks all experience the strong force in the same way, while the electron, muon, tau and neutrinos (which, collectively are referred to as *leptons*) are all untouched by the strong force.

The masses of the particles are replicated in Figure 1. They span at least 11 orders of magnitude, maybe more. (The masses of the neutrinos are not well constrained, as shown in the figure.) Why these particular masses? Why this ordering of masses? We have no idea. That's one of the outstanding questions that we hope might be answered by a deeper theory.

There is one final piece of the Standard Model that sits, lording over everything. This is the Higgs boson. It is, in many ways, the thing that ties everything together. In particular, all the masses listed above can be traced to the interactions of various fermions with the Higgs field.

The Higgs is simultaneously both the simplest and the most complicated field in the Standard Model. It is the simplest because it is the only fundamental (as far as we can tell!) scalar field that we have so far observed, meaning that it is the only field to carry zero spin. It is the most complicated because, in contrast to fermions and gauge fields, scalar fields don't come with many consistency requirements which means that there are a plethora of interaction terms that we can write down and the only way we have to constrain their values is to go out and measure them. It's here that we find the two dozen or so parameters that we can't yet explain. And it's here that things get messy and interesting.

This, then, is the Standard Model, part beauty, part beast. A glorious and astonishingly successful theoretical edifice that, so far, has stood firm against everything that experimenters have thrown at it. Yet few believe that it can really be the last word in physics. The Standard Model, like the periodic table before it, surely holds clues for what lies beyond. Our duty as physicists is to understand the Standard Model as best we can, to learn its secrets and, if possible, to let it guide us to a still deeper understanding of the world. The purpose of this course is to take you, at least part way, on this journey.

# 1 Symmetries

A large chunk of the structure of the Standard Model follows from understanding the various symmetries at play. Among these symmetries are

- Poincaré symmetries of spacetime, which restrict us to scalars, fermions, and gauge fields. These are the basic building blocks of the Standard Model.
- Gauge symmetries, better referred to as "gauge redundancies". These dictate the interactions of the spin 1 fields. Indeed, we've already seen that the Standard Model is usually advertised by specifying the gauge group

$$G = U(1) \times SU(2) \times SU(3) \ . \tag{1.1}$$

- Global symmetries. These act on the fermions and include baryon number and lepton number, as well as various approximate flavour symmetries.
- Discrete symmetries. Prominent among these are parity, time-reversal, and charge conjugation. These three symmetries are critically important in the structure of the Standard Model because, we shall see, none of them are actually good symmetries of our universe! But this is one case where not having symmetries puts even stronger constraints on the theory than having symmetries. This is because of something called "anomaly cancellation" that will be described in Section 4.

Of these, the various global symmetries arise because of the specific matter content of the Standard Model and so we will postpone a discussion of them until we have more details in place. (We'll first get there in Section 3 when we describe features of the strong force.) However, the other three symmetries – Poincaré, gauge, and discrete – are ingredients that arise in pretty much all relativistic field theories. For this reason, it makes sense to explore them in some detail in preparation for what's to come.

## 1.1 Spacetime Symmetries

On the length scales appropriate for particle physics, spacetime is effectively flat. This means that the arena for our story is Minkowski space $\mathbb{R}^{1,3}$, equipped with the Minkowski metric

$$\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1) \ . \tag{1.2}$$

We label a point in Minkowski space as $x^\mu = (x^0, x^1, x^2, x^3)$. The set of symmetries of Minkowski space include Lorentz transformations of the form $x^\mu \to \Lambda^\mu{}_\nu x^\nu$ where

$$\Lambda^T \eta \Lambda = \eta \ . \tag{1.3}$$

Embedded among these are a couple of discrete transformations: parity with $\Lambda = \mathrm{diag}(1,-1,-1,-1)$ and time reversal with $\Lambda = \mathrm{diag}(-1,1,1,1)$. These are important enough that we will discuss them separately in Section 1.4. The transformations that are continuously connected to the identity have $\det\Lambda = 1$ and $\Lambda^0{}_0 > 0$ and form the *Lorentz group* $SO(1,3)$. (The restriction to $\Lambda^0{}_0 > 0$ is sometimes written as $SO^+(1,3)$.)

Our main goal in this section is to understand some things about the representations of the Lorentz group and its extension to the Poincaré group which also includes spacetime translations. Among these representations, spinors are the most fiddly and subtle and we will describe some of their properties in Section 1.2.

### 1.1.1 The Lorentz Group

Strictly speaking, the group $SO(1,3)$ doesn't have any spinor representations. However, there is a closely related group called $\mathrm{Spin}(1,3)$ that does admit spinors. This is the double cover, in the sense that

$$SO(1,3) \cong \mathrm{Spin}(1,3)/\mathbb{Z}_2 \tag{1.4}$$

where that $\mathbb{Z}_2$ is related to the famous minus sign that spinors pick up under a $2\pi$ rotation, a minus sign that vectors like $x^\mu$ are oblivious to. The fact that there are spinors in our world is the statement that the true symmetry group is $\mathrm{Spin}(1,3)$ rather than $SO(1,3)$.

The groups $\mathrm{Spin}(1,3)$ and $SO(1,3)$ share the same Lie algebra $so(1,3)$. A Lorentz transformation acting on a 4-vector can be written as

$$\Lambda = \exp\left(-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right) \tag{1.5}$$

where $\omega_{\mu\nu}$ are six numbers that specify what Lorentz transformation we're doing, while $M^{\mu\nu} = -M^{\nu\mu}$ are a choice of six $4\times 4$ anti-symmetric matrices that generate the different Lorentz transformations. The matrix indices are suppressed in the above expressions; in their full glory we would write $(M^{\mu\nu})^\rho{}_\sigma$. So, for example

$$(M^{01})^\rho{}_\sigma = i\begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix} \quad \text{and} \quad (M^{12})^\rho{}_\sigma = i\begin{pmatrix} 0&0&0&0\\ 0&0&-1&0\\ 0&1&0&0\\ 0&0&0&0 \end{pmatrix}. \tag{1.6}$$

(Note that the generators differ by a factor of $i$ from those defined in the Quantum Field Theory lectures. This is compensated by an extra factor of $i$ in the exponent (1.5).) The matrices $M^{\mu\nu}$ generate the algebra $so(1,3)$,

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\left(\eta^{\nu\rho}M^{\mu\sigma} - \eta^{\nu\sigma}M^{\mu\rho} + \eta^{\mu\sigma}M^{\nu\rho} - \eta^{\mu\rho}M^{\nu\sigma}\right)\ . \tag{1.7}$$

The six different Lorentz transformations naturally decompose into three rotations $J_i$ and three boosts $K_i$, defined by

$$J_i = \frac{1}{2}\epsilon_{ijk}M_{jk} \quad \text{and} \quad K_i = M_{0i} \tag{1.8}$$

where the $j,k=1,2,3$ indices are summed over, and $\epsilon_{123}=+1$. The rotation matrices are Hermitian, with $J_i^\dagger = J_i$ while the boost matrices are anti-Hermitian with $K_i^\dagger = -K_i$. This ensures that the rotations in (1.5) give rise to a compact group while the boosts are non-compact. From the Lorentz algebra, we find that these generators obey

$$[J_i, J_j] = i\epsilon_{ijk}J_k \ , \quad [J_i, K_j] = i\epsilon_{ijk}K_k \ , \quad [K_i, K_j] = -i\epsilon_{ijk}J_k \ . \tag{1.9}$$

The rotations form an $su(2)$ sub-algebra. That, of course, is to be expected and is related to the fact that $SO(3) \cong SU(2)/\mathbb{Z}_2$.

We can, however, find two mutually commuting $su(2)$ algebras sitting inside $so(1,3)$. For this we take the linear combinations

$$A_i = \frac{1}{2}(J_i + iK_i) \quad \text{and} \quad B_i = \frac{1}{2}(J_i - iK_i) \ . \tag{1.10}$$

Both of these are Hermitian, with $A_i^\dagger = A_i$ and $B_i^\dagger = B_i$. They obey

$$[A_i, A_j] = i\epsilon_{ijk}A_k \ , \quad [B_i, B_j] = i\epsilon_{ijk}B_k \ , \quad [A_i, B_j] = 0 \ . \tag{1.11}$$

But we know all about representations of $SU(2)$: they are labelled by an integer or half-integer $j \in \frac{1}{2}\mathbb{Z}$ which, in the context of rotations, we call "spin". The dimension of the representation is then $2j+1$. The fact that we can find two $su(2)$ sub-algebras of the Lorentz algebra tells us that all representations must carry two such labels

$$(j_1, j_2) \quad \text{with} \quad j_1, j_2 \in \frac{1}{2}\mathbb{Z} \ . \tag{1.12}$$

Moreover, we know that this representation must have dimension $(2j_1+1)(2j_2+1)$. We'll flesh out the meaning of these representations more below. But for now, we can identify the simplest such representations just by counting: we have

$$\begin{aligned}
(0,0) &: \quad \text{scalar} \\
(\tfrac{1}{2},0) &: \quad \text{left-handed Weyl spinor} \\
(0,\tfrac{1}{2}) &: \quad \text{right-handed Weyl spinor} \\
(\tfrac{1}{2},\tfrac{1}{2}) &: \quad \text{vector} \\
(1,0) &: \quad \text{self-dual 2-form} \\
(0,1) &: \quad \text{anti-self-dual 2-form}
\end{aligned} \tag{1.13}$$

What we call the physical spin of a particle is the quantum number under rotations $\vec{J}$: this is $j = j_1 + j_2$. The spin-statistics theorem ensures that particles with $j \in \mathbb{Z}$ are bosons, while those with $j \in \mathbb{Z} + \frac{1}{2}$ are fermions.

There's something a little odd about the our discovery of two $su(2)$ sub-algebras. After all, it certainly isn't true that the Lorentz group is isomorphic to two copies of $SU(2)$. This is because $SU(2)$ is a compact group: keep doing a rotation and you will eventually get back to where you started. Indeed, two copies of the group $SU(2)$ give the rotation group of Euclidean space $\mathbb{R}^4$:

$$\mathrm{Spin}(4) \cong SU(2) \times SU(2) \quad \text{with} \quad SO(4) \cong \mathrm{Spin}(4)/Z_2 \ . \tag{1.14}$$

In contrast, the Lorentz group is non-compact: keep boosting and you get further and further from where you started. How does this manifest itself in the two $su(2)$ algebras that we've found in (1.11)?

The answer is a little subtle and is to be found in the reality properties of the generators $A_i$ and $B_i$. Recall that all integer, $j \in \mathbb{Z}$, representations of $SU(2)$ are real, while all half-integer spin, $j \in \mathbb{Z} + \frac{1}{2}$, are pseudoreal (which means that, while not actually real, the representation is isomorphic to its complex conjugate). However, the $A_i$ and $B_i$ in (1.11) do *not* have these properties. You can see in (1.6) that both $J_i$ and $K_i$ are pure imaginary. This, in turn, means that the generators $A_i$ and $B_i$ are complex conjugates of each other

$$(A_i)^\star = -B_i \ . \tag{1.15}$$

This is where the difference lies that distinguishes $SO(4)$ from $SO(1,3)$. The Lie algebra $so(1,3)$ does not contain two, mutually commuting copies of the real Lie algebra $su(2)$, but only after a suitable complexification. This means that certain complex linear combinations of the Lie algebra $su(2) \times su(2)$ are isomorphic to $so(1,3)$. To highight this, the relationship between the two is sometimes written as

$$so(1,3) \cong su(2) \times su(2)^\star \ . \tag{1.16}$$

For our purposes, it means that the complex conjugate of a representation $(j_1, j_2)$ exchanges the two quantum numbers

$$(j_1, j_2)^\star = (j_2, j_1) \ . \tag{1.17}$$

Both the scalar representation $(0,0)$ and the vector representation $(\frac{1}{2}, \frac{1}{2})$ are real, while the left- and right-handed Weyl spinors $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ are exchanged under complex conjugation. This last statement, which is important, will be elaborated upon in Sections 1.2 and 1.4. In the context of quantum field theory, if a field appears in a theory then so too does its complex conjugate. This means that if you have a left-handed spinor, you also have a right-handed complex conjugated spinor.

### 1.1.2 The Poincaré Group and its Representations

The continuous symmetries of Minkowski space comprise of Lorentz transformations together with spacetime translations. Combined, these form the *Poincaré group*. Spacetime translations are generated, as usual, by the momentum 4-vector $P^\mu$. Their commutation relations with themselves and with the Lorentz generators $M^{\mu\nu}$ are given by

$$[P^\mu, P^\nu] = 0 \quad \text{and} \quad [M^{\mu\nu}, P^\sigma] = i\left(P^\mu \eta^{\nu\sigma} - P^\nu \eta^{\mu\sigma}\right) \tag{1.18}$$

The latter of these is equivalent to the statement that $P^\mu$ transforms as a 4-vector under Lorentz transformations. These commutation relations should be considered in conjunction with the Lorentz algebra (1.7),

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\left(\eta^{\nu\rho} M^{\mu\sigma} - \eta^{\nu\sigma} M^{\mu\rho} + \eta^{\mu\sigma} M^{\nu\rho} - \eta^{\mu\rho} M^{\nu\sigma}\right) \tag{1.19}$$

Together, (1.18) and (1.19) form the algebra of the Poincaré group.

Given an algebra, our next task is to explore its representations. There are different ways that we could approach this. Ultimately, we will be interested in the way that the Poincaré group acts on fields that make up the Standard Model. But first, to build some intuition, we will understand how the Poincaré group acts on single particle states in the Hilbert space.

To set the scene, let's first recall how we construct irreducible representations of the rotation group. We work with the algebra $so(3) \cong su(2)$ rather than the group. This is, of course, defined by the familiar commutation relations

$$[J_i, J_j] = i\epsilon_{ijk} J_k \ . \tag{1.20}$$

To construct representations, the first thing we do is look to the Casimirs. These are operators that commute with all generators of the group. For $su(2)$, there is just a single Casimir,

$$C = \sum_{i=1}^{3} J_i^2 \ . \tag{1.21}$$

Irreducible representations are labelled by the eigenvalue of the Casimir. For $su(2)$, the eigenvalue of $J^2$ is $j(j+1)$ with the spin $j$ taking values in $j = 0, \frac{1}{2}, 1, \ldots$. Each representation has dimension $2j+1$, with the states within a multiplet identified by their eigenvalue under, say, $J_3$ whose eigenvalue lies in the range $|j_3| \leq j$. The result is the familiar one from quantum mechanics: states are labelled by two quantum numbers $|j, j_3\rangle$

Now let's turn to the Poincaré group. The irreducible representations are what we call "particles". Again, they are characterised by the Casimirs. I won't tell you how to construct Casimirs, but will instead just present you with the result. First, we introduce the *Pauli-Lubański vector*,

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} P_\nu M_{\rho\sigma} \ . \tag{1.22}$$

This can be thought of as a relativistic version of angular momentum. You can easily check this commutes with momentum $[W_\mu, P_\nu] = 0$. The remaining non-trivial commutation relations are somewhat more laborious to show:

$$[W_\mu, M_{\nu\rho}] = i(\eta_{\mu\nu} W_\rho - \eta_{\mu\rho} W_\nu) \quad \text{and} \quad [W_\mu, W_\nu] = -i\epsilon_{\mu\nu\rho\sigma} W^\rho P^\sigma \ . \tag{1.23}$$

The last of these commutation relations is quadratic on the right-hand side and so we're not looking at a Lie algebra here, but something more complicated. (This is reminiscent of the Runge-Lenz vector which is a conserved quantity for the Kepler problem; there too, the Poisson bracket structure returns something quadratic on the right-hand side.)

The two Casimirs of the Poincaré group are formed from the momentum $P_\mu$ and the Pauli-Lubański vector $W_\mu$,

$$C_1 = P_\mu P^\mu \quad \text{and} \quad C_2 = W_\mu W^\mu \ . \tag{1.24}$$

This is our starting point: representations of the Poincaré group are labelled by the eigenvalues of $C_1$ and $C_2$, together with the eigenvalues of any other operators that we can find to make a maximally commuting set, analogous to $J_3$ for the angular momentum.

The most important of these "other operators" is the momentum $P^\mu$ itself. All states will be labelled by the eigenvalue $p^\mu$ which is simply the 4-momentum of the particle. The first Casimir is then just the rest mass of the particle, $C_1 = p_\mu p^\mu = m^2$. By acting with rotations and boosts $M_{\mu\nu}$, we can change the momentum to take any value subject to the constraint $p_\mu p^\mu = m^2$. In the rotation analogy, the different values of $p^\mu$ are like the different values of $j_3$ in the multiplet. However, in contrast to rotations, representations of the Poincaré group will necessarily be infinite dimensional, labelled (among other things) by the continuous variable $p^\mu$. This difference can be traced to the fact that the Poincaré group is non-compact while the rotation group is not.

What happens next depends on whether we're dealing with massive or massless particles. We describe each in turn, followed by a somewhat mysterious massless representation that no one really knows what to make of.

**Massive Representations**

First, consider the situation when $C_1 = m^2 \neq 0$. It's fruitful to pick a representative value of the momentum $p^\mu$ and the simplest choice is to boost to the rest frame of the particle so that $p^\mu = (m, 0, 0, 0)$. In this frame, the Pauli-Lubański vector is

$$W^0 = 0 \quad \text{and} \quad W^i = -mJ^i \ . \tag{1.25}$$

with $J^i$ the generators of rotations. Note that the rotation generators $J^i$ are precisely those elements of the Lorentz group that *don't* change the value of our chosen momentum $p^\mu = (m, 0, 0, 0)$. That means that these generators $J^i$ must act on whatever other degrees of freedom are carried by the particles. We want to ask: what are the allowed extra degrees of freedom?

But this is a question that we already answered above because our problem has reduced to finding a representation of the Lie algebra $su(2)$, generated by $J^i$. The second quadratic Casimir of the Poincaré group is $C_2 = -m^2 J^2$ and so is specified by the eigenvalue of $J^2$ which, as we reviewed above, is $j(j+1)$ for some $j \in \frac{1}{2}\mathbb{Z}$. The full multiplet is then filled out by the different values of $j_3$ with $|j_3| \leq j$.

We've seen that, if we fix the momentum to the specific value $p^\mu = (m, 0, 0, 0)$, then we're left with finding representations of the rotation group. But, importantly, it doesn't matter which value of the momentum we started with: had we picked a different $p^\mu$ (still with $p_\mu p^\mu = m^2$), then we'd have got the same result. This suggests that we can lift the $SU(2)$ representation that we found for our given $p^\mu$ to a representation of the full Poincaré group. And, indeed, this is the case.

There is a theorem underlying this result which we won't prove. Instead, I'll just give you some names of things. Once we fix the momentum $p^\mu$, the elements of the Lorentz group that don't change $p^\mu$ form a group known as the *little group*. For massive particles, the little group is $SU(2)$. One can then show that representations of the little group uplift to representations of the full Poincaré group. This is what's known as an *induced representation*.

The upshot is something familiar: massive particles are characterised by their mass $m$ and spin $j$. Given these Casimirs, states in this representation of the Poincaré group are labelled by $|p_\mu, j_3\rangle$.

**Massless Representations**

The story is slightly different for massless particles, for which the first Casimir vanishes: $C_1 = m^2 = 0$. We again choose a representative momentum. This time we can't boost to the rest frame, but we can choose the momentum to take the form $p^\mu = (E, 0, 0, E)$ where $E$ is the energy of the particle. A short calculation shows that, in this frame, the Pauli-Lubański now takes the form

$$W_\mu = E \begin{pmatrix} -M_{12} \\ M_{23} - M_{02} \\ M_{31} + M_{01} \\ M_{12} \end{pmatrix} = E \begin{pmatrix} -J_3 \\ J_1 - K_2 \\ J_2 + K_1 \\ J_3 \end{pmatrix} \ . \tag{1.26}$$

Here we've replaced the $M_{\mu\nu}$ with the appropriate rotation generator $J_i$ or boost generator $K_i$ defined in (1.8). Once again, each of the components of $W_\mu$ leaves our initial momentum $p^\mu = (E, 0, 0, E)$ unchanged, a fact that you can check by looking at the explicit form of the generators (1.6). In other words, these components of $W_\mu$ are once again our little group. (This has happened twice now and it is no coincidence: the structure of the Pauli-Lubański vector was designed so that this holds.)

What group do the components of $W^\mu$ actually generate? We can look at their commutation relations which, using (1.9), are

$$[W_1, W_2] = 0 \ , \quad [W_3, W_2] = -iEW_1 \ , \quad [W_3, W_1] = iEW_2 \ . \tag{1.27}$$

This is the Euclidean group in $\mathbb{R}^2$, sometimes written as $ISO(2)$, with $W_1$ and $W_2$ the generators of translations and $W_3$ the generator of rotations. Again, the little group doesn't act on our chosen $p^\mu = (E, 0, 0, E)$, but it may act on any other degrees of freedom that our state carries. Said differently, those other degrees of freedom must fall into a representation of the 2d Euclidean group.

Here a subtlety rears its head. For reasons that we will explain below, things turn out to be simplest if we consider representations of the little group on which the translation generators $W_1$ and $W_2$ act trivially. If we ignore these translations, the remaining little group is just the $U(1)$ of rotations generated by $J_3$. Representations of this $U(1)$ are labelled by a single eigenvalue $h$ such that the states transform as

$$e^{i\theta J_3}|h\rangle = e^{ih\theta}|h\rangle \ . \tag{1.28}$$

The eigenvalue $h$ is called the *helicity* and is the analog of spin for massless particles. At time, we'll be lazy and just refer to both as "spin". For a general null $p$, the helicity tells us the eigenvalue of the state under a rotation along the direction of motion,

$$e^{i\theta\,\hat{\mathbf{p}}\cdot\mathbf{J}}|p_\mu;h\rangle = e^{ih\theta}|p_\mu;h\rangle\ . \tag{1.29}$$

Because the $U(1)$ generated by $J_3$ was a subgroup $U(1) \in SU(2)$, we know that this helicity is quantised to take values

$$h \in \frac{1}{2}\mathbb{Z}\ . \tag{1.30}$$

This is the statement that, under a rotation of $\theta = 2\pi$, the states are either left the same (for $h \in \mathbb{Z}$) or pick up a minus sign (for $h \in \mathbb{Z} + \frac{1}{2}$).

There's something missing in the story above. For massive representations, we've seen that the states are labelled by $m$ and $j$ and fill out a multiplet $|p_\mu, j_3\rangle$ with $|j_3| \leq j$. This multiplet has dimension $2j+1$. (Ok, the multiplet is really infinite dimensional because of the $p_\mu$, but for a fixed $p_\mu$ the multiplet has dimension $2j+1$.)

However, for massless particles there is just a single state $|p_\mu;h\rangle$. This is because the helicity describes the representation of the Abelian group $U(1)$ generated by $J_3$ rather than the non-Abelian group $SU(2)$ and irreducible representations of Abelian groups are one-dimensional.

The problem with this is that it doesn't fit with what we know about massless particles. For example, the photon has helicity $h = 1$ and has two polarisation states, as does a graviton with $h = 2$. A massless spinor with $h = \frac{1}{2}$ also has two degrees of freedom. Why aren't we seeing this doubling in our representation theory analysis?

What we're missing is the additional requirement that the spectrum of states is invariant under $CPT$. These are discrete symmetries that we will look at more closely in Section 1.4. For massive particles, this doesn't buy us anything new: the set of states $|p_\mu, j\rangle$ is already invariant under $CPT$. However, for massless particles $CPT$ flips $h \mapsto -h$ and tells us that massless states must come in pairs

$$|p_\mu;h\rangle \quad \text{and} \quad |p_\mu;-h\rangle\ . \tag{1.31}$$

This is the origin of the two polarisation states of the photon or graviton, or the two helicities of a massless Weyl spinor. Note that a massless scalar has helicity $h = 0$ and so is $CPT$ self-conjugate. This means that there's no requirement from $CPT$ to add an additional degree of freedom in this case.

### Weird Continuous Spin Representations

We brushed over something above. When looking at massless representations, we found that the little group coincides with the 2d Euclidean group (1.27). But then, without justification, we restricted ourselves to representations on which the translation generators $W_1$ and $W_2$ act trivially. Here we give the justification.

Let's look at representations of the 2d Euclidean group (1.27) for which translations $W_1$ and $W_2$ act non-trivially. Because $[W_1, W_2] = 0$, we can simultaneously diagonalise these generators so that they act on states $|w_1, w_2\rangle$ such that

$$W_i|w_1, w_2\rangle = w_i|w_1, w_2\rangle \quad \text{for} \quad i = 1, 2 \ . \tag{1.32}$$

The second Casimir is then

$$C_2 = W^\mu W_\mu = -(w_1^2 + w_2^2) \ . \tag{1.33}$$

For the massless representations above, we assumed that $w_1 = w_2 = 0$. Now we want to understand what happens when they are non-zero. Since $C_2$ is fixed, we write $w_1 = \rho \cos\alpha$ and $w_2 = \rho \sin\alpha$ with $C_2 = -\rho^2$ and we should think of the collection of states $|w_1, w_2\rangle$ as parameterised by the angle $\alpha \in [0, 2\pi)$ with the action

$$W_1|\alpha\rangle = \rho\cos\alpha|\alpha\rangle \quad \text{and} \quad W_2|\alpha\rangle = \rho\sin\alpha|\alpha\rangle \ . \tag{1.34}$$

It remains to determine the action of $W_3 = EJ_3$ on these states. This is given by

$$e^{i\theta J_3}|\alpha\rangle = e^{ih\theta}|\alpha + \theta\rangle \quad \Longrightarrow \quad J_3|\alpha\rangle = h|\alpha\rangle - i\frac{d}{d\alpha}|\alpha\rangle \ . \tag{1.35}$$

You can check that the actions (1.35) and (1.34) do indeed furnish a representation of the 2d Euclidean algebra (1.27). But, from the perspective of particle physics, it's a very weird representation. This is because particle states $|p_\mu, \alpha; h\rangle$ are labelled by their momentum $p_\mu$ and an additional angle $\alpha \in [0, 2\pi)$. This means that for every choice of momentum $p_\mu$, there's still an infinite dimensional Hilbert space, labelled by the continuous parameter $\alpha$ rather than a discrete, bounded parameter like $j_3$. Said differently, it's as if we have an uncountably infinite number of species of particle. These are known as *continuous spin representations*.

We've certainly never observed particles corresponding to these states and they would have very strange properties (such as infinite heat capacity). Nonetheless, one can't help but wonder if nature may make use of them somewhere.

### 1.1.3 The Coleman-Mandula Theorem

It's not unusual for quantum field theories to exhibit further continuous symmetries. Say, a global $U(1)$ symmetry that rotates the phase of a complex field, or perhaps a non-Abelian $SU(N)$ symmetry under which a multiplet of fields transforms. The generators of these symmetries – which we'll denote collectively as $T$ – correspond to some conserved charge and are always Lorentz scalars which means that they necessarily commute with the Poincaré generators,

$$[P^\mu, T] = [M^{\mu\nu}, T] = 0 \ . \tag{1.36}$$

One could ask: is it possible for something less trivial to happen, with the new generators transforming in some fashion under the Poincaré group? For example, this would happen if the additional generators $T$ themselves carried some spacetime index. If this were possilble, the Poincaré group would be subsumed into a larger group. And that sounds interesting.

A theorem due to Coleman and Mandula greatly restricts this possibility. Roughly speaking, the theorem states that, in any spacetime dimension greater than $d = 1 + 1$, the symmetry group of any interacting quantum field theory must factorise as

$$\text{Poincaré} \times \text{Internal} \ . \tag{1.37}$$

We won't prove the Coleman-Mandula theorem here. The gist of the proof is to look at 2-to-2 scattering (meaning two incoming particles scatter into two outgoing particles). Poincaré invariance already greatly restricts what can happen, with only the scattering angle left undetermined. Any internal symmetries that factorise, as in (1.37), put restrictions on the kinds of interactions that are allowed, for example enforcing conservation of electric charge. But if the generators $T$ were to carry a spacetime index then they would put further constraints on the scattering angle itself and that would be overly restrictive, at best allowing scattering to occur only at discrete angles. But if one assumes that the scattering amplitudes are analytic functions of the angle then the amplitude must vanish for all angles and the theory is free.

Like all no-go theorems in physics, the Coleman-Mandula theorem comes with a number of underlying assumptions. Some of these are eminently reasonable, such as locality and causality. But it may be possible to relax other assumptions to find interesting loopholes to the Coleman-Mandula theorem. Two such loopholes have proven to be extremely important.

- **Conformal Invariance:** The Coleman-Mandula theorem assumes that the theory has a mass gap, meaning that all particles are massive. Indeed, the theorem is a statement about symmetries of the S-matrix which is really only well defined for massive particles where we don't have to worry about IR divergences. For theories of massless particles something interesting can, and often does, happen.

The first interesting thing is that interacting massless theories typically exhibit scale invariance. This means that physics is unchanged under the symmetry $x^\mu \to \lambda x^\mu$. The associated symmetry generator is called $D$ for "dilatation". This can only be a symmetry of a theory that has no dimensionful parameters, which is the main reason it can occur only for massless theories.

The second interesting thing is more surprising. For reasons that are not entirely understood, theories that exhibit scale invariance also exhibit a further symmetry known as *special conformal transformations* of the form

$$x^\mu \to \frac{x^\mu - a^\mu x^2}{1 - 2a\cdot x + a^2 x^2} \ . \tag{1.38}$$

This transformation depends on a vector parameter $a^\mu$ and the associated generator is a 4-vector $K^\mu$. The resulting conformal algebra extends the Poincaré algebra (1.18) and (1.19) with the non-trivial commutators

$$\begin{aligned} [D, K^\mu] = -iK_\mu \quad &, \quad [D, P^\mu] = iP^\mu \\ [K^\mu, P^\nu] &= 2i(D\eta^{\mu\nu} - M^{\mu\nu}) \\ [M^{\mu\nu}, K^\sigma] &= i\left(K^\nu\eta^{\mu\sigma} - K^\mu\eta^{\nu\sigma}\right) \quad . \end{aligned} \tag{1.39}$$

Interacting conformal field theories crop up in many places in physics. In their Euclidean incarnation, they describe critical points, or second order phase transitions, that were the focus of our lectures on Statistical Field Theory. In $d = 1+1$ dimensions the conformal group has rather more structure and a detailed introduction can be found in the lectures on String Theory.

- **Supersymmetry:** The second loophole to the Coleman-Mandula theorem is supersymmetry. This is a symmetry that relates bosons to fermions. The generator that enacts this magical transformation is denoted as $Q_\alpha$ and carries a spacetime spinor index $\alpha = 1, 2$. (We will learn more about spinors in Section 1.2.) This is exactly the kind of thing that the Coleman-Mandula theorem is supposed to rule out. However, supersymmetry evades the theorem because the generators $Q_\alpha$ do not form a Lie algebra: instead they form what is known as a *super-Lie algebra*, with the commutation relations of the Poincaré group (1.18) and (1.19) augmented by the anti-commutation relation

$$\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} = 2\sigma^\mu_{\alpha\dot{\alpha}} P_\mu \ . \tag{1.40}$$

Here $\sigma^\mu_{\alpha\dot\alpha}$ are a collection of $2 \times 2$ matrices defined in (1.44). (We'll see a lot more about what the $\alpha$ and $\dot\alpha$ spinor indices mean shortly.) You can learn (a lot!) more about this algebra and its consequences for various field theories in the lectures on Supersymmetry.

Neither conformal symmetry nor supersymmetry play a role in the Standard Model. However, both arise in different ways when it comes to ideas for what lies beyond the Standard Model.

## 1.2 Spinors

Scalars are basic. They have no internal structure and, as such, come with very little baggage. There's a lot of fun that we can have with them, largely by writing down potentials that do interesting things, and we'll see examples of this when we discuss spontaneous symmetry breaking in Section 2. But there's little that is subtle about scalars: what you see is what you get.

In contrast, any field with higher spin is awash with subtleties. For massless spin 1 particles, like photons, these subtleties are all about gauge invariance and we will discuss them in Section 1.3. Here our interest is in spin $\frac{1}{2}$ particles, known as spinors. These are the fields that describe all matter particles in the Standard Model, meaning the quarks and leptons. They are subtle largely because anything that comes back to itself with a minus sign after a $2\pi$ rotation is always going to be a little strange.

### 1.2.1 Dirac vs Weyl Spinors

We start by reviewing some features of spinors that we met in the lectures on Quantum Field Theory. However, our focus in going to be a little different. In particular, to prepare us for the Standard Model, we will need to look more closely at the properties of Weyl spinors.

In the lectures on Quantum Field Theory, we learned about the 4-component Dirac spinor $\psi$. This comes hand in hand with a collection of gamma matrices that obey the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \ . \tag{1.41}$$

The Clifford algebra admits a unique irreducible representation, up to conjugation. But that "up to conjugation" caveat hides all manner of headaches as it provides ample opportunity for physicists to use annoying conventions. Here we use the chiral basis of gamma matrices,

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \quad \text{and} \quad \gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{1.42}$$

where we've introduced two collections of $2 \times 2$ matrices,

$$\sigma^\mu = (\mathbb{1}, \sigma^i) \quad \text{and} \quad \bar{\sigma}^\mu = (\mathbb{1}, -\sigma^i) \tag{1.43}$$

where $\sigma^i$ with $i = 1, 2, 3$ are the familiar Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{1.44}$$

The bar on $\bar{\sigma}^\mu$ in (1.43) doesn't denote complex conjugation: these are simply a different collection of $2 \times 2$ matrices from $\sigma^\mu$.

In the Quantum Field Theory lectures, we showed that the generators of Lorentz transformations for a Dirac spinor are

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu] = \begin{pmatrix} \sigma^{\mu\nu} & 0 \\ 0 & \bar{\sigma}^{\mu\nu} \end{pmatrix} \ . \tag{1.45}$$

(As with our earlier definition of $M^{\mu\nu}$, this differs by a factor of $i$ from the conventions in the Quantum Field Theory lectures.) Here we've defined

$$\begin{aligned} \sigma^{\mu\nu} &= \frac{i}{4} \left( \sigma^\mu \bar{\sigma}^\nu - \sigma^\nu \bar{\sigma}^\mu \right) \\ \bar{\sigma}^{\mu\nu} &= \frac{i}{4} \left( \bar{\sigma}^\mu \sigma^\nu - \bar{\sigma}^\nu \sigma^\mu \right) \ . \end{aligned} \tag{1.46}$$

Because both of these expressions are anti-symmetrised in $\mu$ and $\nu$, each is a collection of six $2 \times 2$ matrices.

The generators $S^{\mu\nu}$ defined in (1.45) are block diagonal. This is telling us that they are *not* an irreducible representation of the Lorentz group. Instead, it's formed of two distinct representations, one generated by $\sigma^{\mu\nu}$ and the other generated by $\bar{\sigma}^{\mu\nu}$. Indeed, you can check that each of these obeys the Lorentz algebra (1.5)

$$[\sigma^{\mu\nu}, \sigma^{\rho\sigma}] = i \left( \eta^{\nu\rho} \sigma^{\mu\sigma} - \eta^{\nu\sigma} \sigma^{\mu\rho} + \eta^{\mu\sigma} \sigma^{\nu\rho} - \eta^{\mu\rho} \sigma^{\nu\sigma} \right) \tag{1.47}$$

with a similar expression for $\bar{\sigma}^{\mu\nu}$ . Correspondingly, the 4-component Dirac spinor $\psi$ also decomposes into two 2-component spinors

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \ . \tag{1.48}$$

These are referred to as left-handed and right-handed spinors respectively. In the language of our earlier table of representations (1.13), $\psi_L$ sits in the $(\frac{1}{2}, 0)$ representation while $\psi_R$ sits in the $(0, \frac{1}{2})$ representation. A Dirac spinor is a combination of both representations $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$.

Under a Lorentz transformation, a left-handed Weyl spinor transforms as

$$\psi_L \ \rightarrow \ S\psi_L \quad \text{with} \quad S = \exp\left(-\frac{i}{2}\omega_{\mu\nu}\sigma^{\mu\nu}\right) \ . \tag{1.49}$$

Here $\omega_{\mu\nu}$ are the same set of six numbers that specify the Lorentz transformation (1.5). There is a similar expression for $\psi_R$, with $\sigma^{\mu\nu}$ replaced by $\bar{\sigma}^{\mu\nu}$.

You can check that $\operatorname{tr}\sigma^{\mu\nu} = 0$ and so, using $\det(e^A) = e^{\operatorname{tr} A}$, we have $\det S = 1$. In fact, $S \in SL(2, \mathbb{C})$, and what we've done in constructing the Weyl spinor representation of the Lorentz group is highlight the group isomorphism $\text{Spin}(1,3) \cong SL(2, \mathbb{C})$.

**(Left-Handed)$^\star$ = Right-Handed**

The two representations – one for a left-handed Weyl spinor, the other for a right-handed Weyl spinor – are related by complex conjugation.

It's not immediately obvious because, as we've seen, the generators are $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$ and it's not true that these generators are complex conjugates: $(\sigma^{\mu\nu})^\star \neq \bar{\sigma}^{\mu\nu}$. To see the relation, we need an additional conjugation by the anti-symmetric tensor

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ . \tag{1.50}$$

You can then check that

$$\epsilon^T(\sigma^{\mu\nu})^\star\epsilon = \bar{\sigma}^{\mu\nu} \ . \tag{1.51}$$

Operationally, the complex conjugation flips the sign of $(\sigma^2)^\star = -\sigma^2$ leaving the other Pauli matrices alone: $(\sigma^i)^\star = \sigma^i$ for $i = 1, 3$. But the conjugation by $\epsilon = i\sigma^2$ then flips the sign of $\sigma^i$ with $i = 1, 3$, leaving $\sigma^2$ alone.

This simple algebraic relation has an important physical implication. If you have a left-handed particle described by a Weyl spinor $\psi_L$, then its anti-particle is described by the conjugate spinor $\psi_L^\dagger$ (which we also write as $\bar{\psi}_L$) and is right-handed.

**Building Scalars from Spinors**

If we're given two left-handed spinors, $\psi_L$ and $\chi_L$, then we can build a scalar. We'll adorn our spinors with indices, so we have $(\psi_L)_\alpha$ and $(\chi_L)_\alpha$ with $\alpha = 1, 2$. We also add indices to our anti-symmetric matrix

$$\epsilon^{\alpha\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ . \tag{1.52}$$

We then define the scalar quantity

$$\psi_L\,\chi_L := \epsilon^{\alpha\beta}(\psi_L)_\beta(\chi_L)_\alpha = (\psi_L)_2(\chi_L)_1 - (\psi_L)_1(\chi_L)_2 \ . \tag{1.53}$$

To see that this does indeed transform as a scalar, we look at

$$\psi_L\,\chi_L \to \ S_\alpha{}^\gamma S_\beta{}^\delta \epsilon^{\alpha\beta}(\psi_L)_\delta(\chi_L)_\gamma = (\det S)\epsilon^{\gamma\delta}(\psi_L)_\delta(\chi_L)_\gamma = \psi_L\,\chi_L \tag{1.54}$$

where, in the first equality we've used the fact that $S_\alpha{}^\gamma S_\beta{}^\delta \epsilon^{\alpha\beta} = \det S\,\epsilon^{\gamma\delta}$, which you can confirm simply by checking all the cases $\gamma, \delta = 1, 2$. In the second equality we've used the fact that $\det S = 1$.

This is an important lesson: you can form a scalar from two left-handed spinors. In terms of the representation theory of the previous section, what we're seeing here is the tensor product $(\frac{1}{2}, 0) \otimes (\frac{1}{2}, 0) = (0, 0) \oplus (1, 0)$, where the scalar (1.53) picks out the singlet $(0, 0)$.

The anti-symmetric tensor $\epsilon^{\alpha\beta}$ is an invariant tensor for the group $SL(2, \mathbb{C})$. In that sense, it plays a role that is similar to the delta function $\delta^{ab}$ for the group $SO(N)$, or the Minkowski metric $\eta^{\mu\nu}$ for the group $SO(1, 3)$. In particular, it allows us to form a scalar product between two spinors as in (1.53). The fact that this product is anti-symmetric, rather than symmetric, fits nicely with the fact that, in quantum field theory, spinors are anti-commuting variables. This means that we have,

$$\psi_L\,\chi_L = (\psi_L)_2(\chi_L)_1 - (\psi_L)_1(\chi_L)_2 = -(\chi_L)_1(\psi_L)_2 + (\chi_L)_2(\psi_L)_1 = \chi_L\,\psi_L \ . \tag{1.55}$$

In particular, this means that we can form a scalar from just a single left-handed Weyl spinor

$$\psi_L\,\psi_L = (\psi_L)_2(\psi_L)_1 - (\psi_L)_1(\psi_L)_2 = 2(\psi_L)_2(\psi_L)_1 \ . \tag{1.56}$$

Again, there are similar expressions for right-handed spinors.

There's quite a bit more to say about the two different representations of the Lorentz group and their properties. You can read about this (and the corresponding dotted and undotted indices) in the first section of the lectures on Supersymmetry. But the simple summary above will suffice for our purposes.

### 1.2.2 Actions for Spinors

Our next goal is to understand how to construct Lagrangians for spinors. Again, our starting point will be the Dirac spinor that we met in Quantum Field Theory. There we saw that the Lorentz invariant action is

$$S_{\text{Dirac}} = -\int d^4x \, \left( i\bar{\psi}\gamma^\mu \partial_\mu \psi - M\bar{\psi}\psi \right) \, . \tag{1.57}$$

For a Dirac spinor, the bar notation means $\bar{\psi} = \psi^\dagger \gamma^0$. Decomposed in terms of Weyl fermions (1.48),

$$S_{\text{Dirac}} = -\int d^4x \, \left( i\bar{\psi}_L \bar{\sigma}^\mu \partial_\mu \psi_L + i\bar{\psi}_R \sigma^\mu \partial_\mu \psi_R - M(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \right) \, . \tag{1.58}$$

First an important, but trivial, notational point: the bar for a Weyl spinor means something different from a bar for a Dirac spinor. It is simply a more elegant way of writing $\bar{\psi}_L = \psi_L^\dagger$.

Second, note that the mass term couples the left- and right-handed Weyl spinors. Combining our observations above, we know that the complex conjugate $\bar{\psi}_R$ is a left-handed spinor, and so in writing $\bar{\psi}_R \psi_L$ we've combined two left-handed spinors into a scalar. Similarly, $\bar{\psi}_L \psi_R$ combines two right-handed spinors into a scalar.

It's worth pausing to look at the symmetries of the action (1.58). Crucially, these symmetries are different for massless and massive fermions. In the absence of the mass term, so $M = 0$, the action has a $U(1)^2$ symmetry, under which the two fermions rotate separately, $\psi_L \to e^{i\alpha}\psi_L$ and $\psi_R \to e^{i\beta}\psi_R$. When we turn on the mass term, only the diagonal combination, with $\alpha = \beta$ survives. This is a general story, and one that will be particularly important for understanding the Standard Model: massless fermions always have more symmetries than massive fermions.

The mass in (1.58) can take values $M \in \mathbb{R}$. (There's no positivity requirement.) Upon quantisation, with $M \neq 0$, we get a particle of spin $+\frac{1}{2}$ and charge $+1$ under the surviving $U(1)$, together with a distinct anti-particle of spin $+\frac{1}{2}$ and charge $-1$, both with mass $|M|$.

The mass term in (1.58) which combines two different spinors, $\psi_L$ and $\psi_R$, is known as a *Dirac mass*. It's not the only thing we can write down. Suppose that we have just a left-handed spinor $\psi_L$. Then it's perfectly possible to write down an action with a mass term,

$$\mathcal{S}_{\rm Weyl} = -\int d^4x\, \left( i\bar{\psi}_L \bar{\sigma}^\mu \partial_\mu \psi_L + \frac{m}{2}\psi_L\psi_L + \frac{m^\star}{2}\bar{\psi}_L\bar{\psi}_L \right) \quad . \tag{1.59}$$

This is known as a *Majorana mass*. Here we can take $m \in \mathbb{C}$.

Again, the massive theory has less symmetry than the massive theory, with the $U(1)$ that rotates the phase of $\psi_L$ broken when $m \neq 0$. This means that there's no $U(1)$ quantum number to distinguish particles from anti-particles and, upon quantisation, the theory describes a single spin $\frac{1}{2}$ particle with mass $|m|$ that is now its own anti-particle.

Because the Majorana mass term explicitly breaks the $U(1)$ symmetry, it is not allowed if the $U(1)$ is gauged. Relatedly, it's not possible to write down such a term for any fermion $\psi_L$ that transforms in a complex representation of a gauge group. It is, however, possible to write down such terms for fermions in real representations.

## 1.3 Gauge Invariance

In the Standard Model, forces are associated to massless spin 1 particles, known collectively as *gauge bosons*. As we now explain, much of the dynamics of these forces is fixed by gauge invariance.

### 1.3.1 Maxwell Theory

The key ideas of gauge invariance are familiar from electromagnetism. There, the fundamental field is the 4-vector $A_\mu(x)$, known as the *gauge potential*. Crucially, not all components of $A_\mu(x)$ are physical: instead, we should identify any two gauge potentials that are related by a gauge transformation of the form

$$A_\mu \to A_\mu + \partial_\mu \alpha \tag{1.60}$$

for any function $\alpha(x)$. The transformation (1.60) is sometimes called a *gauge symmetry*. It's not a good name. A "symmetry" describes a situation in which two physically distinct configurations share the same physics. But that's not what's going on in (1.60). Instead, the two configurations related by a gauge transformation describe the *same* physical configuration. A fairly decent analogy is to think of two gauge potentials that are related by (1.60) in the same way as you would view two different coordinate systems. A much better name would be *gauge redundancy*.

As we proceed, we'll see that a great deal of the structure of the Standard Model is determined by the requirements of gauge invariance. Yet, in many ways, this is a strange idea on which to rest our most important theories of physics. Gauge invariance is, at heart, merely an ambiguity in how we choose to present the laws of physics. Why should it play such an important role?

One reason is that the ambiguity allows us to demonstrate various properties that we care about but which, naively, might appear incompatible. These properties include Lorentz invariance and locality and, in the quantum theory, unitarity. We already got a glimpse of this in the lectures on Quantum Field Theory when we quantised Maxwell theory. One choice of gauge makes unitarity manifest while another makes Lorentz invariance manifest. The gauge ambiguity allows us to flit from one choice to another, allowing us to both have our cake and eat it.

Relatedly, we know that the photon has two polarisation states. But try writing down a field which describes the photon that has only two indices and which transforms nicely under the $SO(3,1)$ Lorentz group; its not possible. So instead we introduce the field $A_\mu$ which makes Lorentz invariance manifest and then use the gauge symmetry to kill two of four resulting states.

The physical information in $A_\mu$ can be found in the *field strength*

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \; . \tag{1.61}$$

The field strength is invariant under the gauge transformation (1.60). The field strength houses the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$. If we write $A^\mu = (\phi, \mathbf{A})$, then we have

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A} \; . \tag{1.62}$$

The dynamics of the gauge field is described by the action

$$S_{\text{Maxwell}} = -\frac{1}{4}\int d^4x \; F_{\mu\nu}F^{\mu\nu} \; . \tag{1.63}$$

The resulting equations of motion are

$$\partial_\mu F^{\mu\nu} = 0 \; . \tag{1.64}$$

This coincides with two of the Maxwell equations: Gauss' law $\nabla \cdot \mathbf{E} = 0$ and Ampère's law $\nabla \times \mathbf{B} = \partial \mathbf{E}/\partial t$. The other two follow immediately from constructing $F_{\mu\nu}$ in terms of the gauge potential. To see this, we first introduce the *dual field strength*

$$^\star F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \; . \tag{1.65}$$

This is similar to $F_{\mu\nu}$, but with $\mathbf{E}$ and $\mathbf{B}$ swapped (one of them with a minus sign). Then, by the anti-symmetry of $\epsilon^{\mu\nu\rho\sigma}$, together with the definition (1.61), we have the Bianchi identity

$$\partial_\mu {}^\star F^{\mu\nu} = 0 \ . \tag{1.66}$$

Expanding this out gives the remaining two Maxwell equations: the one that says magnetic monopoles don't exist $\nabla\cdot\mathbf{B} = 0$, and the law of induction $\nabla\times\mathbf{E}+\partial\mathbf{B}/\partial t = 0$.

The necessity to keep gauge invariance means that it's not possible to augment the action (1.63) with a mass term of the form $m^2 A_\mu A^\mu$. This would break gauge invariance and cause trouble down the line. Naively, this would appear to guarantee that the photon must always be massless. In fact, there is a way to give the photon a mass, known as the *Higgs mechanism*. This will be discussed in Section 2.3.

### Coupling to Matter

Underlying electromagnetism is a $U(1)$ gauge group. That's not so obvious in the description above, where the "symmetry" (really redundancy) manifests itself only as a shift of the gauge field (1.60) depending on a function $\alpha(x)$. However, the $U(1)$ness of electromagnetism becomes more apparent when we couple to charged fields.

Fields that are charged under electromagnetism are necessarily complex. Consider, for example, a complex scalar field $\phi(x)$ of charge $e$. When the gauge field transforms as (1.60), the scalar field has a corresponding transformation

$$\phi \to e^{ie\alpha}\phi \ . \tag{1.67}$$

Here we see the group emerging more clearly, with $e^{ie\alpha(x)} \in U(1)$. Because the transformation parameter $\alpha(x)$ is a function, we really have a $U(1)$ symmetry/redundancy for each point $x$ in space. This is what it means to have a $U(1)$ "gauge group": it is a much larger group than the global symmetries that appear elsewhere.

We can construct theories that are invariant under the transformation (1.67) by replacing partial derivatives with the covariant derivative

$$\mathcal{D}_\mu\phi = \partial_\mu\phi - ieA_\mu\phi \ . \tag{1.68}$$

This has the nice property that $\mathcal{D}_\mu\phi$ transforms covariantly under a gauge transformation, a fact that requires a couple of quick lines of calculation:

$$\begin{aligned} \mathcal{D}_\mu\phi &\to (\partial_\mu - ieA_\mu - ie\partial_\mu\alpha)\, e^{ie\alpha}\phi \\ &= e^{ie\alpha}\,(\partial_\mu - ieA_\mu)\,\phi \\ &= e^{ie\alpha}\mathcal{D}_\mu\phi \ . \end{aligned} \tag{1.69}$$

The key to this calculation is that the derivative hitting $\partial_\mu(e^{ie\alpha})$ exactly cancels the shift of the gauge field (1.60). Taking the complex conjugate of (1.68), we have

$$\mathcal{D}_\mu\phi^\dagger = (\partial_\mu + ieA_\mu)\phi^\dagger \; . \tag{1.70}$$

From this, we see that the meaning of the covariant derivative $\mathcal{D}_\mu$ depends on the object it's hitting: it's $-ieA_\mu$ for the scalar in (1.68), but $+ieA_\mu$ for the conjugate scalar in (1.70). You can check that, under a gauge transformation, $\mathcal{D}_\mu\phi^\dagger \rightarrow e^{-ie\alpha}\mathcal{D}_\mu\phi^\dagger$. This ensures that we can form a gauge invariant action

$$S_{\text{scalar}} = \int d^4x \, \Big(\mathcal{D}_\mu\phi^\dagger\mathcal{D}^\mu\phi - V(|\phi|)\Big) \tag{1.71}$$

where we take the potential to depend only on $|\phi|^2 = \phi^\dagger\phi$. In particular, this means that we disallow terms in the potential of the form $\phi^2 + \phi^{\dagger\,2}$ which are real but are not gauge invariant.

If we have multiple scalar fields, then they can carry different charges. When the gauge group is $U(1)$, these charges should be integer multiples of each other, meaning that each field transforms as

$$\phi \rightarrow e^{ieq\alpha}\phi \quad \text{with} \quad q \in \mathbb{Z} \; . \tag{1.72}$$

It is possible to write down theories in which the charges $q$ are not integer valued. (For example, one could imagine one scalar field with $q = 1$ and another with $q = \sqrt{2}$.) Strictly, the gauge group should be viewed as $\mathbb{R}$ in this case, rather than $U(1)$. The differences between a $U(1)$ gauge group and an $\mathbb{R}$ gauge group are rather subtle, and manifest themselves only in the presence of magnetic monopoles, or in spacetimes of non-trivial topology. We won't get into these issues here.

Everything that we've said above for scalars also holds for fermions, both Weyl and Dirac. In either case, we replace the partial derivatives in the relevant action (either (1.59) or (1.58)) with covariant derivatives and off we go.

### 1.3.2 A Refresher on Lie Algebras

There is an important extension of Maxwell theory in which the gauge group $U(1)$ is replaced by a compact Lie group $G$. Here we give a lightening review of the relevant aspects of Lie groups and Lie algebras.

A Lie group is a group that is also a differentiable manifold[1]. This means, among other things, that a group element is labelled by some continuous parameters. We've already met examples of Lie groups in both the rotation group and the Poincaré group.

Lie groups have the property that, for elements continuously connected to the identity, we can write each $U \in G$ as

$$U = e^{i\theta^A T^A} \tag{1.73}$$

Here the $\theta^A$ are just numbers that tell us which group element we're working with, while the $T^A$ are generators of the group. If you like, the $T^a$ tell us the infinitesimal action of the group, with $g \approx \mathbb{1} + i\theta^A T^A + \mathcal{O}(\theta^2)$ when $\theta$ is small. A general group element (1.73) can then be constructed by exponentiating the infinitesimal action.

It turns out that, with the exception of some global information, the structure of the Lie group is captured in the behaviour of those infinitesimal generators $T^A$. They form the associated Lie algebra $\mathfrak{g}$, given by

$$[T^A, T^B] = i f^{ABC} T^C \ . \tag{1.74}$$

Here $A, B, C = 1, \ldots, \dim G$ and $f^{ABC}$ are the fully anti-symmetric structure constants which distill the information about the group $G$. The factor of $i$ on the right-hand side is taken to ensure that the generators are Hermitian: $(T^A)^\dagger = T^A$.

(Mathematicians usually prefer the convention where there is no $i$ on the right-hand side and the generators are anti-Hermitian, largely because there are examples like $SO(N)$ where everything in the game is real and a factor of $i$ makes things needlessly complex. In contrast, physicists tend to include the factor of $i$ on the right-hand side because they're usually working in the realm of quantum mechanics where things will ultimately become complex anyway.)

The $T^A$ in (1.74) are abstract objects but we will shortly want to identify them with matrices. This means, among other things, that we want the commutator in (1.74) to have the same properties as matrix commutation, among them the Jacobi identity

$$[T^A, [T^B, T^C]] + [T^B, [T^C, T^A]] + [T^C, [T^A, T^B]] = 0 \ . \tag{1.75}$$

This puts constraints on the structure constants $f^{abc}$ which must, in turn, obey

$$f^{ADE} f^{BCD} + f^{BDE} f^{CAD} + f^{CDE} f^{ABD} = 0 \ . \tag{1.76}$$

[1]For many physicists, Lie groups are the only groups they know. A mathematician friend of mine told me that a physicist's definition of a finite group is a Lie group without manifold structure.

| $G$ | $SU(N)$ | $SO(N)$ | $Sp(N)$ | $E_6$ | $E_7$ | $E_8$ | $F_4$ | $G_2$ |
|---|---|---|---|---|---|---|---|---|
| $\dim G$ | $N^2-1$ | $\frac{1}{2}N(N-1)$ | $N(2N+1)$ | 78 | 133 | 248 | 52 | 14 |
| $\dim F$ | $N$ | $N$ | $2N$ | 27 | 56 | 248 | 6 | 7 |

**Table 2**. The classification of compact, semi-simple Lie algebras $G$, together with their dimension and the dimension of the fundamental representation $F$.

We will be interested in *simple*, *compact* Lie groups. Here "simple" means that we don't have any trivial $U(1)$ factors floating around that commute with everything else. We can always include such factors if we wish (and we will wish for the Standard Model) but we'll be best served if we ignore them at this stage. Meanwhile, "compact" means that if you continue to rotate in the group then you ultimately come back to where you started from (or close to where you started from). For example, the group of rotations is compact, while the Lorentz group is non-compact because if you keep boosting in a given direction then you just move faster and faster.

There is a classification of simple compact Lie algebras. The possible options for the group $G$, together with the dimension of the group, are shown in Table 2[2]. All of these groups are referred to as *non-Abelian* meaning that things don't commute with each other. In contrast, $U(1)$ is an Abelian group.

As we mentioned above, the $T^A$ in (1.74) are initially viewed as just abstract objects. But it's interesting to ask when they can take a more concrete form in the guise of matrices. These are the *representations* of the algebra. For each algebra $G$, there is an infinite list of numbers which are the dimensions of the matrices that can be used to represent $G$. The smallest such matrix is called the *fundamental representation* and we will denote it as $F$. The dimension of $F$ for each Lie group $G$ are also shown in Table 2.

In what follows, we will (with a slight abuse of notation) use $T^A$ to refer to the generators fundamental representation. When we have occasion to use other representations $R$, we will refer to the generators as $T^A(R)$ (In later sections, we'll also refer to these as $T^A_R$.). In fact, for the Standard Model we will only need two different representations: the fundamental and the adjoint. The adjoint is a representation that has dimension

[2] We're using the convention $Sp(1) = SU(2)$. Other authors sometimes write $Sp(2N)$, or even $USp(2N)$ to refer to what we've called $Sp(N)$, preferring the argument to refer to the dimension of the fundamental representation $F$ rather than the rank of the Lie algebra $\mathfrak{g}$.

$\dim(\text{adj}) = \dim G$ with the generators given by

$$T^A(\text{adj})_{BC} = -if^{ABC} \ . \tag{1.77}$$

Don't be lulled into thinking that you don't need to consider other representations: they will appear in other situations, including when we discuss flavour symmetry in QCD in Section 3.

The Lie algebra comes with what, in fancy language, is called a Killing form. But, by the time we're thinking about matrices, this Killing form is just the trace. The generators of any simple Lie algebra obey $\text{Tr}\, T^A = 0$. (This is what it means for the Lie algebra to be "simple".) We take the generators in the fundamental representation $F$ to satisfy

$$\text{Tr}\, T^A T^B = \frac{1}{2}\delta^{AB} \tag{1.78}$$

This can be viewed as tantamount to fixing the normalisation of the structure constants $f^{ABC}$. Having fixed the normalisation in the fundamental representation, other representations $T^A(R)$ will have different normalisations.

Before we proceed, an example. The simplest non-Abelian Lie group is $SU(2)$, which has $\dim(SU(2)) = 3$ and structure constants given by $f^{ABC} = \epsilon^{ABC}$. In this case, the fundamental representation is (up to an overall normalisation) the $2 \times 2$ Pauli matrices

$$T^A = \frac{1}{2}\sigma^A \ . \tag{1.79}$$

These indeed obey $[T^A, T^B] = i\epsilon^{ABC}T^C$, together with the normalisation condition (1.78).

The group $SU(3)$ also plays a prominent role in the Standard Model. (In fact, as we will see, it plays two prominent roles!) We will describe the structure constants and the generators in Section 3.

### 1.3.3 Yang-Mills Theory

Now we can turn to some physics. Yang-Mills theory is a generalisation of Maxwell theory in which the group $U(1)$ is replaced by a simple, compact Lie algebra $G$. To specify the Yang-Mills theory, we need only specify the choice of $G$ together with a coupling constant $g > 0$ that will dictate the strength of the interactions. (The coupling constant $g$ plays the same role as the charge $e$ in Maxwell theory. As we will later see, the phrase "coupling constant" is not particularly accurate because it will turn out not to be constant!)

For each element of the algebra, we introduce a gauge field $A_\mu^A$ with $A = 1, \ldots, \dim G$. These are then packaged into the Lie algebra-valued gauge potential

$$A_\mu = A_\mu^A T^A \tag{1.80}$$

A down-to-earth perspective is to think of the $T^A$ as matrices in the fundamental representation. This means, for example, that for $G = SU(N)$, the gauge potential $A_\mu$ is a 4-vector where each component is a traceless $N \times N$ matrix.

The fields $A_\mu^A$ are collectively referred to as *gauge bosons*. (They have other, more specific, names in the Standard Model when we apply these ideas to the two nuclear forces.) As in Maxwell theory, not all the information in $A_\mu$ is physical and any two field configurations related by a gauge transformation should be viewed as equivalent. This time, however, the gauge transformation is a little more intricate.

The action of the gauge symmetry is associated to a Lie group valued function over spacetime,

$$\Omega(x) \in G\ . \tag{1.81}$$

The set of all such transformations is known as the *gauge group*. As in Maxwell theory, we will sometimes be sloppy and refer to the Lie group $G$ as the gauge group, but strictly speaking it is the much bigger group of maps from spacetime into $G$. The action on the gauge field is

$$A_\mu \to \Omega A_\mu \Omega^{-1} + \frac{i}{g}\Omega\, \partial_\mu \Omega^{-1}\ . \tag{1.82}$$

The first term is the expected transformation for an adjoint-valued field. The second, inhomogeneous, term is an additional piece that is characteristic of gauge transformations.

To make contact with gauge transformations in electromagnetism, suppose that we have $G = U(1)$ and write $\Omega(x) = e^{ie\alpha(x)}$. Then, using the fact that everything commutes, we have

$$\Omega A_\mu \Omega^{-1} + \frac{i}{e}\Omega \partial_\mu \Omega^{-1} = A_\mu + \partial_\mu \alpha \tag{1.83}$$

and the gauge transformation (1.82) reproduces the familiar gauge transformation of Maxwell theory.

As in Maxwell theory, we can construct a field strength. Here too there is an extra ingredient arising from the fact that $A_\mu$ is a matrix and the generalisation of (1.61) is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \ . \tag{1.84}$$

In contrast to Maxwell theory, the field strength includes a non-linear term, proportional to the coupling $g$. This will prove to be important: it is this non-linear term that makes Yang-Mills theory significantly richer and more interesting than Maxwell theory. Like $A_\mu$, the field strength is a Lie algebra-valued field and we could also expand it as $F_{\mu\nu} = F^A_{\mu\nu} T^A$.

So far, I've not explained *why* (1.84) is the right field strength. The main reason is that it transforms nicely under the gauge transformation (1.82)

$$F_{\mu\nu} \to \Omega\, F_{\mu\nu}\, \Omega^{-1} \ . \tag{1.85}$$

To see this, you could just plug (1.82) into (1.84) but it's mildly laborious; we will offer a shortcut to this result presently.

The transformation (1.85) means that, in contrast to electromagnetism, the Yang-Mills "electric field" $E_i = F_{0i}$ and "magnetic field" $B_i = -\frac{1}{2}\epsilon_{ijk}F_{jk}$ are *not* gauge invariant. To construct something physical, you can multiply together some number of $E_i$ and $B_j$ and then take the trace, which ensures that the $\Omega$ and $\Omega^{-1}$ in (1.85) cancel and you get something gauge invariant. (You need something that is at least quadratic in $F_{\mu\nu}$ because, for simple Lie groups, $\mathrm{Tr}\, F_{\mu\nu} = 0$.)

The gauge transformations above involve the Lie group valued object $\Omega(x)$. But one of the key properties of Lie groups is that their structure is largely determined by the elements that are infinitesimally close to the identity. This suggests that it's fruitful to look at gauge transformations that are everywhere close to the identity. These can be written as

$$\Omega(x) \approx 1 + ig\alpha^A(x)T^A + \ldots \tag{1.86}$$

where the $\alpha^a$ are taken to be everywhere small. From (1.82), the infinitesimal transformation of the gauge field is $A_\mu \to A_\mu + \delta A_\mu$ with

$$\delta A_\mu = \partial_\mu \alpha - ig[A_\mu, \alpha] \tag{1.87}$$

where $\alpha = \alpha^a T^a$ is the Lie algebra-valued infinitesimal transformation. It's convenient to write this as $\delta A_\mu = \mathcal{D}_\mu \alpha$ where the covariant derivative is defined to be

$$\mathcal{D}_\mu \alpha = \partial_\mu \alpha - ig[A_\mu, \alpha] \ . \tag{1.88}$$

This is the covariant derivative acting on the Lie algebra-valued (i.e. adjoint) field $\alpha$. We'll soon see different covariant derivatives acting on other representations.

Now we can check how infinitesimal gauge transformations act on the field strength (1.84). We have

$$
\begin{aligned}
\delta F_{\mu\nu} &= \partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu - ig[A_\mu, \delta A_\nu] - ig[\delta A_\mu, A_\nu] \\
&= \mathcal{D}_\mu \delta A_\nu - \mathcal{D}_\nu \delta A_\mu \\
&= [\mathcal{D}_\mu, \mathcal{D}_\nu]\alpha \ .
\end{aligned}
\tag{1.89}
$$

We see that we're left with the task of computing the commutator of two covariant derivatives, acting on the adjoint field $\alpha$. This is a worthwhile and straightforward, calculation. We have

$$
[\mathcal{D}_\mu, \mathcal{D}_\nu]\alpha = -ig[F_{\mu\nu}, \alpha] \ . \tag{1.90}
$$

This gives $\delta F_{\mu\nu} = ig[\alpha, F_{\mu\nu}]$ which is indeed the expected infinitesimal gauge transformation arising from (1.85).

### The Yang-Mills Action

The dynamics of the Yang-Mills field is the obvious generalisation of the Maxwell action,

$$
S_{\rm YM} = -\frac{1}{2}\int d^4x \ \mathrm{Tr}\, F^{\mu\nu}F_{\mu\nu} \ . \tag{1.91}
$$

Naively, the only difference lies in that overall trace, which ensures that the action is invariant under gauge transformations (1.85). This also accounts for the overall normalisation of the action, which comes with a factor of 1/2 rather than the 1/4 seen in (1.63) because an additional factor of 1/2 comes from the trace in (1.78). This means that the Yang-Mills and Maxwell action come with the same normalisation.

However, the key difference between the two actions is buried in our notation: while the Maxwell action is quadratic in $A_\mu$, the Yang-Mills action includes terms that are cubic and quartic in $A_\mu$, both coming from the commutator in the definition of the field strength (1.84).

The classical equations of motion are derived by minimizing the action with respect to each gauge field $A^a_\mu$. It is a simple exercise to check that they are given by

$$
\mathcal{D}_\mu F^{\mu\nu} = 0 \ . \tag{1.92}
$$

Here the covariant derivative is defined as in (1.88): $\mathcal{D}_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu} - ig[A_\mu, F^{\mu\nu}]$. These are the *Yang-Mills equations*. In contrast to the Maxwell equations, they are non-linear. This means that the Yang-Mills fields interact with themselves.

There is also a Bianchi identity that follows from the definition (1.84) of $F_{\mu\nu}$ in terms of the gauge field. This is best expressed by first introducing the dual field strength

$$ {}^{\star}F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \ . \tag{1.93} $$

and noting that this obeys the identity

$$ \mathcal{D}_\mu {}^{\star}F^{\mu\nu} = 0 \ . \tag{1.94} $$

Both (1.92) and (1.94) are non-linear equations. However, the non-linearities come in the form of commutators like $[A_\mu, A_\nu]$. This means that if we focus on field configurations that sit purely with a subgroup $U(1) \subset G$, then the commutators vanish and the equations reduce to those of Maxwell theory. So although the general solutions to the Yang-Mills equations are surely complicated, we can always import any solution to Maxwell theory and embed it in some $U(1)$. In particular, Yang-Mills theory admits solutions akin to electromagnetic waves that travel at the speed of light.

Although we can always embed solutions of Maxwell theory in the Yang-Mills field, there's nothing that tells us that these solutions are stable. For that, one has to work harder and look at fluctuations of the other fields that do not live in your favourite $U(1)$. (For what it's worth, a constant electric field is stable in Yang-Mills theory, while a constant magnetic field is *unstable*.) We won't discuss these stability issues further in these lectures, largely because our interest lies in what happens in quantum Yang-Mills rather than in the classical theory.

Just as for Maxwell theory, the need to keep gauge invariance means that we can't add a mass term like $A_\mu A^\mu$ or $\mathrm{Tr}\, A_\mu A^\mu$ to the action (1.91). This strongly suggests that quantum Yang-Mills is, like Maxwell theory, a theory of massless particles. This strong suggestion is, it turns out, completely wrong! When we quantise the Yang-Mills action (1.91), we find a theory of interacting massive particles, rather than massless particles. The reason for this can be traced to the interaction terms in Yang-Mills, but is not fully understood. Indeed, proving it from first principles remains one of the most important open problems in mathematical physics. We will discuss this further in section 3.

**Coupling to Matter**

As with electromagnetism, we can couple the Yang-Mills field to matter. We do this by requiring that the matter fields live in some representation $R$ of the gauge group. This means that the matter fields come in some vector of dimension $\dim R$.

For each such representation, we have generators $T^A(R)$ which we can can think of as square matrices of dimension $\dim R$. Dressed resplendent in all their indices, they take the form

$$T^A(R)^a{}_b \quad \text{with } 'a, b = 1, \ldots, \dim R \;\; \text{and} \;\; A = 1 \ldots, \dim G \;. \tag{1.95}$$

Consider a scalar field in the representation $R$. Under a gauge transformation $\Omega(x) = e^{ig\alpha^A(x)T^A}$, the scalar transforms as

$$\phi^a \to (\Omega_R)^a{}_b \phi^b \quad \text{with} \quad (\Omega_R)^a{}_b = \Big( \exp\big(ig\alpha^A T^A(R)\big)\Big)^a{}_b \;. \tag{1.96}$$

Some representations $R$ are real, and some are complex. For example, the fundamental representation of $SU(N)$ is complex, and so $\phi$ must be a complex $N$-dimensional vector. Meanwhile, the adjoint representation of any group $G$ is always real and, correspondingly, $\phi$ can be real.

To write down an action for $\phi$ that is invariant under the gauge transformation (1.96), we follow our Maxwellian noses and construct the covariant derivative,

$$\mathcal{D}\phi^a = \partial_\mu \phi^a - igA^A_\mu \, T^A(R)^a{}_b \phi^b \;. \tag{1.97}$$

Under a gauge transformation, this covariant derivative transforms, as the name suggests, covariantly, meaning

$$\mathcal{D}_\mu \phi^a \to (\Omega_R)^a{}_b \mathcal{D}_\mu \phi^b \;. \tag{1.98}$$

We will later see that all matter fields in the Standard Model transform in the fundamental representation. For $SU(N)$, this means that we can think of $\phi^a$ as an $N$-component complex vector, with $a = 1, \ldots, N$, and write the covariant derivative in terms of the $N \times N$ matrix-valued gauge field $A_\mu = A^A_\mu T^A$,

$$\mathcal{D}_\mu \phi^a = \partial_\mu \phi^a - ig(A_\mu)^a{}_b \phi^b \;. \tag{1.99}$$

This expression differs from our previous covariant derivative (1.88) because $\phi$ is in the fundamental representation, while $\alpha$ in (1.88) was in the adjoint. This highlights something we've stressed previously: the meaning of the covariant derivative depends on the representation of the object on which it acts. Once again, covariant derivatives do not commute. This time, for covariant derivatives acting on fundamental fields, we find

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = -igF_{\mu\nu} \;. \tag{1.100}$$

This should be compared to the analogous result (1.90) for covariant derivatives acting on adjoint-valued fields.

As before, it's useful to check some of the formulae for infinitesimal gauge transformations. We have $\delta A_\mu = \mathcal{D}_\mu \alpha$, as in (1.87) and, from (1.96), $\delta\phi = ig\alpha\phi$. Then, suppressing the $a = 1, \ldots, N$ index, the covariant derivative (1.99) transforms as

$$\begin{aligned}
\delta(\mathcal{D}_\mu \phi) &= \partial_\mu \delta\phi - ig\,\delta A_\mu \phi - igA_\mu\,\delta\phi \\
&= ig\partial_\mu(\alpha\phi) - ig(\mathcal{D}_\mu \alpha)\phi + g^2 A_\mu \alpha\phi \\
&= ig\alpha\left(\partial_\mu \phi - igA_\mu \phi\right) \\
&= ig\alpha\mathcal{D}_\mu \phi.
\end{aligned} \tag{1.101}$$

This is, indeed, the infinitesimal version of the gauge transformation (1.98).

With covariant derivatives that transform nicely, it's straightforward to write down an action for the matter fields. As in electromagnetism, we just need to replace the partial derivatives in the action with covariant derivatives and we have something gauge invariant. This holds for scalars, Weyl fermions, and Dirac fermions.

#### A Rescaling

Above we've written the action so that the coupling constant $g$ multiplies the non-linear terms. This means, in particular, that it makes an appearance in the field strength (1.84). It also appears, perhaps rather strangely, as the inverse $1/g$ in the gauge transformation (1.82).

There is a different way to normalise the gauge field that, for many purposes, turns out to be more natural. We define the new gauge field

$$\tilde{A}_\mu = gA_\mu \quad \text{and} \quad \tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu - i[\tilde{A}_\mu, \tilde{A}_\nu]\ . \tag{1.102}$$

We also define the rescaled gauge parameter $\tilde{\alpha} = g\alpha$, so that the group element is $\Omega = e^{i\tilde{\alpha}}$. This then eliminates the gauge coupling from all kinematic quantities like the field strength and covariant derivatives. The only place that the coupling shows up is in an overall coefficient multiplying the entire action,

$$S_{\text{YM}} = -\frac{1}{2}\int d^4x\ \text{Tr}\, F^{\mu\nu}F_{\mu\nu} = -\frac{1}{2g^2}\int d^4x\ \text{Tr}\,\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu} \tag{1.103}$$

In the first way of writing things, the coupling constant $g$ sits in front of the non-linear terms, making it clear that it governs the strength of interactions. But it also governs the strength of interactions in the second way of writing things. To see this, note that in the Euclidean path integral, we sum over all field configurations weighted by $e^{-S/\hbar}$. With the rescaling above, $g^2$ sits in the same place in the action as $\hbar$, which suggests that $g^2 \to 0$ will be a classical limit. Heuristically you should think that, for $g^2$ small, we pay a large price for field configurations that do not minimize the action; in this way, the path integral is dominated by the classical configurations. In contrast, when $g^2 \to \infty$, the Yang-Mills action disappears completely. This is the strong coupling regime, where all field configurations are unsuppressed and contribute equally to the path integral.

### The Analogy with General Relativity

General Relativity is rightly lauded for the way it places geometry into the heart of physics. But the other laws of physics, which combine to form the Standard Model, are no less geometrical. Rather than arising from the geometry of spacetime, they instead arise from a slightly more subtle object known as a *fibre bundle*.

We won't describe the mathematics of fibre bundles in any detail in these lectures, but will instead just point out some analogies between the gauge theories discussed above and the differential geometry that underlies general relativity.

One of the key ideas in general relativity is *diffeomorphism invariance*. This is the statement that physical quantities should not depend on the coordinates that we choose to describe them. Such coordinate transformations are analogous to gauge transformations in Yang-Mills theory.

One of the most important objects in general relativity is the Levi-Civita connection $\Gamma^\mu_{\rho\nu}$. Famously, this is *not* a tensor. Under a coordinate transformation $x \to \tilde{x}$, with

$$\Omega^\mu{}_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \; , \tag{1.104}$$

the Levi-Civita connection transforms as

$$\Gamma^\mu_{\rho\nu} \to (\Omega^{-1})^\mu{}_\tau \Omega^\sigma{}_\rho \Omega^\lambda{}_\nu \Gamma^\tau{}_{\sigma\lambda} + (\Omega^{-1})\Omega^\sigma{}_\rho \partial_\sigma \Omega^\tau{}_\nu \; . \tag{1.105}$$

The first term is how a tensor would transform. The second term is independent of $\Gamma$ and is the characteristic transformation of a connection. But this looks very similar to the transformation of the gauge field (1.82),

$$A_\mu \to \Omega A_\mu \Omega^{-1} + \frac{i}{g}\Omega\, \partial_\mu \Omega^{-1} \tag{1.106}$$

where, again, there is a transformation that befits a tensor, supplemented with the additional derivative term $\partial\Omega$. Indeed, this analogy can be made more precise, and mathematicians refer to the gauge field $A_\mu$ as a *connection*. Both connections find their natural home inside covariant derivatives. In gauge theory, this is the $\mathcal{D}_\mu$ that we've already met, while in general relativity it is the object that acts naturally on vector fields $Y$, with $(\nabla_\nu Y)^\mu = \partial_\nu Y^\mu + \Gamma^\mu_{\nu\rho} Y^\rho$ and is then extended to act on other tensor fields.

Given a Levi-Civita connection, one can construct the Riemann curvature tensor $R^\sigma_{\rho\mu\nu}$. Rearranging some of the indices this can be written as

$$(R_{\mu\nu})^\sigma{}_\rho = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho}\Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho}\Gamma^\sigma_{\nu\lambda} \ . \tag{1.107}$$

Again, we see an immediate similarity with the construction of the field strength in Yang-Mills (1.84) which, including the $a, b = 1, \ldots, \dim F$ indices, reads

$$(F_{\mu\nu})^a{}_b = \partial_\mu (A_\nu)^a{}_b - \partial_\nu (A_\mu)^a{}_b - ig(A_\mu)^a{}_c (A_\nu)^c{}_b + ig(A_\nu)^a{}_c (A_\mu)^c{}_b \ . \tag{1.108}$$

Mathematicians refer to both the Riemann tensor and the field strength $F_{\mu\nu}$ as the *curvature*.

## 1.4 C,P, and T

Discrete symmetries play a crucial role in understanding the structure of the Standard Model. There are three that are particularly important: parity, charge conjugation, and time reversal. In this section, we describe each of these in turn. We end by explaining why the combination of all three is necessarily a symmetry of any local, relativistic quantum field theory.

### 1.4.1 Parity

Parity is an inversion of the spatial coordinates,

$$P : (t, \mathbf{x}) \mapsto (t, -\mathbf{x}) \ . \tag{1.109}$$

This can be viewed as a Lorentz transformation, but not one that is continuously connected to the identity. Roughly speaking the action of parity mimics what a system looks like reflected in the mirror. More precisely, a reflection is implemented by, say, $R : (x, y, z) \mapsto (x, y, -z)$. The parity transformation (1.109), which is a reflection followed by a rotation by 180°, has the advantage that it treats all spatial coordinates on the same footing.

(As an aside: one disadvantage of the parity transformation $P : \mathbf{x} \mapsto -\mathbf{x}$ is that it only works when the number of spatial dimensions is odd. For example, in $d = 2 + 1$ dimensions, the transformation $(x, y) \mapsto (-x, -y)$ is just a rotation by 180°. For this reason, if you're discussing quantum field theories in different dimensions, it's better to talk about reflections which flip the sign of just one spatial direction, rather than parity which flips all of them. In these lectures, we've got no interest in dimension hopping: our interest is strictly in the Standard Model and so we keep with the conventional definition of parity (1.109).)

We would like to understand the circumstances under which a quantum field theory is invariant under parity, and how the fields transform. When we come to discuss the weak force in Section 5, we will find that the laws of our universe are *not* invariant under parity. This is a shocking statement. It means that given a solution to the equations of motion, the parity reflected evolution is *not* a solution!

First, let's ask how electromagnetic fields transform under parity. For this, we can look at the covariant derivative which, regardless of the object it acts on, takes the schematic form

$$\mathcal{D}_\mu = \partial_\mu - iA_\mu \ . \tag{1.110}$$

This ties the behaviour of the gauge field to that of the derivative. Under a parity transformation $\partial_0$ is left unaffected, while the spatial derivatives $\partial_i$ change sign. This tells us that parity must act as

$$P : A_0(t, \mathbf{x}) \mapsto +A_0(t, -\mathbf{x}) \quad \text{and} \quad P : A_i(t, \mathbf{x}) \mapsto -A_i(t, -\mathbf{x}) \ . \tag{1.111}$$

Tracing this through to the definitions of the electric field $\mathbf{E} = -\nabla\phi - \partial\mathbf{A}/\partial t$ and magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$, we have

$$P : \mathbf{E}(t, \mathbf{x}) \mapsto -\mathbf{E}(t, -\mathbf{x}) \quad \text{and} \quad P : \mathbf{B}(t, \mathbf{x}) \mapsto +\mathbf{B}(t, -\mathbf{x}) \ . \tag{1.112}$$

Vectors like $\mathbf{E}$, which transform under parity in the same way as $\mathbf{x}$ are deemed worthy to keep the name "vector". Meanwhile, vectors like $\mathbf{B}$ which don't pick up a minus sign under parity are said to be *pseudovectors*. The most familiar examples of pseudovectors are the magnetic field and angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. These are also the two kinds of vectors that exhibit the most counterintuitive behaviour when we're undergraduates. This is not a coincidence.

In the quantum theory, the parity transformation is enacted by a unitary operator on the Hilbert space that we also call $P$. The fields $A_\mu(x)$ are now also operators and the transformation (1.111) becomes

$$P A_0(t,\mathbf{x}) P^\dagger = A_0(t,-\mathbf{x}) \quad \text{and} \quad P A_i(t,\mathbf{x}) P^\dagger = -A_i(t,-\mathbf{x}) \ . \tag{1.113}$$

In what follows, we will flit between the description of parity and other discrete symmetries as a map, as in (1.111), and as an operator acting on Hilbert space, as in (1.113).

Next, we turn to spinors. It can be somewhat fiddly to figure out how spinors transform under various discrete symmetries, but it's a topic that will play a crucial role as we proceed. The equations of motion for a left-handed massless Weyl spinor $\psi_L$ is

$$\bar{\sigma}^\mu \partial_\mu \psi_L = 0 \tag{1.114}$$

where $\bar{\sigma} = (\mathbb{1}, -\sigma^i)$. Under a parity transformation, the spatial derivative changes sign and the Weyl equation (1.114) is not invariant. This is important: if we have just a single left-handed Weyl spinor $\psi_L$ then this theory is *not* invariant under parity.

We can rescue the situation if, in addition to our left-handed Weyl spinor $\psi_L$, we also have a right-handed Weyl spinor $\psi_R$. This obeys the equation of motion

$$\sigma^\mu \partial_\mu \psi_R = 0 \tag{1.115}$$

where $\sigma^\mu = (\mathbb{1}, \sigma^i)$. The different minus signs in $\sigma^\mu$ and $\bar{\sigma}^\mu$ mean that we can compensate for a parity transformation if we also exchange left- and right-handed spinors, so that

$$P\psi_L(t,\mathbf{x})P^\dagger = \psi_R(t,-\mathbf{x}) \quad \text{and} \quad P\psi_R(t,\mathbf{x})P^\dagger = \psi_L(t,-\mathbf{x}) \ . \tag{1.116}$$

There are also options to put different minus signs (and even phases) on the right-hand side as we describe below.

As we've seen in Section 1.2.1, the two spinors $\psi_L$ and $\psi_R$ naturally sit in a Dirac spinor $\psi = (\psi_L, \psi_R)^T$. The action of parity on Weyl spinors (1.116) translates into the action on the Dirac spinor

$$P\psi(t,\mathbf{x})P^\dagger = \gamma^0 \psi(t,-\mathbf{x}) \quad \text{with} \quad \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ . \tag{1.117}$$

In the lectures on Quantum Field Theory, we saw that a stationary fermion is associated to a solution to the Dirac equation, where the spinor degrees of freedom take the form $\psi = (\xi, \xi)^T$. Here $\xi$ is some 2-component spinor the tells us the orientation of the spin of the particle. Meanwhile, the solution corresponding to an anti-fermion takes the form $\psi = (\xi, -\xi)^T$. This means that the fermion has intrinsic parity $+1$ while the anti-fermion has intrinsic parity $-1$.

Terms in the action are always constructed out of an even number of fermions. Given the transformation (1.117), we can look at the fate of various fermion bilinears under parity. You can check, for example, that

$$P : \bar{\psi}\psi \mapsto \bar{\psi}\psi \quad \text{and} \quad P : \bar{\psi}\gamma^5\psi \mapsto -\bar{\psi}\gamma^5\psi \tag{1.118}$$

where we've suppressed the all-important spinor indices. We say that $\bar{\psi}\psi$ transforms as a *scalar* while $\bar{\psi}\gamma^5\psi$ transforms as a *pseudoscalar*. Similarly, you can check that $\bar{\psi}\gamma^\mu\psi$ is a vector while $\bar{\psi}\gamma^5\gamma^\mu\psi$ is a pseudovector.

You shouldn't be too dogmatic about insisting that (1.116) and (1.117) are the definitive action of parity. Suppose that you have a Dirac fermion with action

$$S = \int d^4x \, \left( i\bar{\psi}\gamma^\mu\partial_\mu\psi - M\bar{\psi}\psi \right) . \tag{1.119}$$

Then this is invariant under parity with the transformation (1.117). Suppose, in contrast, that you're given the action

$$S = \int d^4x \, \left( i\bar{\psi}\gamma^\mu\partial_\mu\psi - M\bar{\psi}\gamma^5\psi \right) . \tag{1.120}$$

This is not invariant under (1.117) because the mass term is parity odd. Nonetheless, that doesn't mean that the theory doesn't have parity symmetry. We just need to look more carefully. You can check that the action (1.120) is invariant under the redefined parity transformation

$$P'\psi(t, \mathbf{x})P'^{-1} = \gamma^5\gamma^0\psi(t, -\mathbf{x}) . \tag{1.121}$$

In terms of Weyl fermions, this inserts an extra minus sign on the right-hand side of one of the transformations in (1.116). Ultimately, given a theory the aim is to find some parity transformation of the fields that leaves the action, and hence the equation of motion, invariant.

So far, we haven't discussed the action of parity on scalar fields. These are more malleable. Given a scalar field $\phi$, the kinetic terms are invariant under either

$$P\phi(t,\mathbf{x})P^\dagger = \pm\phi(t,-\mathbf{x}) \ . \qquad (1.122)$$

In other words, the kinetic terms don't distinguish between scalar (the plus sign) or pseudoscalar (the minus sign). Typically, this gets fixed when we look at the interaction of the scalar field with fermions. For example, a Yukawa term of the form $\phi\bar{\psi}\psi$ means that the scalar $\phi$ is parity even under the transformation (1.117) while a Yukawa term of the form $\phi\bar{\psi}\gamma^5\psi$ means that $\phi$ is parity odd under (1.117).

There are various pay-offs from understanding the way that parity is implemented in a theory. If a theory is invariant under parity then, as we've seen, we can assign transformation laws to the various fields. But, after quantisation, these fields give rise to particles. That means that different species of particles can be thought of as parity even or parity odd. Moreover, this concept of parity is conserved in all interactions and, like all conservation laws, this puts constraints on the kind of things that can happen.

Perhaps surprisingly, it turns out that things are even more constrained when parity is *not* a symmetry of the theory! This is for a much more subtle reason known as an *anomaly*. We will discuss this in Section 4.

### 1.4.2 Charge Conjugation

Charge conjugation is an operation that switches particles with their anti-particles. If a theory is invariant under charge conjugation, then the laws of physics that govern particles coincide with those that govern anti-particles.

This time we start with a complex scalar field $\phi$, coupled to electromagnetism. It will prove simplest to look at actions, rather than equations of motion. Charge conjugation exchanges particles and anti-particles, so we want it to act as

$$C : \phi \mapsto \pm\phi^\dagger \ . \qquad (1.123)$$

The $\pm$ ambiguity is like the ambiguity in the action of parity (1.122) and, as in that case, will typically be fixed by the interactions with other fields. In contrast, there's no ambiguity about the action on the gauge field, which is fixed by looking at the covariant derivatives, $\mathcal{D}_\mu\phi = (\partial_\mu - ieA_\mu)\phi$ and $\mathcal{D}_\mu\phi^\dagger = (\partial_\mu + ieA_\mu)\phi^\dagger$. This means that any transformation (1.123) must be accompanied by

$$C : A_\mu \mapsto -A_\mu \ . \qquad (1.124)$$

As for parity, we can also think of charge conjugation as a quantum operator $C$, in which case (1.123) and (1.124) are replaced by $C\phi C^\dagger = \pm\phi^\dagger$ and $CA_\mu C^\dagger = -A_\mu$ respectively. For non-Abelian gauge fields, charge conjugation acts as $CA_\mu C^\dagger = -A_\mu^\dagger$.

Again, the story for spinors is a little more fiddly. We'll start by looking at a Dirac spinor, rather than a Weyl spinor. The Dirac equation is

$$i\gamma^\mu(\partial_\mu - ieA_\mu)\psi - M\psi = 0\ . \tag{1.125}$$

We will look for an action of charge conjugation that transforms the spinor to

$$C : \psi \mapsto C\psi^\star\ . \tag{1.126}$$

Here $C$ on the right-hand side is a $4\times 4$ matrix that allows for the possibility that the components of the spinor get mixed up under charge conjugation. Note that we've written the transformed spinor as $\psi^\star$, rather than $\psi^\dagger$, to emphasise that it remains a "column vector" rather than a "row vector". (Of course, it's not really a vector at all. It's a spinor!)

The question is: what choice of $C$ ensures that the transformation (1.126), combined with (1.124), is a symmetry? First, we take the complex conjugate of the equation of motion (1.125):

$$-i(\gamma^\mu)^\star(\partial_\mu + ieA_\mu)\psi^\star - M\psi^\star = 0\ . \tag{1.127}$$

This is the equation that $\psi^\star$ obeys. Next, we compare this to what we get if we act with charge conjugation on the original equation (1.125):

$$\begin{aligned} & i\gamma^\mu(\partial_\mu + ieA_\mu)C\psi^\star - MC\psi^\star = 0 \\ \Longrightarrow\quad & iC^{-1}\gamma^\mu C(\partial_\mu + ieA_\mu)\psi^\star - M\psi^\star = 0\ . \end{aligned} \tag{1.128}$$

We see that (1.128) coincides with (1.127) provided that the charge conjugation matrix $C$ obeys

$$C^{-1}\gamma^\mu C = -(\gamma^\mu)^\star\ . \tag{1.129}$$

The charge conjugation matrix depends on your chosen basis of gamma matrices. For the chiral basis of gamma matrices (1.42), all gamma matrices are real except for $\gamma^2$ which is pure imaginary. This means that we should take $C = \pm i\gamma^2$, and the action of charge conjugation is

$$C : \psi \mapsto \pm i\gamma^2\psi^\star \quad \text{with} \quad \gamma^2 = \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}\ . \tag{1.130}$$

For theories that are invariant under charge conjugation, we can assign an eigenvalue $C = \pm 1$ to each particle, usually referred to as *C-parity*. As with actual parity, $P$, this new quantum number restricts the possible interactions. For example, it turns out that the neutral pion $\pi^0$ has $C = +1$ while, from (1.124), the photon necessarily has $C = -1$. This means that the decay to two photons, $\pi^0 \longrightarrow \gamma + \gamma$, is allowed (and indeed, happens over 98% of the time). But the decay to three photons, $\pi^0 \longrightarrow \gamma+\gamma+\gamma$ is forbidden on symmetry grounds.

If we decompose the Dirac fermion into its two Weyl components, $\psi = (\psi_L, \psi_R)^T$, then we can read off from (1.130) the action of charge conjugation on Weyl spinors,

$$C : \psi_L \mapsto \pm i\sigma^2 \psi_R^\star \quad \text{and} \quad C : \psi_R \mapsto \mp i\sigma^2 \psi_L^\star \ . \tag{1.131}$$

We see that charge conjugation, like parity, involves an exchange of two Weyl spinors.

A theory with just a single Weyl fermion is invariant under neither parity nor charge conjugation. However, there's still hope if we combine the two symmetries. We can take the combined action from (1.116) and (1.131) to be

$$CP : \psi_L(t, \mathbf{x}) \mapsto \mp i\sigma^2 \psi_L^\star(t, -\mathbf{x}) \quad \text{and} \quad CP : \psi_R(t, \mathbf{x}) \mapsto \pm i\sigma^2 \psi_R^\star \ . \tag{1.132}$$

A Weyl fermion coupled to a gauge field is invariant under CP. However, as we will see later, it's quite possible for this symmetry to be violated by other interaction terms (for example, Yukawa interactions between fermions and scalars).

### 1.4.3 Time Reversal

Our final discrete symmetry is time reversal, which acts on spacetime coordinates as

$$T : (t, \mathbf{x}) \mapsto (-t, \mathbf{x}) \ . \tag{1.133}$$

There's a subtlety in implementing time reversal symmetry in quantum theories. This manifests itself already in the simplest quantum mechanical systems like, say, a free particle moving in $\mathbb{R}^3$. The Schrödinger equation for the wavefunction $\Psi$ takes the form

$$i\frac{\partial \Psi}{\partial t} = -\nabla^2 \Psi \ . \tag{1.134}$$

Now compare this to the heat equation that describes how conserved quantities, such as temperature $T$, diffuse in a system

$$\frac{\partial T}{\partial t} = \nabla^2 T \ . \tag{1.135}$$

The heat equation most certainly isn't time reversal invariant since the left-hand side picks up a minus sign, while the right-hand side does not. That's to be expected: after all, diffusion is a process that increases entropy and there's a clear arrow of time as things spread out. In contrast, there's no increase in entropy for a single quantum particle and we do expect the physics to be invariant under time reversal. Yet the Schrödinger equation is almost identical to the heat equation in structure. How can one be time reversal invariant, and the other not?

Almost identical, but not quite. The key is that factor of $i$ in the Schrödinger equation that is not there in the heat equation. Suppose that $\Psi(t)$ is a solution to the Schrödinger equation. Then $\Psi(-t)$ is *not* a solution but the factor of $i$ means that $\Psi^\star(-t)$ is. That's the clue that we need: time reversal in quantum mechanics acts as

$$T : \Psi(t) \mapsto \Psi^\star(-t) \ . \tag{1.136}$$

Viewed as an operator acting on the Hilbert space, this complex conjugation translates into the requirement that $T$ is an *anti-unitary* operator, rather than the more familiar unitary operator. This means that, acting on states, we have

$$T(\alpha|\psi_1\rangle + \beta|\psi_2\rangle) = \alpha^\star T|\psi_1\rangle + \beta^\star T|\psi_2\rangle \ . \tag{1.137}$$

In addition, the operator obeys

$$\langle T\psi_1|T\psi_2\rangle = \langle\psi_1|\psi_2\rangle^\star \ . \tag{1.138}$$

See the lectures on Topics in Quantum Mechanics for more discussion of the action of the time reversal in quantum mechanics.

This anti-linear behaviour changes some of the transformation properties of fields. For example, you might naively think, following (1.111), that $A_0$ would be odd under time reversal and $A_i$ even. But, in fact, it's the opposite way around because there's an additional factor of $i$ in the covariant derivative $\mathcal{D}_\mu = \partial_\mu - ieA_\mu$ and that gets conjugated. It means that the action of time reversal on the gauge field is

$$T : A_0(t, \mathbf{x}) \mapsto +A_0(-t, \mathbf{x}) \quad \text{and} \quad T : A_i(t, \mathbf{x}) \mapsto -A_i(-t, \mathbf{x}) \ . \tag{1.139}$$

Tracing this through to the electric field $\mathbf{E} = -\nabla A_0 - \partial\mathbf{A}/\partial t$ and magnetic field $\mathbf{B} = \nabla \times \mathbf{B}$, we have

$$T : \mathbf{E}(t, \mathbf{x}) \mapsto +\mathbf{E}(-t, \mathbf{x}) \quad \text{and} \quad T : \mathbf{B}(t, \mathbf{x}) \mapsto -\mathbf{B}(-t, \mathbf{x}) \ . \tag{1.140}$$

This makes sense: it's the same transformation that we get from the Lorentz force law $m\ddot{\mathbf{x}} = q(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B})$.

What about fermions? Once again, the action of time reversal can mix the different components of a Dirac spinor. As we now show, it turns out that (for our chiral basis of gamma matrices (1.42)) the correct transformation is

$$T : \psi(t, \mathbf{x}) \mapsto \Theta\psi(-t, \mathbf{x}) \quad \text{where} \quad \Theta = \gamma^1\gamma^3. \tag{1.141}$$

As for other transformations, we could also include a minus sign on the right-hand side. To see (1.141) is indeed a symmetry, consider the action of time reversal on the Dirac equation (1.125). Remembering that time reversal also acts by complex conjugation (so, for example, changes $\gamma^\mu$ to $(\gamma^\mu)^\star$), we have

$$\begin{aligned} & -i\big( - (\gamma^0)^\star\mathcal{D}_0 + (\gamma^i)^\star\mathcal{D}_i\big)\Theta\psi - M\Theta\psi = 0 \\ \Longrightarrow \quad & i\Theta^{-1}\big((\gamma^0)^\star\mathcal{D}_0 - (\gamma^i)^\star\mathcal{D}_i\big)\Theta\psi - M\psi = 0. \end{aligned} \tag{1.142}$$

This gives us back the original Dirac equation if the matrix $\Theta$ obeys

$$\Theta^{-1}(\gamma^0)^\star\Theta = \gamma^0 \quad \text{and} \quad \Theta^{-1}(\gamma^i)^\star\Theta = -\gamma^i \ . \tag{1.143}$$

It's simple to check that, for the chiral basis of gamma matrices (1.42), $\Theta = \gamma^1\gamma^3$ does the job. We can also translate this to the action on the component Weyl spinors $\psi = (\psi_L, \psi_R)^T$,

$$T : \psi_L(t, -\mathbf{x}) \mapsto -i\sigma^2\psi_L(-t, \mathbf{x}) \quad \text{and} \quad T : \psi_R(t, \mathbf{x}) \mapsto -i\sigma^2\psi_R(-t, \mathbf{x}) \ . \tag{1.144}$$

We see that time reversal, like CP, does not mix the left- and right-handed Weyl spinors.

What would it mean for a quantum field theory to break time-reversal invariance? It sounds rather cool. In practice, however, a breaking of time reversal manifests itself in rather mundane ways. One simple example is the presence of an electric dipole moment for particles. Recall from the lectures on Electromagnetism that an electric dipole moment arises from two, equal and opposite, closely separated charges and gives rise to an electric field that drops off as $1/r^3$.

The dipole moment points in a particular direction. For an elementary particle, this direction must align with the spin otherwise the particle would pick a preferred direction in space and so break Lorentz invariance. But the spin and dipole moment transform differently under both parity and time-reversal. To see this, recall that spin $\mathbf{S}$ is a form of angular momentum $\mathbf{L} = m\mathbf{x} \times \dot{\mathbf{x}}$, which is even under parity and odd under time reversal. Hence, we have

$$\begin{aligned} P : \mathbf{S} \mapsto \mathbf{S} \quad & \text{and} \quad T : \mathbf{S} \mapsto -\mathbf{S} \\ P : \mathbf{E} \mapsto -\mathbf{E} \quad & \text{and} \quad T : \mathbf{E} \mapsto \mathbf{E} \ . \end{aligned} \tag{1.145}$$

This means that discovery of a dipole moment for a fundamental particle would imply that the laws of physics break both parity and time reversal invariance. The search for the electric dipole moment of the neutron remains one of the most direct ways to test for time-reversal breaking in the strong nuclear force. So far, no such breaking has been found. (We discuss this further in Section 3.4.) As we will see later, the weak force does break both parity $P$ and, to a lesser extent, time reversal $T$. This results in a theoretical prediction for the electric dipole moment of the electron, albeit one that is far below current experimental bounds.

### 1.4.4 CPT

There are theories that are invariant under our three discrete symmetries, $C$, $P$ and $T$, and other theories that break them. As we will see, the Standard Model is in the latter class and all three symmetries are broken.

However, there is a theorem that says that all relativistic quantum field theories must necessarily be invariant under the combined action of $CPT$. In other words, if you look at anti-particles in the mirror, with their motion reversed, then you will have a symmetry on your hands.

One somewhat workaday proof of the CPT theorem is to simply write down all possible Lorentz invariant terms and check that they are indeed invariant under CPT. As we've seen, the most subtle transformations are those of spinors. For example, combining our previous results (1.117), (1.126) and (1.141), we find that a Dirac spinor is transformed by the anti-unitary operation

$$CPT : \psi(x) \mapsto -\gamma^5 \psi^\star(-x) \quad \text{with} \quad \gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.146}$$

You can check that all fermion bilinears are invariant under this transformation. For example,

$$\bar{\psi}\psi = \psi^\dagger \gamma^0 \psi \mapsto \psi^T \gamma^5 \gamma^0 \gamma^5 \psi^\star = -\psi^T \gamma^0 \psi^\star = \bar{\psi}\psi \tag{1.147}$$

where, in the final equality, we reordered the fermions and picked up a minus sign for our troubles due to their Grassmann nature. The pseudoscalar $\bar{\psi}\gamma^5\psi$ is also invariant by a similar argument, while both $\bar{\psi}\gamma^\mu\psi$ and $\bar{\psi}\gamma^\mu\gamma^5\psi$ transform as vectors, rather than pseudovectors (meaning that they pick up minus signs) which ensures that any kinetic term we write down is invariant. (For this, you will need to use the fact that $\gamma_1^T = -\gamma_1$ and $\gamma_3^T = -\gamma_3$ while $\gamma_0^T = \gamma_0$ and $\gamma_2^T = \gamma_2$.)

A slightly more elegant, but not entirely convincing, demonstration of CPT follows from Wick rotating to Euclidean space. Here we sketch the basic idea. The full Lorentz group in Minkowski space is really $O(1,3)$ and contains four disconnected components, with the actions of parity and time reversal taking us from one component to the other. In contrast, in Euclidean space the group becomes $O(4)$ and this contains only two disconnected components. If you follow the Lorentzian $CPT$ under a Wick rotation, it becomes simply a rotation in $SO(4)$, i.e. a transformation that is connected to the identity. (The need to include $C$ here is roughly because particles are like anti-particles travelling backwards in time.) This means that if your Euclidean theory is to have $SO(4)$ rotational invariance, then your Lorentzian theory must enjoy $CPT$.

The statement that $CPT$ is a symmetry of all relativistic quantum field theories is eminently falsifiable. Here's an example from neutrino physics. We will learn later that neutrinos oscillate from one flavour to another as they travel through space. So, for example, a muon neutrino $\nu^\mu$ will have some probability to convert into an electron neutrino $\nu^e$, a process that we write as

$$\nu^\mu \longrightarrow \nu^e \ . \tag{1.148}$$

We could also consider the CP conjugate process, namely

$$\bar{\nu}^\mu \longrightarrow \bar{\nu}^e \ . \tag{1.149}$$

There is no reason for the amplitudes for these two processes to be equal if CP is broken. However, there is also the time reversed process of (1.148)

$$\nu^e \longrightarrow \nu^\mu \ . \tag{1.150}$$

This too may have a different amplitude to (1.148) if time reversal is broken. However, CPT tells us that the amplitude for (1.149) and the amplitude for (1.150) are necessarily equal. Indeed, all experimental tests so far have failed to find any violation of CPT.