"WHAT IS AN ELECTRON?"

OR

PARTICLES AS REPRESENTATIONS

ROK GREGORIC

Abstract. In these informal notes, we attempt to offer some justification for the mathematical physicist's answer: "A certain kind of representation.". This requires going through some reasonably basic physical ideas, such as the barest basics of quantum mechanics and special relativity.

Contents

0.1. **What these notes are.** This are higly informal notes on basic physics. Being a mathematician-in-training myself, I am afraid I am incapable of writing for any other audience. Thus these notes must fall into the unfortunate genre of "physics for mathematicians", an idiosyncracy on the level of "music for the deaf" or "visual art for the blind" - surely possible in some fashion, but only after overcoming substantial technical obstacles, and even then likely to be accussed by the cognoscenti of "missing the point" of the original. And yet we beat on, boats against the current, ...

0.2. **How they came about.** As usual, this write-up began its life in the form of a series of emails to my good friend Tom Gannon. Allow me to recount, only a tinge apocryphally, one major impetus for its existence: the 2020 New Year's party.

Hosted at Tom's apartment, it featured a decent contingent of UT math department's grad student compartment. Among others in attendence were Ivan Tulli, a senior UT grad student working on mathematics inperceptably close to theoretical physics, and his wife Anabel, a genuine experimental physicist from another department, in town to visit her husband.

At some point in the evening, after a couple of drinks have sufficiently clouded his mind so as to be oblivious to what this course of action would inevitably bring down upon him, Tom turned to the Tullis and posed them a question, loud enough for all the room to hear:

"What is an electron?"

As should perhaps be expected, he got two different answers. Ivan waxed poetically about unitary representations of the Poincare group, while Anabel succintly listed a few

Date: January 20, 2020.

University of Texas at Austin.

measurable quantities: mass, spin, charge, and said that an electron is a particle with those quantities at some specified values.

While Tom admitted to prefer Anabel's question, he was also aware that it was Ivan's answer which was much closer to the relationship between physics and Tom's very own field of research: geometric representation theory.

Of course, both answes are really the same, only disguised in slightly different clothing. I spent a fair amount of time the past year trying to learn a bit about that, and so of course, I set out to write Tom an email to explain it to him too. And a few weeks later, here we are.

0.3. **What they do not contain.** Rigor, for one. Correctness may be another..

They are also in no way complete. We are telling a very specific story, and omitting much. For instance, while the question that sparked it was about electrons, we do not delve even the smallest bit into electromagnetism. As such, we can at best hope to describe how the representation-theoretic perspective on particles recovers their mass and spin. Charge, which governs the suceptibility to electromagnetic effects, is outside our reach. Though the story of EM is at least as interesting as the one we will describe, we must make peace with omitting it - these notes are already too long as they are!

As always, I can not guarantee that everything below is correct, only that I thought itwas when I wrote it. Use at your own risk!

1. Quantum Mechanics in a Nutshell

1.1. **Why though?** Presumably you're vaguely aware of classical physics, which agrees quite well with experiments in the size range of human perception (so: from the sizes say barely perceptible to the human eye, up to the sizes of say planetary motion in the solar system).

But when we try to talk about the really small, the predictions that classical physics makes just start failing to agree with experimental observations. So a new model is necessary. Enter quantum mechanics (QM for short)!

1.2. **The Hilbert space.** The basic tenant of QM is that states of a system come in the form of a complex Hilbert space $\mathcal{H}$. Unless the system is ridiculously simple (e.g.: quantum bits or qubits, relevant for quantum computing, and the like), this Hilbert space will be infinitely dimensional.

What precisely $\mathcal{H}$ is depends a bit on the question, and also possibly on how we are mathematically describing the system (e.g. different choices of coordinates or parametrizations, etc.). A very common case in real-life examples, e.g. to study very small particles and the like, is to take $\mathcal{H} = L^2(\mathbf{R}^3)$, with its usual inner product

$$\langle \psi_2 | \psi_1 \rangle = \int_{\mathbf{R}^3} d^3x \, \overline{\psi_2(x)} \psi_1(x).$$

The intuition is that an element $\psi \in L^2(\mathbf{R}^3)$ is a wave-function of a particle, i.e. a distribution spread out through the entire physical space $\mathbf{R}^3$, whose value at a point is the probability that the particle is located there.

In general, if the classical space in which the things you're thinking about are moving is a manifold M, taking $\mathcal{H} = L^2(M)$ is a good bet. You might want to rather do $\mathcal{H} = L^2(M; E)$, the L^2-sections of some complex vector bundle $E \to M$, which allows for a little more freedom - the values of E_x for $x \in M$ may parametrize internal degrees of freedom (spin, polarization, etc.) of the particle, other than just its position (which is already encoded in the base-point $x \in M$).

Really the states of the system are represented by the unit vectors in $\mathcal{H}$ (even more precisely: elements of the projectivization $\mathbf{P}(\mathcal{H}) := (\mathcal{H} \setminus \{0\})/\mathbf{C}^\times$, but let's not be that precise), alas the whole Hilbert space structure is relevant for the reasons explained in the next section.

First though, the requirement of unitality on $\psi \in \mathcal{H}$ becomes, in the above-mentioned case $\mathcal{H} = L^2(\mathbf{R}^3)$, the requirement that

$$\int_{\mathbf{R}^3} d^3x \, |\psi(x)|^2 = 1.$$

This is certainly a sensible requirement if ψ is to be interpreted as a probability distribution for the location of a particle; surely all the probabilities should sum (= integrate) up to 1, since a particle should be somewhere in $\mathbf{R}^3$. That's just how probabilities are: they sum up to one! :)

1.3. **Probability amplitudes.** As is perhaps not surprising in light of the discussion of the special example $\mathcal{H} = L^2(\mathbf{R}^3)$ above, the key thing we do with any Hilbert space of spaces $\mathcal{H}$ in QM is compute probability amplitudes.

If we know the system is in a state $\psi_1 \in \mathcal{H}$, and $\psi_2 \in \mathcal{H}$ is another state (and as said before, both should be unital), and we have a way to measure if the system is in state ψ_2, we may ask what is the probability that the measurement will be affirmative. This *probability amplitude* is computed as

$$|\langle \psi_2 | \psi_1 \rangle|^2$$

where $\langle \psi_2 | \psi_1 \rangle$ denotes the inner product of the two vectors.

Technically, we may view $|\psi_1\rangle$ as an element of the Hilbert space $\mathcal{H}$, and $\langle \psi_2|$ as an element of the dual Hilbert space $\mathcal{H}^*$, with the inner product providing a way to associate a linear functional to a vector. That's the notion a lot of physicists prefer, but I'll use it only sometimes.

Anyway, it might seem counter-intuitive that a system that we know to be in some state might also be measured to be in a different state. That's ostensibly due to the fact that the "states", in the sense of (unital) elements $\mathcal{H}$ are not necessarily *exclusive* states. E.g. ψ_1 could be something like "the particle is moving to the right" and ψ_2 could be "particle is spinning around an axis" - they don't need to exclude each other.

1.4. **Superposition.** But if we chose an orthonormal basis $\{\psi_i\}_{1 \leq i \leq \infty}$ for $\mathcal{H}$, then we would have $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ and so these states would be mutually exclusive. Since any state $\psi \in \mathcal{H}$ may be written as

$$\psi_i = \sum_{i=1}^{\infty} a_i \psi_i, \qquad a_i \in \mathbf{C}$$

(note that this is a *bona fide* convergent infinite sum!), it follows that the states ψ_i also suffice to entirely describe the state of the system. That is, specifying all the amplitudes $\langle \psi_i | \psi \rangle = a_i$ uniquely specifies ψ itself! But note that there are some restrictions on what these coefficients can be, e.g. if we want ψ to be unital, we must have $\sum_{i=1}^{\infty} |a_i|^2 = 1$.

In the situation above, say if we had (to simplify) $\psi = \frac{1}{2}\psi_1 + \frac{\sqrt{3}}{2}\psi_2$ in terms of an ONB $\{\psi_i\}_{1 \leq i \leq \infty}$, then we would say that the system is in a *superposition* of states ψ_1 and ψ_2. That's all that superposition is: that not every element in a vector space is just a basis vector!

1.5. **Phase.** Write $\langle \psi_2 | \psi_1 \rangle = re^{i\theta}$ for two arbitrary unit elements $\psi_1, \psi_2 \in \mathcal{H}$ (so unlike in the previous paragraph, they are *not* necessarily elements of a basis). You will have noted that in 1.3 we only gave physical meaning to the radius r (really its square) as a probability. The factor $e^{i\theta}$, called the *phase*, has a less clear physical interpretation, though it is quite essential.

You see, it is necessary to explain the (experimentally observed!) phenomenon of *interference*.

You have probably heard of this in the context of the double slit experiment, but it shows up in essentially any experiment on a small enough scale. The point is that different states

can interfere with each other, much alike waves, in ways that one wouldn't expect from one's classical intuition (where all quantities are real numbers!).

Mathematically, the point is that if you have two complex numbers $a, b \in \mathbf{C}$, then the absolute value of $a + b$ won't just depend on $|a|$ and $|b|$, but also on the phase factors. That's what is broadly behind the phenomenon of interference, but I don't want to go into more detail here.

1.6. **Observables.** An observable, or an observable quantity, is some property of a system that we can (or at least, conceivably could) measure. That might be position, velocity, angular momentum, energy, spin, polarization, etc.

We would like to compute the *expectation value* $\langle \mathcal{O} \rangle_\psi$ of an observable $\mathcal{O}$, provided the system is in state ψ. Experimentally we obtain such an expectation value $\langle \mathcal{O} \rangle_\psi$ by performing the measurement experiment which determines the value of $\mathcal{O}$ when we prepare the state of the system to be ψ, repeating the experiment a large number of times, and then averaging the results. Hence $\langle \mathcal{O} \rangle_\psi$ is the *expectation value*, the value we can probabilistically expect to get, and not the "actual definitive value that we will always get when doing the same experiment" (such as is always the case in classical physics). So far as we can tell, the latter simply does not exist on very small scales. That is the major difference between classical and quantum - in the quantum realm, the best we can do is predict probabilities! :)

But how do we *compute* such an expectation value in the Hilbert space framework of QM? The answer is that observables are modeled as certain sort of (we'll get back to this in a second!) linear operators on $\mathcal{H}$, and the the expectation value for $\mathcal{O}$ when in the state ψ is defined as $\langle \mathcal{O} \rangle_\psi := \langle \psi | \mathcal{O} | \psi \rangle$. That is to say, take the state ψ, apply the operator $\mathcal{O}$, and then pair (via the Hilbert space's inner product) against ψ again.

You could similarly interpret $\langle \psi_2 | \mathcal{O} | \psi_1 \rangle$ as the expectation value for the observable $\mathcal{O}$, if we prepared at the start the system to be in the state ψ_1, and we found at the end the system to be in state ψ_2.

Of course we expect measurements of real-world quantities to be real numbers! Since $\overline{\langle \psi | \mathcal{O} | \psi \rangle} = \langle \psi | \mathcal{O}^\dagger | \psi \rangle$, where $\mathcal{O}^\dagger$ is the adjoint operator to $\mathcal{O}$, this reality requirement translates to the requirement that the linear operator $\mathcal{O}$ on $\mathcal{H}$ be *self-adjoint*, or in physics-speak: *Hermitian*. This is the extra condition required on $\mathcal{O}$ alluded to two paragraphs ago. Thus quantum observables are precisely Hermitian operators on the Hilbert space.

The Hermitian condition $\mathcal{O}^\dagger = \mathcal{O}$ is really a functional-analytic analogue of reality. The role of the adjoint operation $(-)^\dagger$ is analogous to complex conjugation, and so an operator being Hermitian corresponds to a complex number $z \in \mathbf{C}$ satisfiying $\overline{z} = z$. Of course, that just means that $z \in \mathbf{R}$.

1.6. $+\frac{1}{2}$ **Digression: measurement in quantum land.** This is really just a remark, with no effect on what follows. But since we encountered in the previous section the first piece of quantum weirdness, that being that things don't seem to be deterministic and we can at best determine probabilities and expectation values, I thought maybe this would also be a good spot to mention the other piece of weirdness.

It is closely related, and has to do with the actual values of measurements on small scales. Say you have fixed an observable $\mathcal{O}$, and you set up an experiment to determine the value of $\mathcal{O}$. As discussed, you've given up the idea of always getting the same result from the same initial conditions, but averaging a lot of results will reproduce the expectation value. So far so good. But then you notice a weird thing happening: the actual individual results of the measurements you do each time, while not always being equal, are however not entirely random either. Instead, there are only certain values which ever come up as results of this experiment, forming a discrete subset of $\mathbf{R}$.

This is familiar to even the layman: electric charge, say is "quantized", in that it always comes in integer multiples of the basic electric charge e. But it is not an exception, but

the rule: the same happens with all quantities we could imagine to measure! Now just what is going on?

It turns out that, if the observable $\mathcal{O}$ being measured is given by a Hermitian operator on the Hilbert space, then the only possible values that measuring it can return are its *spectrum*, i.e. the eigenvalues of the operator $\mathcal{O}$. This is quite shocking, but it is how nature works. Since we are in QM only in the business of predicting expectation values, and not results to experiments themselves, this has little direct effect.

On the other hand, this phenomenon of quantities becoming "quantized" on small scales, i.e. coming in discrete packages (or "quanta") instead of being continuously distributed (as classical physics would predict) was one of the first experimental observations that couldn't be explained classically. QM of today, with all its weird Hilbert spaces and stuff, is what we as a species came up with to explain this (and many other) experimentally detected phenomena.

But enough philosophy, let's get back to observables!

1.7. **Simple examples of observables.** Let's do two of the simplest examples. Let us consider some particle moving around in 3-space, thus $\mathcal{H} = L^2(\mathbf{R}^3)$. As discussed in 1.2, the (unital) element of the Hilbert space are viewed as wave-functions, encoding the probability distributions for the location of the particle. In the ambient position space $\mathbf{R}^3$, let us choose an origin and lay down coordinates x^1, x^2, x^3.

(1) Then we can ask where in space the particle is located, and that reduces to three measurements: its x^1-coordinate position, its x^2-coordinate position, and its x^3-coordinate position. Lo and behold: three observables! The way this works is: the observable "x^j-coordinate position", which we mostly abusively just call x^j, but let's be nice and call it $\hat{x}^j$, is the operator on $\mathcal{H}$ which sends an L^2-function $\psi : \mathbf{R}^3 \to \mathbf{C}$ to the function $x^j\psi : \mathbf{R}^3 \to \mathbf{C}$. That is to say, $\hat{x}^j$ is multiplication by the function $x^j : \mathbf{R}^3 \to \mathbf{R} \subseteq \mathbf{C}$, reading off the x^j-coordinate.

(2) We can also ask (and hopefully measure!) what the velocity, or better, momentum of the particle is. Momentum (think velocity, if it makes you feel safer) is a vector quantity, i.e. has direction as well as magnitude, so we should get three new observables: the x^1-coordinate of momentum, the x^2-coordinate of momentum, and the x^3-coordinate of momentum. These should be Hermitian operators $\hat{p}_1$, $\hat{p}_2$, $\hat{p}_3$ on $\mathcal{H}$. It turns out that those are given as

$$\hat{p}_j = \tfrac{\hbar}{i}\partial_j$$

where $\partial_j = \frac{\partial}{\partial x^j}$ is the partial derivative, and $\hbar$ is the Planck constant, a specific number that has to be there to agree with experimentally determined values.

For the time being I encourage you to take this *ad hoc* definition of momentum on faith, but there is a deeper (and quite insightful) reason for it, that we will say a little bit about later when we discuss symmetries.

Also, a reader familiar with such things will surely be happy to recognize the algebra of observables spanned by position and momentum as the differential operators. Indeed, this is the major way in which differential operators enter the math of quantum stuff.

1.8. **Energy.** Another (key) thing we should be able to measure is the energy of a system. This should thus also be an come in the form of an observable, called the *Hamiltonian*: a Hermitian operator $\hat{H}$ on the Hilbert space of states $\mathcal{H}$. (Between $\mathcal{H}$, $\hbar$, and $\hat{H}$, the letter H suffers from quite a bit of overuse in quantum stuff.)

Since energy is arguably the main thing we measure from a system, we make it play a distinguished role. For instance, we often use the ONB for $\mathcal{H}$ consisting of the unital eigenvectors for $\hat{H}$ (such an eigenbasis exists for any Hermitian operator by the so-called Spectral Theorem). Let's denote those by $|E\rangle$, where this eigenvector has eigenvalue E

(and $\in \mathbf{R}$ too, due to Hermitianness of $\hat{H}$). That is to say, $\hat{H}|E\rangle = E|E\rangle$. Then we can develop any state $\psi \in \mathcal{H}$ in terms of the energy eigenbasis as

$$\psi = \sum_E a_E |E\rangle,$$

which is useful for instance because the (expected) energy of the system in that state is then

$$\langle \hat{H} \rangle_\psi = \langle \psi | \hat{H} | \psi \rangle = \sum_E |a_E|^2.$$

Key takeaway: there is a distinguished observable $\hat{H}$, the Hamiltonian, which encodes energy.

Example. In the setting of the example from 1.7 where $\mathcal{H} = L^2(\mathbf{R}^3)$, the standard choice for a Hamiltonian is $\hat{H} = \frac{1}{2m}|\hat{\mathbf{p}}|^2$ (where the bold font indicates that $\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)$ is the momentum vector). This is inspired by the standard kinematic Hamiltonian $H = \frac{1}{2m}|\mathbf{p}|^2$ from classical mechanics, which is ostensibly just a restatement Newton's Law $\mathbf{F} = \partial_t \mathbf{p}$ in terms of energy instead of force. The dynamics imposed (as explained in the next section) by this Hamiltonian models the free motion of a point particle of mass m in 3-space, in the settings of both classical and quantum mechanics. In terms of the formula $\hat{p}_j = i\hbar\partial_j$ for quantized momentum from 1.7, the standard kinematic quantum Hamiltonian is

$$\hat{H} = \tfrac{1}{2m}\left(\hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2\right) = -\tfrac{\hbar^2}{2m}\Delta.$$

a constant multiple of the Laplacian.

1.9. **Schrodinger's equation.** So far we've only discussed states and observables, but not how their dynamical behavior. That is, we do not yet know how a QM system evolves with time.

The axiom governing time evolution in quantum mechanics is another thing you have heard of before: *Schrodinger's equation.* It says that for every $\psi \in \mathcal{H}$ we have

$$i\hbar\, \partial_t \psi = \hat{H}\psi,$$

thus identifying derivative with respect to time (the rate of change with respect to time) with the Hamiltonian operator, at least modulo the constants $i\hbar$ that have to be multiplied along to make things right from the agreeing-with-experiments POV.

Just as with the formula for momentum in 1.7, we will re-interpret Schrodinger's equation in the following section on symmetries, to hopefully make it less *ad hoc.*

1.10. **Time evolution.** But for now, let's take Schrodinger's equation at face value. It identifies the time derivative as $\partial_t = -\frac{i}{\hbar}\hat{H}$. Since time derivative should be the infinitesimal generator of time evolution, this suggests defining *time evolution* as the operator $\hat{U}(t) = e^{-\frac{i}{\hbar}t\hat{H}}$. This is the operator exponential, which is to say if $\psi \in \mathcal{H}$ is written in the energy eigenbasis from 1.8 as $\psi = \sum_E a_E|E\rangle$, then time evolution is given by

$$\hat{U}(t)\psi = \sum_E e^{-\frac{i}{\hbar}tE} a_E |E\rangle.$$

Thus if ψ is the starting state of our system in question, whose dynamics is governed by the Hamiltonian $\hat{H}$, then after time t the system will be in state $\hat{U}(t)\psi$.

As per usual, the actual thing of relevance to experiments, i.e. that we could measure, would be the transitional amplitude $\langle \psi_2 | \hat{U}(t) | \psi_1 \rangle$, encoding the probability that, if we prepare the system in state ψ_1, then after time t elapses, we will find system to be in state ψ_2.

2. Symmetry

At last, we come to symmetries of a quantum system, the topic we have been very much building towards this whole time, and which shall be constantly in the background from here on out!

2.1. **Unitary operators.** So what should a symmetry of a quantum system be? Well, since the system is given by a Hilbert space $\mathcal{H}$, and its inner product has physical relevance as probability amplitudes, then surely any symmetry of the system should preserve it.

Thus we wish to consider maps $U : \mathcal{H} \to \mathcal{H}$ which preserves the Hilbert space structure, in the sense that they are linear, and that they satisfy

$$\langle U\psi_2 | U\psi_1 \rangle = \langle \psi_2 | \psi_1 \rangle$$

for all $\psi_1, \psi_2 \in \mathcal{H}$. Since the definition of the adjoint operator shows that

$$\langle U\psi_2 | U\psi_1 \rangle = \langle \psi_2 | U^\dagger U | \psi_1 \rangle,$$

this requirement is equivalent to demanding that $U^\dagger U = 1$ (where we just write 1 for the identity map $\mathrm{id}_{\mathcal{H}}$). We say that U is a *unitary operator*.

Note that the unitarity condition implies that the adjoint $U^\dagger$ is a compositional inverse to U, hence making every unitary operator invertible. Hence unitary operators on the Hilbert space $\mathcal{H}$ form a group, which we will denote $\mathrm{U}(\mathcal{H})$.

2.2. **Groups of symmetries.** Of course we usually don't care for just one single symmetry, but for a whole family of them. More precisely, a group. Thus a G-symmetry will consist of a family of unitary operators $\rho(g) : \mathcal{H} \to \mathcal{H}$ indexed on the points $g \in G$, and satisfying

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2), \qquad \rho(1) = 1.$$

That is to say, a group homomorphism $\rho : G \to \mathrm{U}(\mathcal{H})$, or yet equivalently, a *unitary representation* of G.

2.3. **Compatibility with the Hamiltonian.** A quantum system is not fully described by the Hilbert space $\mathcal{H}$ though. We found in 1.9 that the dynamics are encoded by the Hamiltonian $\hat{H}$. The latter is an observable, which is to say, a Hermitian operator on $\mathcal{H}$.

In order for a unitary operator $U : \mathcal{H} \to \mathcal{H}$ (or a family thereof, i.e. a unitary representation of a Lie group on $\mathcal{H}$) to be a *bona fide* symmetry of the system, it must be compatible with the Hamiltonian. It turns out that the appropriate requirement is the commutator vanishing

$$[\hat{H}, U] = 0.$$

To say the same thing from another (perhaps more insightful) perspective: we should require that a real symmetry of the system does not depend on the specific time at which we look at it, but is instead a symmetry throughout. This means requiring that U be invariant under time evolution. Since we saw in 1.10 that time evolution acts on states through the operators $\hat{U}(t) = e^{-\frac{i}{\hbar}t\hat{H}}$, this amounts to asking that conjugation with $\hat{U}(t)$ preserves U. But then that is the condition that

$$U = \hat{U}(t)^\dagger U \hat{U}(t) = e^{\frac{i}{\hbar}t\hat{H}} U e^{-\frac{i}{\hbar}t\hat{H}},$$

which is equivalent to requiring that $U = \hat{H}^{-1} U \hat{H}$, which is precisely the commutator requirement above.

2.4. **Symmetries of a quantum system.** To surmise the discussion so far: without considering the dynamics, the maximal group of symmetries is $\mathrm{U}(\mathcal{H})$. But when taking into account the dynamics garnered by the choice of a Hamiltonian $\hat{H}$, the maximal group of symmetries becomes the stabilizer subgroup $\mathrm{Stab}_{\mathrm{U}(\mathcal{H})}(\hat{H})$ for the conjugation action of $\mathrm{U}(\mathcal{H})$ on quantum observables $\mathrm{Obs}(\mathcal{H})$, which is to say, Hermitian operators on the Hilbert space $\mathcal{H}$.

Okay, we saw representations enter the picture, though as symmetries. There is still a little ways to go before we understand how this makes sense of the matematicial physicist's (or more accurately: physical mathematician's) motto of "particles are representations". We'll get there, but first, let us pay off some debts from 1.7 and 1.9.

2.4. **Noether's Theorem.** There is a general principle in physics which is:

Symmetries correspond to conserved quantities.

A formalization of this theorem in the realm of classical mechanics goes by the name Noether's Theorem, as it was first singled out and proved by Emmy Noether. While I don't know for a fact, I imagine that this had to have been one of the stronger motivations for her to later play such an important role in the development of group theory, and algebra in general.

In QM this becomes particularly simple, so let's see how it works through the next few sections.

2.5. **Noether's Theorem in QM: the setup.** Consider a G-symmetry of a quantum system, given by a Hilbert space $\mathcal{H}$ and a Hamiltonian $\hat{H}$. The symmetry is encoded by a representation, i.e. a homomorphism $\rho : G \to \mathrm{U}(\mathcal{H})$ of Lie groups (infinite dimensional as they may be). Now note that a Lie group homomorphism induces a corresponding map of Lie algebras $d\rho_1 : \mathfrak{g} \to \mathfrak{u}(\mathcal{H})$. This Lie algebra map recovers ρ upon passing to Lie group exponential maps, in the sense that the obvious square

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{d\rho_1} & \mathfrak{u}(\mathcal{H}) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\rho} & \mathrm{U}(\mathcal{H}) \end{array}$$

commutes. Observe that the Lie algebra exponential $\exp : \mathfrak{u}(\mathrm{H}) \to \mathrm{U}(\mathcal{H})$ is just the operator exponential, as we have already encountered when discussing time evolution in 1.10.

2.6. **Infinitesimal generators of symmetry.** Physically we call $d\rho_1$ the *infinitesimal generator* of the action. Or more precisely, we view elements $\xi \in \mathfrak{g}$ themselves as being infinitesimal generators of the G-symmetry, acting on the Hilbert space $\mathcal{H}$ by the operator $d\rho_1(\xi)$. Now, the unitary Lie algebra $\mathfrak{u}(\mathcal{H})$ is not hard to determine; since $\mathrm{U}(\mathcal{H})$ is cut out insede $\mathrm{Aut}(\mathcal{H})$ by the equation $U^\dagger U = 1$, we may derivate to obtain the equation $X + X^\dagger = 0$ cutting out $\mathfrak{u}(\mathcal{H}) = T_1\mathrm{U}(\mathcal{H})$ inside $\mathfrak{aut}(\mathcal{H}) = \mathrm{End}(\mathcal{H})$. In conclusion: $\mathfrak{u}(\mathcal{H})$ consists of linear operators $X : \mathcal{H} \to \mathcal{H}$ which satisfy $X^\dagger = -X$, i.e. which are anti-Hermitian.

2.7. **Hermitian vs anti-Hermitian operators.** The appearance of anti-Hermitian operators tickles our little gray cells - it reminds us of the Hermitian operators which we already encountered in our journey in 1.6 in the guise of quantum observables. Of course in QM we always work over $\mathbf{C}$, and so we have a way to turn anti-Hermitian operators into Hermitian ones. Indeed, since taking adjoint works on scalars by conjugating them, we have for any $X \in \mathfrak{u}(\mathcal{H})$ that $(iX)^\dagger = -iX^\dagger = iX$. Hence multiplication by i, and in fact by $\hbar i$ for any non-zero real constant $\hbar$, provides an isomorphism

$$\mathfrak{u}(\mathcal{H}) \xrightarrow{i\hbar} \mathrm{Obs}(\mathcal{H}).$$

You might very justifiably ask why we chose to introduce this auxiliary constant $\hbar$. Mathematically there is no need for it, and indeed it essentially amounts to an unfortunate choice of convention at some point of physics history. But it is quite standard, so let's keep it here (as said, it effects nothing).

In fact, mathematicians and mathematical physicists often prefer to work in what's called "natural units", in which $\hbar = 1$ (along with other physical constants, such as the speed of light c), and it politely disappears from formulas. But let's not be that sophisticated!

2.8. **Noether's Theorem in QM: conclusion.** Combining the discussion of 2.5 - 2.7, we find that any symmetry $\rho : G \to \mathrm{U}(\mathcal{H})$ gives rise to a linear map

$$J : \mathfrak{g} \xrightarrow{d\rho_1} \mathfrak{u}(\mathcal{H}) \xrightarrow{i\hbar} \mathrm{Obs}(\mathcal{H}).$$

That is to say, for every infinitesimal generator of the symmetry $\xi \in \mathfrak{g}$, we have an associated observable

$$J(\xi) = i\hbar\, d\rho_1(\xi),$$

called its *Noether current*. These are the promised "conserved quantities" of Noether's theorem.

Indeed, so far we haven't used the compatibility of ρ with the Hamiltonian $\hat{H}$ in this discussion. Using it, in the form $[\hat{H}, \rho] = 0$, implies rather easily (it's actually a good exercise on how the association of Noether's currents to symmetries works, and I recommend trying it) that $[\hat{H}, J(\xi)] = 0$ holds also. That is to say, Noether's currents too also compatible with the Hamiltonian, and hence by the argument of 2.3, invariant under time evolution. Invariant under time evolution = not changing with time = conserved.

2.9. **Noether's Theorem: in reverse.** If the symmetry group G is simply connected, we can also go in reverse. Given a "Noether current" $J : \mathfrak{g} \to \mathrm{Obs}(\mathcal{H})$, compatible with the Hamiltonian, and such that $\frac{1}{i\hbar} J = -\frac{i}{\hbar} J$ is a Lie algebra homomorphism, then there exits a unique symmetry $\rho : G \to \mathrm{U}(\mathcal{H})$ whose Noether current is J. Explicitly, the representation ρ takes an element $g = \exp(\xi) \in G$ with $\xi \in \mathfrak{g}$ to the operator

$$\rho(g) = e^{-\frac{i}{\hbar} J(\xi)},$$

and knowing what it should do on the image of the Lie group exponential suffices to determine ρ.

Thus Noether's Theorem gives (at least for simply connected Lie groups) a full bijection between symmetries and conserved quantities. Thus instead of thinking about a symmetry, one can just about the corresponding conserved observables.

2.10. **Symmetry derivation of Schrodinger's equation.** Alright, we're finally at the point where we can pay off our debts from 1.7 and 1.9, and give insightful explanations of the formula for momentum, and of Schrodinger's equation. Let us do the latter first, devoting this section to it.

Suppose we didn't subscribe to the idea that time evolution in QM should be governed by a choice of an observable $\hat{H}$ and Schrodinger's equation, as we imposed in 1.9 and 1.10. Instead, let's think about what time evolution *should* be.

We have a Hilbert space of states $\mathcal{H}$, and time evolution $\hat{U}(t)$ should certainly take in states which the system is at the start, and return the what state the system will evolve into after t amount of time has passed. As explained in 2.1, the requirement that time evolution respect the physically-relevant Hilbert space structure, translates into the requirement that $\hat{U}(t)$ be a unitary operator. Furthermore evolving by time t and then by time s should be the same as evolving by time $t + s$ - we are looking for an $\mathbf{R}$-symmetry of $\mathcal{H}$.

That is to say, time evolution should come in the form of a Lie group homomorphism $\hat{U} : \mathbf{R} \to \mathrm{U}(\mathcal{H})$. By the discussion of 2.5 - 2.7 (where we didn't require picking a Hamiltonian), this gives rise (and is, due to simple connectivity of the additive group $\mathbf{R}$, equivalent to) a Noether current $J : \mathbf{R} \to \mathrm{Obs}(\mathcal{H})$, where we use that $\mathbf{R} = T_0\mathbf{R}$ is its own Lie algebra. Since J is a linear map, it is uniquely specified by the image $\hat{H} := J(1) \in \mathrm{Obs}(\mathcal{H})$, a distinguished observable.

As the notation indicates, we have just recovered the Hamiltonian. Indeed, using only the defining formula for Noether currents from 2.8, lo and behold

$$\hat{H} = J(1) = i\hbar\, d\hat{U}_0 = i\hbar \tfrac{d}{dt}\big|_{t=0}\hat{U}(t) = i\hbar\,\partial_t,$$

the Schrodinger equation in all its terrible glory!

So while the way we introduced the dynamics of a QM system in 1.9 and 1.10 through the Hamiltonian and Schrodinger's equation seemed ad hoc and random, we see that it is actually forced upon us by symmetry considerations, from the simple observation that time evolution must be an $\mathbf{R}$-symmetry.

2.11. **Symmetry derivation of the formula for momentum.** One debt down, one to go! Recall the setup of 1.7 of the quantum system describing the motion of a particle in $\mathbf{R}^3$. We have $\mathcal{H} = L^2(\mathbf{R}^3)$ with its usual inner product, viewing the unital elements as probability distributions for the position of the particle in 3-space.

One obvious symmetry this system possesses (disregarding for the moment the dynamical behavior, i.e. we have not yet chosen a Hamiltonian) is the unitary representation $\rho : \mathbf{R}^3 \to \mathrm{U}(\mathcal{H})$ given by positional translation. That is, for any point $x_0 \in \mathbf{R}^3$ the corresponding operator sends an L^2-function $\psi(x)$ to the function $\psi(x - x_0)$.

Let us determine the Noether currents for this action! Once again the additive Lie group $\mathbf{R}^3$ is its own Lie algebra. Since the Noether current map $J : \mathbf{R}^3 \to \mathrm{Obs}(\mathcal{H})$ is linear, it suffices to determine where a basis is sent. This corresponds to picking coordinates x^1, x^2, x^3 on $\mathbf{R}^3$ along the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. We wish to determine the specially-named observables $\hat{p}_j := J(\mathbf{e}_j)$, whose name suggests that they are (the respective components of) momentum. By definition of Noether currents we have

$$\hat{p}_j = i\hbar\, d\rho_0(\mathbf{e}_j),$$

and so we compute the derivative

$$(d\rho_0(\mathbf{e}_j))(\psi(x)) = \tfrac{d}{dt}\big|_{t=0}\rho(t\mathbf{e}_j)(\psi(x)) = \tfrac{d}{dt}|_{t=0}\psi(x - t\mathbf{e}_j) = -\partial_j\psi(x)$$

where $\partial_j = \frac{\partial}{\partial x^j}$ is the j-th partial derivative with respect to the chosen coordinates. Putting those together recovers the formula

$$\hat{p}_j = \tfrac{\hbar}{i}\partial_j$$

for quantized momentum, that we gave in section 1.7 unmotivated.

We have now seen that momentum results from translational $\mathbf{R}^3$-symmetry of the system. Thus momentum will only be a conserved quantity (a constant of motion) if the Hamiltonian $\hat{H}$ defining the dynamics is translationally invariant. Of course, this tends to be the case. For instance, it is certainly the case for the standard kinetic Hamiltonian from the example in 1.8.

2.12. **Angular momentum.** While I don't want to write it out any of the details here (since I fear that I have probably showcased too many of the them so far already!), it is also a reasonably simple game to figure out that rotational symmetry on $\mathcal{H} = L^2(\mathbf{R}^3)$ that is, $\mathrm{SO}(3)$-symmetry induced by rotations on $\mathbf{R}^3$ in the arguments of L^2-functions, give rise to the conserved quantity called *angular momentum*.

Since the Lie algebra $\mathfrak{so}(3)$ is isomorphic to $\mathbf{R}^3$ with the cross product (by sending the infinitesimal generator of the rotation around the coordinate axis $\mathbf{e}_j$ to the unit vector $\mathbf{e}_j$), it suffices to give formulas for Noether currents $\hat{L}_j := J(\mathbf{e}_j)$ for the standard basis of

3-space. Somewhat analogously for the formula for quantum momentum, these turn out to be

$$\begin{aligned}\hat{L}_1 &= \tfrac{\hbar}{i}(x^2\partial_3 - x^3\partial_2),\\ \hat{L}_2 &= \tfrac{\hbar}{i}(x^3\partial_1 - x^1\partial_3),\\ \hat{L}_3 &= \tfrac{\hbar}{i}(x^1\partial_2 - x^2\partial_1).\end{aligned}$$

If you so desired, you could express this in terms of curl. But slightly more naturally, collecting all three components together into $\hat{\mathbf{L}} = (\hat{L}_1, \hat{L}_2, \hat{L}_3)$, we may write $\hat{\mathbf{L}}\psi = \frac{\hbar}{i}\mathbf{x}\times\nabla\psi$. In terms of the expression for quantized momentum we derived in 2.12, this can be rewritten as $\hat{\mathbf{L}} = \mathbf{x}\times\hat{\mathbf{p}}$. That may finally harken your memory back to your introductory phisics class, where you probably used this formula (in the setting of non-quantum mechanics) to define angular momentum. What goes around, goes around, goes around, comes all the way back around.

3. Toward Particles

We are almost there, to explain why elementary particles (or better, their particle types) may be thought of as certain sort of irreducible representations. What the word "elementary" is supposed to indicate, is that it should be possible to obtain all other things in the Universe (at least on small scales - we're purposefully ignoring gravity, as it is much too weak to have any effect on the microscopic scale) by combining (possibly very large quantities) of those few elementary particles. But how *do* we combine quantum systems together?

3.1. **Composite systems.** Sometimes there are more than one particles involved in a system. This is quite a general situation in physics.

For instance, you could be looking at two particles moving around, and possibly even interacting in some way. Or think something like trying to model the solar system - although that wouldn't really be a quantum thing, scale-wise.

Quite generally, suppose we have two or more quantum systems interlocking, and we wish to treat them all together as a single quantum system. Treating each part individually and disregarding the rest, we may mathematically describe it by a Hilbert space $\mathcal{H}_i$ and a Hamiltonian $\hat{H}_i$. The index i here ranges over all the parts involved, say $i = 1, 2, \ldots, n$. To describe the total composite system, we should take the Hilbert space to be $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. Technically these tensor products, obviously over $\mathbf{C}$, must also be appropriately completed, but let us ignore that issue. We may view each Hamiltonian $\hat{H}_i$ as a Hermitian operator on the whole Hilbert space $\mathcal{H}$, by identifying it with the "diagonal" operator $1\otimes\cdots\otimes 1\otimes\hat{H}_i\otimes 1\otimes\cdots\otimes 1$ with $\hat{H}_i$ in the i-th spot. In this way, we may set the Hamiltonian of the whole system to be $\hat{H} = \hat{H}_1 + \cdots + \hat{H}_n + \hat{H}_I$, where the interaction Hamiltonian $\hat{H}_I$ must be supplied externally, and governs the way that the individual quantum systems may effect each other.

In what follows, we will only be interested in the Hilbert spaces of states, and entirely neglect the Hamiltonians. Of course the Hamiltonians are where the dynamics and as such real physics enters! But there is already surprisingly much about the Hilbert spaces.

3.2. **Indistinguishability.** Actually there are some more subtleties with composite systems than indicated in the preceding discussion. In the above description, while $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ includes all the n systems, we can still tell which one is which. That is, we know which came first, which second, etc. In effect, the individual parts are labeled, and in particular, distinguishable.

Alas, in experimental experience, this is simply *not the case* for very small particles such as electrons or photons. That is, when we are looking at say two electrons, we *can* determine that there are two (by preforming some sort of measurement - the number of

particles is an observable!), but *not* which of the two is which. We say that electrons (and other small particles) are *indistinguishable*.

This isn't as weird as it sounds - the way we usually "label" things is by changing something about them slightly, but in a way where this change will not effect anything in regards to what we are considering (coloring balls with different colors will not effect how they bounce off the ground, say). But for small almost-elementary particles, any changes will be pretty drastic, and will have a noticeable effect on how the particle will behave!

In any case, this implies that modeling n particles, whose individual states are described by a Hilbert space $\mathcal{H}_1$ (the subscript 1 is meant to indicate that this is the Hilbert space for one particle), by the (always implicitly completed) tensor power $\mathcal{H} = \mathcal{H}_1^{\otimes n}$, will not work.

3.3. **The Bose statistic.** If we can't tell n distinct particles apart, that means that there is a symmetric group Σ_n's worth of ambiguity. Thus we may try to replace the tensor power $\mathcal{H} = \mathcal{H}_1^{\otimes n}$ with the symmetric power

$$\mathcal{H} = \mathrm{Sym}^*(\mathcal{H}_1) = (\mathcal{H}_1^{\otimes n})/\Sigma_n$$

(as always, implicitly completed). If this works to describe the system of n particles of the same type, we say that said particle type *satisfies the Bose statistic*. Or more simply, we call such particles *bosons*.

This works brilliantly in certain cases, e.g. for photons, and various other particles. In the Standard Model: the gluons, the W and Z bosons, and Higgs boson (what a shock - particles with "boson" in their name are bosons!). These are precisely the "force-carrier" particles - interactions with them give rise to the fundamental forced. Photons mediate the electromagnetic force, gluons the strong nuclear force (they are called glueons because the "glue" the atomic nucleus together; physics has some top tier naming choices!), the W and Z bosons the weak nuclear force (responsible for radioactive phenomena).

3.4. **The Fermi statistic.** So all is well for force-carrying particles, but those are only half the story. The other half is matter, and it is different.

This difference is expressed by the *Pauli indeterminacy principle:*

> *Two electrons (or any other matter particles for that matter) may not simultaneously occupy the same state.*

We may express mathematically this in terms of the wave-function $\psi \in \mathcal{H}_1$ (think $L^2(\mathbf{R}^3)$, though it should really be L^2-sections of a bundle to account for the spin polarization of electrons) describing the probability density for the electron. Indeed, Pauli's principle states that, in the corrected version of the naive Hilbert space of two electrons $\mathcal{H} = \mathcal{H}_1^{\otimes 2}$, we must have $\psi \otimes \psi = 0$.

Being mathematicians, we instantly recognize that, since we are working over $\mathbf{C}$ and in particular in characteristic $\neq 2$, this squaring to zero requirement is equivalent to anti-symmetry. That is to say, in the correct Hilbert space for two electrons, we must have $\psi_1 \otimes \psi_2 = -\psi_2 \otimes \psi_1$. Contrast this to the symmetric square $\mathrm{Sym}^2(\mathcal{H}_1)$, in which $\psi_1 \otimes \psi_2 = \psi_2 \otimes \psi_1$. Thus the correct Hilbert space for n electrons may be obtained from the Hilbert space $\mathcal{H}_1$ for the single electron, as the exterior power

$$\mathcal{H} = \Lambda^*(\mathcal{H}_1) = (\mathcal{H}_1^{\otimes n})/\Sigma_n.$$

The difference from the case in the previous section is that, where in the previous section we took the quotient by the action of the symmetric group Σ_n by the obvious permutation action on the factors of $\mathcal{H}^{\otimes n}$, here we quotient by the alternating action.

The particles for which this works are called *fermions*, and are said to *satisfy the Fermi statistic*. As alluded to above, all the elementary matter particles in the Standard Model are of this sort: electrons, quarks, muons, and neutrinos (really electrons, muons, and

neutrinos are grouped together as leptons, but we don't intend to describe how SM works here anyway, so we can be a little vague).

BTW: two relevant remarks:

(1) This anti-commutativity of the Fermi statistic is the place where super-things (i.e. $\mathbf{Z}/2$-graded algebra, geometry, etc.) enter physics. This isn't supersymmetry though, which is a different (less experimentally verified) beast that I don't want to get into.
(2) The Pauli indeterminacy principle is super important for our world, because it is what ensures the stability of matter! That is, since particles comprising matter are all fermions, they can't occupy the same states, so they don't all try to collapse on each other, instead (often) forming stable structures.

3.5. **Question: When is a system a composite?** Though Bose and Fermi statistics are the only things we encounter in the Standard Model, there are technically still other ways we could try forming composite quantum systems out of individual ones. E.g. there are other ways to combine some number of Hilbert spaces than through tensor, symmetric, or exterior powers. The point that I would like to draw your attention to though is that all of those constructions are standard ways of building new representations out of known ones. This is an important hint.

Now let's turn the story around on its head: suppose you are given a quantum system, by its Hilbert space $\mathcal{H}$ (remember: we are purposefully suppressing all questions about the Hamiltonian). How can we tell if it is a composite of some "smaller" quantum systems? Surely we should be able to answer this question, if we are to look for elementary particles ... Actually, the problems start earlier, on a much more fundamental issue.

3.6. **Question: When is the physics the same?** The issue is: how can we tell if several different mathematical descriptions are describing them same physical problem?

This non-uniqueness comes from factors such as different choices of units, coordinates, base-points, ground states, etc. There is much much more that can go "wrong", which is to say, we can simply approach the way we describe the same phenomenon in different ways.

Very naively: if some particle is moving in a circle, we could obtain a complete description of it by viewing its space of possible positions as S^1, and take the Hilbert space to be something like $L^2(S^1)$. But we could also describe it by keeping track of the distance it travels, in which case the Hilbert space would be something like $L^2(\mathbf{R}^1)$, with a periodicity condition imposed through other means in the system. Or we could view the circle as lying on a plane, or inside 3-space, hence getting candidate Hilbert spaces like $L^2(\mathbf{R}^2)$ or $L^2(\mathbf{R}^3)$. Certainly if the physical phenomenons that several quantum systems are describing is the same, then all the obtained numbers, which will be comparable with experiments, should be the same. Some of those numbers are the probability amplitudes $\langle\psi_2|\psi_1\rangle$. It follows that whatever the definition of "physical equivalence" should be, it should certainly preserve the Hilbert space structure.

You might be tempted to point out that all separable infinitely-dimensional Hilbert spaces are isomorphic to each other. But the issue is that there are no *canonical* isomorphisms. (Also, taking this argument at face value indicates that all quantum systems whatsoever are the same, which is certainly nonsense.) So that's the wrong alley to bark up - i.e. the notion of Hilbert space isomorphism is *not enough* to capture the notion of describing the same physical system.

3.7. **Answer: Symmetry!** One more thing we may do is to ask what *symmetries* the system possesses. The idea then being, that we will describe the system as something having the appropriate symmetries for the physical theory in question.

We discussed symmetries at length in chapter 2, so we know that (since we are ignoring the Hamiltonian) a symmetry of a quantum system with Hilbert space $\mathcal{H}$ consists of Lie group G and a unitary representation $G \to \mathrm{U}(\mathcal{H})$.

The key example would be particles inside our real-life friendly physical 3d space. However we phrase theory of a particle, since it is supposed to be a very small thingy in $\mathbf{R}^3$, and the precise way we lay down coordinates shouldn't matter, we should in particular have at least SO(3)-worth of symmetry, corresponding to rotating the coordinates (or if you want, 3-space itself). Indeed, you will expect more, also transnational symmetry, since space shouldn't have a distinguished base point, which is to say, if we shifted everything simultaneously in any direction, that shouldn't effect the physics. So altogether we might expect a symmetry in the shape of the Galilean group $G = \mathrm{SO}(3) \ltimes \mathbf{R}^3$, the group of affine isometries of Euclidean 3-space.

Now certainly if the unitary G-representation in question is reducible, then we break up the Hilbert space into $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, each of which are themselves G-reps. This is viewed as splitting up the system into smaller pieces! Thus to find the smallest possible systems, i.e. the fundamental particles, we should be looking for irreducible G-representations.

So let's say this makes some amount of sense. Of course there is no per se guarantee that keeping track of symmetries should be the right thing to keep track of. Alas, that turns out to be the case. :)

The hard question is which group G we should consider. If we really believed that we live in $\mathbf{R}^3$, the Galilean group $\mathrm{G} = \mathrm{SO}(3) \ltimes \mathbf{R}^3$ would seem like a great guess. But we don't *really* believe that ...

There are actually two disjoint problems:

(1) The rotation group SO(3) is wrong. While it agrees well with real-life physical experience, it is insufficient to describe spin phenomena detected at the quantum level. It needs to be replaced with the "larger" spin group Spin(3).
(2) The translation group $\mathbf{R}^3$ is wrong. We do not actually live in 3-space, but instead in 4-space of the form $\mathbf{R}^{1,3}$ - this is the lesson of Special Relativity, and if we wish to be consistent with it (as we should, since particle physics should consider small but fast-moving objects, and the closer to the speed of light movement is, the less accurate non-relativistic mechanics becomes), we should extend the translation group to $\mathbf{R}^{1,3}$.

Of course if we accept 2., then we presumably also have to replace the rotation group with its $\mathbf{R}^{1,3}$-analogue. We discuss 2., and what goes into it, in the next chapter, then 1. in the next chapter, before we will put them together.

4. Special Relativity in a Nutshell

So far we've been cheating a little. But at least one way in which we have cheated is a very standard way of cheating, that permeates most elementary physics: we were doing non-relativistic math!

In order to fix that, let us very briefly review a little about how the Special Theory of Relativity (SR for short) works. This will be a non-quantum story, concerning real-life scales of magnitude. Of course, we will afterwards say a bit about how to implement the lessons of SR into quantum theory.

4.1. **Why bother with this relativistic nonsense though?** The experimental evidence for SR is quite overwhelming; since many things these days travel fast enough for the relativistic effects to be non-negligible, and our measurement requirements are very precise, phenomena such as time contraction and length dilation can be quite conclusively perceived, and must be taken into account for things such as satellites to work as accurately as they do.

The starting point of SR is the mind-boggling (but experimentally extremely well-established) fact that the speed of light (in a vacuum) is constant for all observers. That is to say, no matter how fast you are yourself traveling, you will perceive the speed of light to travel at the same speed c.

Now it had been understood since the advent of Maxwell's equations that light should be viewed as some kind of wave. But in what medium? The great hypothesis of 18th century physics (perhaps reminiscent of string theory today?) was the existence of luminferous aether - a medium filling all of space through which light would travel, but which would interact with essentially nothing else. The conclusive evidence against it was provided the famous Michelson-Morley experiment - and the way it disproved the existence of the aether was by establishing the constancy of the speed of light!

From this base tenant, and some minimal physical assumptions, one can follow Einstein in deriving the mathematical description of SR. But since we are mathematicians, let's just skip straight to the chase, and give the most mathematically elegant way of talking about it (due to Minkowski).

4.2. **Minkowski spacetime.** Classical physics postulates (quite reasonably!) that the physical world is of the form $\mathbf{R} \times \mathbf{R}^3$. Hence physical events may be described by the *independent* quantities $t \in \mathbf{R}$, the location in time at which the event is occurring, and $\mathbf{x} \in \mathbf{R}^3$, its spatial position.

In SR on the other hand, the physical world is instead modeled mathematically by the *Minkowski spacetime* $\mathbf{R}^{1,3}$. This is again just $\mathbf{R}^4$ as a vector space (as indicated by the common name 4-*vectors* for its elements), but it comes equipped with an inner product of signature $(1,3)$, called by physicists the *spacetime interval*. That is to say, in the usual coordinates, if we have $x^\mu = (x^0, \mathbf{x}) = (x^0, x^1, x^2, x^3)$ and $y^\mu = (y^0, \mathbf{y}) = (y^0, y^1, y^2, y^3)$, then their spacetime interval is

$$x_\mu y^\mu := x^0 y^0 - \mathbf{x} \cdot \mathbf{y} = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3.$$

There is some disagreement in the physics literature as to whether the signature should be chosen as $(1,3)$ or $(3,1)$, but so far as I can tell, that is merely an aesthetic choice.

The idea is to view the coordinates x^1, x^2, x^3 as the spatial part of the event x^μ, and its time location to be specified as $x^0 = ct$. More precisely, we call a vector $x^\mu \in \mathbf{R}^{1,3}$

(1) *timelike* if $x_\mu x^\mu > 0$,
(2) *spacelike* if $x_\mu x^\mu < 0$,
(3) *lightlike* if $x_\mu x^\mu = 0$.

At least the first two should seems reasonable. Why we should consider something whose spacetime interval with itself is zero, as similar to light, we explain a little later (it should make sense of at least by the end of 4.8).

So far this looks almost precisely like the classical story. But the surprise in SR is that different observers won't agree on which part is the time component and which are the spatial components. Indeed, the direction in which "future" points inside $\mathbf{R}^{1,3}$ depends on the velocity at which one is traveling.

Indeed, fix two events in spacetime. Then different observers won't agree on how far apart the two events took place, nor how much time passed between them - it may even happen that one will perceive them as occurring simultaneously, while the other will see one happen before the other one. The notion of simultaneity is not physical! Wild, right?! :)

The only thing that the two observers *certainly will* agree upon is the spacetime interval between the two events - hence why the spacetime interval plays such a distinguished role!

In order to make any sense of this, we should say a little more about what observers are in SR.

4.3. **Worldlines.** Let us describe the *worldline* of a physical point particle, i.e. the path inside spacetime that it traces out as time unfolds. (Note: if you work on a big enough scale, you may very well model yourself as a point particle. That is to say, this is *not* about the microscopically small!) Such a worldline is a curve $x^\mu : \mathbf{R} \to \mathbf{R}^{1,3}$, such that for every $\tau \in \mathbf{R}$ (and really $\mathbf{R}$ could be replaced by a smaller interval too) the4-*velocity* $u^\mu := \frac{dx^\mu}{d\tau}$ is timelike or lightlike (we say that it is *causal*).

Let us assume for the moment that u^μ is actually always timelike, and not lightlike: thus that $u_\mu u^\mu > 0$. We will return to the lightlike case in 4.7.

We furthermore demand (as can always be achieved by reparametrization) that the curve be parametrized naturally, i.e. by arclength, so that $u_\mu u^\mu = c^2$ (the analogous situation is Euclidean signature is probably familiar from any basic differential geometry course, though with 1 instead of the universal constant c. The c is totally inessential though - if we chose "natural units" then we could have $c = 1$. But let's be classical.) - this is why we demanded timelike and not lightlike 4-velocity. Such a natural parameter τ we call the *proper time* that the observer moving along the worldline experiences. It is the time that the observer measures on his own clock, i.e. it is how time evolves from the perspective that the observer is not moving.

This is sort of like the age-old question: is it you who are moving, or is it the rest of the universe? Well, if you assume the second and measure time, then this is your proper time. In classical physics that happens to agree with "global time", i.e. the time component of $\mathbf{R} \times \mathbf{R}^3$. But in SR we must abandon the whole notion of global time, so such a hope is naive at best.

4.4. **In your perception: spacetime = time + (rest) space.** From your perspective, i.e. along the worldline that your life traces in spacetime, where you are again idealized as being spatially a point (if cows can be smooth spheres, why not!), your 4-velocity direction u^μ determines what you perceive as the *future* and the *past*. If some 4-dimentional being could ask you to point which "direction" future is, then you would point in the direction u^μ.

Conversely, if they asked you to indicate what you perceive as spatial directions, you would trace out the orthogonal complement of u^μ with respect to the spacetime interval. That, i.e. those x^μ such that $u^\mu x_\mu$, is called your *rest space*. The word "rest" is here again to indicate that those are the spatial directions perceived, if you assume that you are yourself not moving, i.e. at rest.

So from your perspective, or more generally the perspective of any point observer traveling along a worldine, spacetime decomposes as $\mathbf{R}^{1,3} = \mathbf{R}u^\mu \oplus (u^\mu)^\perp$ into a time component and space component. This decomposition (and even its origin point!) will of course change as τ changes, and of course obviously, it depends heavily on what u^μ is.

Given any event in spacetime, what you would do to measure it is orthogonality project the vector connecting it to your current origin onto your time direction and rest space, and measure those as the point in time and spatial location at which the event occurred.

4.5. **Relativistic 4-velocity vs. classical velocity.** Alright, that was the perspective from an observer's eyes. How about from the "global" perspective? After all, it is this global perspective that is taken in classical physics

Suppose thus as always that $x^\mu = (x^0, \mathbf{x})$ with $x^0 = ct$ the temporal and $\mathbf{x}$ the spatial location of the point particle. Then we get

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau}$$

and, introducing the notation $\gamma = \frac{dt}{d\tau}$ for the derivative of classical time t with respect to τ, we obtain the expression

$$u^0 = c\gamma, \qquad u^j = \gamma v^j,$$

for spatial indices $j = 1, 2, 3$ and "classical velocity" $\mathbf{v} = \frac{d\mathbf{x}}{dt}$. Thus 4-velocity may be expressed in terms of classical velocity as

$$u^\mu = (c\gamma, \gamma\mathbf{v}),$$

thus the auxiliary number γ, called the *Lorenz* factor, entirely encodes the difference between the classical and relativistic world.

So we care about the Lorenz factor. To express it, let us use the above formula for 4-velocity and the natural particularization assumption, we find that

$$c^2 = u_\mu u^\mu = (c\gamma)^2 - \gamma^2 v^2,$$

where we have done as physicists and denoted the magnitude of velocity $v = |\mathbf{v}|$. Turning this around, it gives a formula for the Lorenz factor, as

$$\gamma = \frac{1}{\sqrt{1 - (\frac{v}{c})^2}}.$$

This formula is very insightful! Among other things, it shows that the Lorenz factor depends only on the velocity v, and is as such sometimes denoted $\gamma = \gamma(v)$.

(CAVEAT: how did we conclude that γ is the positive branch of the square root? Really there is a little more assumed about observers: since time components split into "past" and "future", we assume that observers can tell which of the two is "future", and which is "past". Technically this amounts to picking a so-called time orientation on spacetime. This is very harmless, but it should be pointed out, lest confusion arise.)

4.6. **Physical consequences of the Lorenz factor formula.** Let us collect some things we may conclude from the Lorenz factor formula we derived in the last paragraph.

When $\frac{v}{c} << 1$, i.e. when the speed of motion v is much smaller than the speed of light c, then $\gamma \approx 1$ is a good approximation. In that approximation, we get $u^\mu = (c, \mathbf{v})$. So at very low speeds, relativistic corrections are negligible, and the predictions of classical physics (which use $\mathbf{v}$) will be good.

But the moment the motion in question beings approaching speeds v non-negligibly close to the speed of light c, the Lorenz factor starts effecting thing. In particular, the Lorenz factor starts growing. Since by definition $\gamma = \frac{dt}{d\tau}$, this means that the the proper time τ, the time "measured of the observer", begins slowing down compared to "objective time" t. The faster you go, the slower you perceive time. This is the oh-so famous (experimentally thoroughly verified) phenomenon of time dilatation! :)

On the other hand, the formula $\mathbf{u} = \gamma\mathbf{v}$ for the spatial component of the 4-velocity implies that spatial lengths the observer perceives (measured relative to the time + rest space decomposition, which is tuned to u^μ) contract against the "objectively measured" lengths. This is the equally famous phenomenon of length contraction.

4.7. **Traveling at the speed of light.** We demanded that the worldline along which an observer traveled in the above discussion of the several preceding sections to point everywhere in timelike directions. How about if they pointed into a lightlike direction? Then the Lorenz factor γ explodes to ∞, indicating that the classical approximation $\mathbf{v}$ becomes entirely useless.

But suppose it wasn't, just for a second. Then we would still have $u^\mu = (c\gamma, \gamma\mathbf{v})$, and since $0 = u_\mu u^\mu = \gamma^2(c^2 - v^2)$, it must be that $v = c$. That is to say, even though the classical velocity $\mathbf{v}$ doesn't make much sense, the number v, the speed = magnitude of velocity, does, and is equal to the constant c, the speed of light.

By design, this is the fastest achievable speed in SR - but it is also the speed where certain things we're used to from everyday experience decidedly stop making sense. In comparison, they only have to be appropriately scaled when traveling at speeds close to c, but when *actually traveling* at that speed, a lot of classical ideas totally give out. The notion of future and past still make sense (the 4-vector u^μ versus $-u^\mu$), but little else does.

For instance, if we were to try defining the rest space as the orthogonal complement to u^μ, then u^μ would itself belong to it. So the future direction, that into which u^μ is pointing, is a perceived "spatial" direction? Nonsense!

4.8. **Constancy of the speed of light.** The main assumption of SR is that the speed of light is constant for all observers. So let's take a moment to verify that this is indeed the case with the Minkowski space approach.

Suppose that a worldline x^μ_ℓ is pointing everywhere in a lightlike direction, so that it is encoding something traveling at the speed of light, a photon say. Let x^μ be another observer moving slower than light, hence traversing a worldline with everywhere timelike direction. Suppose that they meet at some point, so that when they occupy the same sport, x^μ can take a measurement of the velocity of x^μ_ℓ. At what speed does x^μ perceive x^μ_ℓ as traveling?

In terms of 4-velocities, this amounts to asking what is the magnitude of the orthogonal projection of u^μ_ℓ onto the rest space of x^μ, i.e. the orthogonal complement of the 4-vector u^μ. Technically there is one small caveat more: before orthogonally projecting, the 4-vector u^μ_ℓ must first be scaled so that its time-component wrt x^μ, i.e. the component into the u^μ – *direction*, is equal to u^0. That is to say, the measurement that the observer x^μ makes is performed under the (albeit mistaken!) assumption that x^μ_ℓ perceives time as pointing in the same direction and going just as fast as x^μ perceives it.

Arguing just like in 4.5, this means that (after appropriate rescaling!), if $u^\mu_\ell = (u^0_\ell, \mathbf{u}^0_\ell)$ is the time + rest space decomposition with respect to u^μ, then

$$u^0_\ell = u^0 = c\gamma, \qquad \mathbf{u}^0_\ell = \gamma \mathbf{v}_\ell,$$

where $\mathbf{v}_\ell$ is the "classical velocity" that x^μ perceives x^μ_ℓ as traveling at, and $\gamma = \gamma(v)$ is the Lorenz factor *with respect to the velocity of* x^μ. The fact that u^μ_ℓ is lightlike is a geometric fact independent of the coordinates chosen, thus

$$0 = (u_\ell)_\mu (u_\ell)^\mu = (c\gamma)^2 - \gamma^2 v_\ell^2 = \gamma^2 (c^2 - v_\ell^2).$$

Since the observer x^μ is traveling at a speed $v \not\leq c$, the Lorenz factor $\gamma = \gamma(v)$ is finite, and so we must have $v_\ell = c$.

In conclusion, the observer x^μ measures the speed of anything moving along a lightlike-direction worldline to always be equal to the speed of light c. We made no assumptions whatsoever on the observer x^μ, so this is a universal fact: the speed of light is constant for all observers!

With this, we have finally also justified why 4-vectors with vanishing spacetime interval should indeed be called lightlike.

4.9. **Momentum, mass, and energy.** We've been talking about the 4-velocity quite a lot by now. But in our discussion in quantum land in sections 1.7 and 2.11, we preferred to discuss momentum instead of velocity. Recall from classical physics that, if $\mathbf{v}$ is the velocity vector, and the object in question has mass m, then its momentum is defined as $\mathbf{p} = m\mathbf{v}$.

Cool trick; let's do it in SR too! We define the 4-*momentum* simply as $p^\mu = mu^\mu$, where the mass $m \in \mathbf{R}_{\geq 0}$ is simply an additional constant that needs to be specified for every individual object. In light of the formulas of 4.5, the relationship between the 4-momentum p^μ and the classical momentum $\mathbf{p}$ is given as $p^\mu = (mc\gamma, \gamma\mathbf{p})$.

If the worldline in question points everywhere in lightlike directions, which is to say the motion it describes is smaller than the speed of light, then by 4.3 we have $u_\mu u^\mu = c^2$. This implies that

$$p_\mu p^\mu = (mc)^2.$$

That might remind you of something.

Recall the famous Einstein energy-mass relation $E = mc^2$. Turning this on its head, we might be tempted to define the energy by it. Alas, this formula in this form only holds

when the object is at rest. Based on our experiences with SR so far, implies that the correct formula might include a Lorenz factor γ or something similar.

Indeed, the 4-momentum p^μ is also called the *energy-momentum 4-vector*. The relation to energy is given through the time-component as $p^0 = \frac{E}{c}$, which is to say $E = mc^2\gamma$ - precisely the Lorenz factor correction we guessed! Now the spacetime interval formula for p^μ above gives

$$(mc)^2 = p_\mu p^\mu = \left(\frac{E}{c}\right)^2 - p^2,$$

where we follow the physicists in using p^2 for norm of the spatial component of relativistic momentum $\gamma\mathbf{p}$.

Indeed, we mostly just discard the classical momentum as an obsolete notion, and use the notation $\mathbf{p}$ for the spatial component of the relativistic 4-momentum, perhaps calling it relativistic 3-momentum if there is any room for confusion. With that, the use of $p = |\mathbf{p}|$, the magnitude of the relativistic 3-momentum in the above formula, becomes unambiguous. With this convention, the components of the 4-momentum are $p^\mu = (\frac{E}{c}, \mathbf{p})$.

Note that the *energy-momentum relation* $p_\mu p^\mu = (mc)^2$, which we have already seen to be equivalent to

$$\left(\frac{E}{c}\right)^2 = (mc)^2 + p^2,$$

reduces for a stationary observer, which is to say when $v = 0$ and so $p = \gamma v = 0$, to the (unjustly!) more famous energy-mass relation $E = mc^2$, as promised. :)

4.10. **Lorenz transformations.** Okay, so that is basically how SR works. But how does the change of coordinates work in this setting?

In classical physics, we just change the coordinates on the spatial factor in $\mathbf{R} \times \mathbf{R}^3$, leaving the time factor unchanged. Natural changes of coordinates are thus on $\mathbf{R}^3$, and since lengths in space are physically relevant, the relevant transformations are the affine isometries $\mathrm{SO}(3) \ltimes \mathbf{R}^3$, the rotational factor $\mathrm{SO}(3)$ and translational factor $\mathbf{R}^3$. Since we know that in SR the speed of time has to be constant in all reference frames, i.e. even under the change of coordinates, this won't work in SR.

On the other hand, we know that the one physically meaningful quantity on which all observers agree is not the length of time or length in space, but the length in spacetime, in the form of the spacetime interval. Recalling that this is but the signature $(1,3)$ inner product, this suggests considering the group of *Lorenz transformations* $\mathrm{SO}^\uparrow(1,3)$. (The arrow subscript $\uparrow$ indicates that we require the transformations to be *ortochronous*: preserve the time orientation - they should not be able to switch the past and the future! This is fully analogous to the S prefix, which amounts to requiring the transforms to preserve (spatial) orientation.)

Actually those are only the linear isometries of $\mathbf{R}^{1,3}$, so the full change of coordinates group should actually be the *Poincare group* $\mathcal{P}_0 := \mathrm{SO}^\uparrow(1,3) \ltimes \mathbf{R}^{1,3}$. But we all understand how the translational part works, so let us focus on the Lorenz group.

Of course, tipping our hand, the goal of this whole digression into special relativity was precisely to find the Lorenz group. Indeed, to answer the question raised at the end of 3.7, if we believe that our world is relativistic (and we do!), then perhaps the Lorenz group (or perhaps the Poincare group $\mathcal{P}_0$) should be the group of symmetries we prescribe get a notion of physical equivalence between quantum systems. Simply: we'll be looking at the unitary irreducible representations of the Lorenz group! But first, let us briefly discuss what the elements of this group look like, as linear operators on $\mathbf{R}^{1,3}$.

4.11. **A quick tour of Lorenz transformations.** Let $g_{\mu\nu}$ be the matrix representing the spacetime interval, i.e. the $(1,3)$-signature inner product, in some coordinates $x^\mu on \mathbf{R}^{1,3}$ (i.e. with some choice of basis). That is to say, the spacetime interval is expressed as

$$x_\mu y^\mu = g_{\mu\nu} x^\nu y^\mu.$$

Any given Lorenz tranforms $\Lambda \in \mathrm{SO}^{\uparrow}(1,3)$ correspond with respect to this choice of basis by the usual formula $(\Lambda x)^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$ to a 4×4 real matrix $\Lambda^{\mu}{}_{\nu}$. Now the condition $(\Lambda x)_{\mu}(\Lambda y)^{\mu} = x_{\mu} x^{\mu}$ is easily seen to be equivalent to the requirement

$$\Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} g_{\mu\nu} = g_{\rho\sigma}$$

on the level of matrices. Choosing the usual "objective" basis for $\mathbf{R}^{1,3}$ of time + space, we get $g_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, and it becomes a simple linear algebra exercise to find all the matrices $\Lambda^{\mu}{}_{\nu}$ in question.

But if you don't feel like doing the work, here is what you would find. Lorenz transformations are comprised of two sorts (generated under composition, etc.):

(1) *Spatial rotations* (those you already know well!): note that a rotation $Q \in \mathrm{SO}(3)$ will preserve the norm $v = |\mathbf{v}|$ of the vector $\mathbf{v} \in \mathbf{R}^3$. Thus if $v^{\mu} = (v^0, \mathbf{v}) \in \mathbf{R}^{1,3}$ and we define $Qv^{\mu} = (v^0, Q\mathbf{v})$, then indeed

$$Qv_{\mu} Qv^{\mu} = (v^0)^2 - |Q\mathbf{v}|^2 = (v^0)^2 - v^2 = v_{\mu} v^{\mu},$$

as required of a Lorenz transformation.

(2) *Lorenz boosts*: these are obtained by switching from "stationary" coordinates to the coordinates as perceived by an observer who, according to the previous coordinates, is moving at a uniform velocity in a fixed direction. Taking the direction for simplicity to be the x-direction, and the speed of motion to be v, then the Lorenz boost is given by $(t,x,y,z) \mapsto (t',x',y',z')$ with

$$\begin{aligned} t' &= \gamma(t - \tfrac{vx}{c^2}) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned}$$

Boosts in y and z directions work analogously, and boosting in an arbitrary direction can be accomplished by linearity.

The cool thing about Lorenz boosts is that applying them transforms you from a perspective in which something you care about is moving, into one where that thing is at rest (and "everything else" is moving instead). This is very useful - it means that it often suffices to consider stationary things, hence put ourselves into a time + rest-frame coordinates of $(x^0, \mathbf{0})$, and then boost to wherever else we wish.

4.12. **Lorenz boosts are just hyperbolic rotations.** Btw, though we are presenting them separately, spatial rotations and Lorenz boosts are really much of the same ilk. To see the similarity, consider rotation by an angle θ around the z-axis. It is given by $(t,x,y,z) \mapsto (t',x',y',z')$ where

$$\begin{aligned} t' &= t \\ x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \\ z &= z'. \end{aligned}$$

This is the rotation in the (x,y)-plane. Similarly the Lorenz boost in the x-direction written above may be viewed as a hyperbolic rotation in the (t,x)-plane. In terms of the "hyperbolic angle" $\theta = \mathrm{arctanh}^{-1}\left(\frac{v}{c}\right)$ (which can, unlike the periodic polar angle, be any real number), the Lorenz boost is given as

$$\begin{aligned} ct' &= ct\cosh\theta - x\sinh\theta \\ x' &= ct\sinh\theta + x\cosh\theta \\ y' &= y \\ z' &= z. \end{aligned}$$

The only difference is that, since the $(1,3)$-inner product restricts to a signature $(1,1)$-inner product on the (t,x)-plane, unlike the signature $2 = (0,2)$ on the (x,y)-plane, we must use hyperbolic trigonometric functions (which are attuned to signature $(1,1)$, as the identity $\cosh^2\theta - \sinh^2\theta = 1 shows$) instead of ordinary ones (where the analogous formula is $\cos^2\theta + \sin^2\theta = 1$, hence signature 2).

This concludes our tour of SR, which is to say, addressing issue 2. raised in section 3.7. Next we turn to addressing issue 1.

5. Spin

In this chapter, we pretend to forget the lessons of SR absorbed in chapter 4, and pretend that our world is a nice $\mathbf{R}\times\mathbf{R}^3$. Only in chapter 6 will we explore how to incorporate the ideas discussed in this one into a relativistic context.

5.1. **The elephant in the room.** Alright, time to broach a topic we have strategically avoided for a while now. It's one that you are surely aware of though, even if you don't know what it means: quantum particles possess *spin*.

5.2. **The Stern-Gerlach experiment.** As you've likely come to expect by now, this new complication comes about once again not from some mathematical playing around with the theory built so far, but instead from experimentally observed phenomena the theory fails to explain, providing evidence that the theory so far is incomplete.

The experiment in question was the (very famous) Sten-Gerlach experiment, which found that small particles (in the experiment: silver atoms, but electrons would do just as well) are susceptible to magnetism in an unexpected way. This was most similar to the way that magnet rotating on an axis would be effected by magnetism. Not *quite* the same though: instead of the rotational axis of a classical magnet, which could point in an direction in 3-space, the experimental effects seemed to coincide with only two possible alignments: up and down. This discreteness of possible experimental outcomes was familiar by then though - it had to do with something quantum!

(Another weird thing though was that this would be the case regardless along which axis the measurement was performed: an ordinary rotating magnet would be effected more the closer you would come to the axis of rotation. But the SG measurements always produced either up or down.)

5.3. **Pauli's solution.** The solution proposed by Wolfgang Pauli was to modify the Hamiltonian, to introduce a spin term.

There is a subtlety: in order for everything to work out, the $\mathrm{SO}(3)$-symmetry we discussed in 2.12, giving rise through Noether's theorem to angular momentum observables, has to be enlarged. Enlarged how though? Well, on the level of Noether's current, simply adding some more $\mathfrak{so}(3)$-indexed conserved observables. These observables $\hat{S}_j$, $j = 1,2,3$, are the *spin operators*, and (a constant factor of) their expected value is the spin number you might have heard of in chemistry.

Instead of having only external (so-called orbital) angular momentum, coming from the rotational movement of a particle (this is the one described in section 2.12), certain quantum systems, such as the one describing an electron, also have internal (called: spin) angular momentum.

5.4. **The spin group.** But when translating this extended symmetry from Lie algebras to Lie groups, things just don't work out! The issue is in some factors of $\frac{1}{2}$, necessary in the formulas for the spin observables, which simply lack meaning when exponentiated to $\mathrm{SO}(3)$.

Said more formally: the Lie algebra representation of $\mathfrak{so}(3)$, that the spin operators $\hat{S}_j$ had to satisfy for everything to work out, does not arise from a representation of the Lie group $\mathrm{SO}(3)$.

When phrased like that, the solution is obvious: extend the group of symmetries from the ordinary rotation group SO(3) to the spin group Spin(3), its simply connected double cover.

5.5. **What this means.** This has some bizarre consequences. It means that electrons are perceptible to a quality of rotation which we simply do not perceive on the macroscopic scale. It is possible to rotate an electron around 2π degrees in a way which switches its spin polarization, say from up to down, which then means that another such 2π degree rotation is required to get it to return to its initial state. From our perspective, it had to be rotated around *twice* to return to how it was.

This might remind you of something though: in section 3.5, we discussed the law of Fermi statistics, which means that the wave functions for pairs of certain particle types (such as electrons!) *anticommute* when positions of the two particles are switched. This is in contrast to the Bose statistic of section 3.4, where the wave functions commute for such a switch instead, and which aligns with classical (hence: macroscopic) expectations. Thus from this point of view, fermions must be exchanged twice to return to the same state. This sounds a lot like this spin business, and furthermore switching can be accomplish by a rotation ...

And while that is not a convincing explanation, we have stumbled upon a real connection here: the Spin-Statistics Theorem of Quantum Field Theory (so a little bit more sophisticated than our simple QM language) says that its spin properties and whether the QFT in question satisfies the Bose or the Fermi statistic, determine each other.

5.6. **Okay, but what does it mean for things to have different spin?** To really understand the question in the title, we must talk about the representations in question. Luckily the unitary irreps of Spin(3), which is to say of $\mathfrak{so}(3)$, are quite easy. The punchline is that they are indexed by half-integers, one for each $s \in \frac{1}{2}\mathbf{Z}_{\geq 0}$. (You might ask why half-integers and not integers, and it's mostly a matter convention. Mathematically a bit bizarre, but so deeply entrenched in the physics-speak that it's here to stay). We'll say a little about what these irreps are in the next section, but for now feel free to treat it like a black box.

With that in mind, let's look at your quantum theory describing some quantum particle. As discussed in 3.7, it should in particular give rise to a unitary irrep of whatever symmetry we deem appropriate. In the discussion of 3.7 we suggested SO(3), but in light of the preceding discussion in this chapter, we find that we must consider Spin(3). Then the type of the irrep in question will give a number $s \in \frac{1}{2}\mathbf{Z}_{\geq 0}$, and *this* is the spin of the particle type in question.

The two different qualitative behaviors are when spin is even and odd - this is the context of the Spin-Statistics Theorem alluded to above. If spin (of a particle in a QFT with enough symmetry) is even, then the particle must be a boson, and if it is odd, then it must be a fermion. You will notice that even spin representations descend to SO(3)-representations, and so they behave "like we would expect", based on our macroscopic experience with rotations.

The spins encountered "in nature", more precisely, in the Standard Model of particle physics, are only 0, $\frac{1}{2}$, and 1. All the "normal" particles are either $\frac{1}{2}$ (the fermions) or 1 (the bosons), only the Higgs boson has spin 0. If there were such a thing as a graviton, a particle mediating the gravitational force (and making this work well is very problematic), it would have spin 2.

5.7. **The exceptional isomorphism.** Okay, let's talk a little bit about and around the irreps of the 3-dimensional spin group. This is something of a digression, but a fun and mathematical one. The key is the *exceptional isomorphism* $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$, which can itself be explained in may ways

(1) Recall the quaternions, which are as a vector space $\mathbf{H} = 4\mathbf{C}^2$, but with a cooler algebra structure. We may identify $\mathbf{R}^3$ with the "purely imaginary" quaternions, the 3d subspace of $\mathbf{H}$ spanned by i, j, k. This inclusion preserves the norm: the Euclidean norm agrees with the quaternionic norm, which is in terms of $\mathbf{H} = \mathbf{C}^2$ just $|(z, w)|^2 = |z|^2 + |w|^2$. Thus the SU(2)-action, which $\mathbf{R}^3$ inherits as imaginary quaternions from $\mathbf{C}^2 = \mathbf{H}$, preserves both the Euclidean norm and orientation, inducing a group homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$. Describing everything in formulas, one can see that this is a $2:1$ cover, and hence exhibits an isomorphism $\mathrm{SU}(2) \simeq \mathrm{Spin}(3)$.

(2) Less quaterniony (though of course just a slight redressing), one can introduce the *Pauli matrices*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These three friendly guys are to $\mathfrak{so}(3)$ kind of like what the trusty e, f, g are to $\mathfrak{sl}(2)$. It is easy to check that they are Hermitian, satisfy $\sigma_j^2 = 1$, and $\sigma_i\sigma_j = -\sigma_j\sigma_i = i\sigma_k$ where $i \neq j$, and the index k is the "next one" in the cyclic ordering $1, 2, 3$. Thus

$$[\sigma_i, \sigma_j] = 2i\sigma_k$$

(please forgive the two appearances of i, as the index and as the imaginary unit - I trust it's not too hard to tell them straight though). Thus the matrices $\frac{1}{2i}\sigma_j$ satisfy the commutator relation of $\mathfrak{so}(3)$, and since they are linearly independent, provide a basis for it.

The Pauli matrices define a map $\mathbf{R}^3 \to \mathbf{C}^{2\times 2}$ given by $\mathbf{x} \mapsto \vec{\sigma} \cdot \mathbf{x}$ for the "Pauli vector" $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, which is to say that

$$\vec{\sigma} \cdot \mathbf{x} = \begin{pmatrix} x^3 & x^1 + ix^2 \\ x^1 - ix^2 & -x^3 \end{pmatrix}$$

Clearly this matrix is Hermitian too, and so defines a map $\vec{\sigma} \cdot - : \mathbf{R}^3 \to \mathrm{Obs}(\mathbf{C}^2)$, the space of 2×2 Hermitian complex matrices, i.e. the observables of the quantum system with only two degrees of freedom (the qubit). In any event, the cool thing is that $\det(\vec{\sigma} \cdot \mathbf{x}) = -|\mathbf{x}|^2$, as a super easy computation shows. Now unitary special complex 2×2-matrices act on Hermitian ones by conjugation, so this defines an map $\mathrm{SU}(2) \to \mathrm{SO}(3)$. The conjugation description makes it obvious that this map is a double cover: it sends for any $U \in \mathrm{SU}(2)$ both U and $-U$ to the same element of $\mathrm{SO}(3)$. Thus it also exhibits an isomorphism $\mathrm{SU}(2) \simeq \mathrm{Spin}(3)$.

There are many ways in which description 2. is better, primary one being that it is more explicit. Also, you might have asked why we built those weird halves into the Pauli matrices. Just like those halves with the spin reps ... :) Lastly, as we will see in the next chapter, unlike approach 1. it generalizes from the Euclidean also to the Lorenz signature.

Supposing we believe the exceptional isomorphism, it seems quite reasonable to look for $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$-irreps among two-variable complex polynomials. The action occurs on the level of the variables, and each homogeneous degree $d = 2s \in \mathbf{Z}_{\geq 0}$ is an irrep - the spin $s \in \frac{1}{2}\mathbf{Z}_{\geq 0}$ irrep! In terms of $\mathrm{SO}(3)$, the only ones to make sense are the ones with $s \in \mathbf{Z}_{\geq 0}$, which is to say $d \in 2\mathbf{Z}_{\geq 0}$, and they can be expressed in terms of so-called spherical harmonics: the restrictions to $S^2 \subset \mathbf{R}^3$ of harmonic polynomials in 3 real variables. The details won't matter for our purposes very much, but it is good to know that these are rather friendly and explicit objects.

The spin $\frac{1}{2}$ representation may be identified with the usual action of $\mathrm{SU}(2)$ on $\mathbf{C}^2$. That is to say, in terms of perspective 2. above, the spinor representation is obtained from the Pauli sigma matrices acting on $\mathbf{C}^2$.

5.8. **Pauli matrices and Weyl spinors.** There is a rather interesting consequence of the exceptional isomorphism $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$, which gives us explicit understanding of the spin group. In order to have spin appear in the physics, we must include the sigma matrices into the Hamiltonian of the theory.

But the Hamiltonian $\hat{H}$ must itself be a Hermitian operator on the Hilbert space $\mathcal{H}$ of the wave-functions of the particle's position. If we try to take $\mathcal{H} = L^2(\mathbf{R}^3)$, there is simply no natural way for $\mathrm{SU}(2)$ to act. So instead, we must extend the Hilbert space, perhaps to $\mathcal{H} = L^2(\mathbf{R}^3; \mathbf{C}^2)$ - the L^2-space of $\mathbf{C}^2$-valued functions instead of just $\mathbf{C}$-valued as before. Now $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$ can act point-wise on $\mathbf{C}^2$ in its usual way. This allows us to treat spin $\frac{1}{2}$ particles, also called *(Weyl) spinors*.

5.9. **Spin polarization.** In general for a spin $s \in \frac{1}{2}\mathbf{Z}_{\geq 0}$ particle, if S_s denotes the spin s representation of $\mathrm{Spin}(3)$, the correct Hilbert space is $\mathcal{H} = L^2(\mathbf{R}^3; S_s)$. Thus spin 0 particles indeed have $\mathcal{H} = L^2(\mathbf{R}^3)$, but wave-functions for all higher spin particles have values in bigger vector spaces S_s instead of just the scalars $\mathbf{C}$. This encodes the inner degrees of freedom of a particle: it is not determined merely by its position, but also by its *spin polarization*.

For spin $\frac{1}{2}$-particles, polarization is determined by an element of $S_{\frac{1}{2}} = \mathbf{C}^2$. As a Hilbert space, $\mathbf{C}^2$ represents the QM system of the qubit, the system with only two independent states (recall: only unital vectors count), denoted in and around quantum computing as $|0\rangle$ and $|1\rangle$, but in the context of spinors usually $|\uparrow\rangle$ and $|\downarrow\rangle$: up or down. Thus a particle can be polarized in two ways: up or down, but the QM way of encoding that requires having wave-functions be L^2 maps $\psi : \mathbf{R}^3 \to \mathbf{C}^2$. With this, Pauli's correction to QM made the idea of viewing elements of the Hilbert space as probability distributions for the position of the particle slightly less straight-forward.

5.10. **Particles with spin on manifolds.** This is a wee bit of a digression from our particle story, alas perhaps a worthwhile one.

Since we feel that our physical space (at least on small enough scale, such as concerning a lab in which an experiment takes place) is well-modeled by $\mathbf{R}^3$, we were happy to consider only particles moving on $\mathbf{R}^3$. If instead we wished to consider a spin 0 = scalar particle moving on a fancier smooth manifold M (perhaps obtainable by imposing some constraints, etc.), then we would take the Hilbert space $\mathcal{H} = L^2(M)$, the Hilbert space of L^2-functions $\psi : M \to \mathbf{C}$. Well, in order to be able to integrate over M well (as we need to make sense of the L^2-condition), we should take it to be an oriented manifold, but that is the only requirement.

But how to do higher spin?

When spin is integral, this is straightforward. That is to say, for spin 1 spinors (= vectors) this is super easy: you should take $\mathcal{H} = L^2(M; TM)$, the L^2-sections of the tangent bundle, i.e. L^2-vector fields (though this must be done over $\mathbf{C}$, so there is an implicit extension of scalars $\otimes_{\mathbf{R}}\mathbf{C}$ happening).

But how to do half-integer spin? Well, to be able to do so, we need a consistent choice of Pauli sigma matrices σ_j (satisfying appropriate $d = \dim M$-dimensional analogues of the commutation relations given above in 5.7.2 for $d = 3$) on each tangent space $T_x M$, varying smoothly with $x \in M$. This is equivalent to specifying a spins structure on M.

The mathematical way to say it is: it is if the tangent bundle $TM \to M$ is classified by a map $f_{TM} : M \to \mathrm{BSO}(d)$, then a *spin structure* on M is a factorization $M \to \mathrm{BSpin}(d)$ of f_{TM} through the canonical map $\mathrm{BSpin}(d) \to \mathrm{BSO}(d)$ induced by the $2:1$ covering group homomorphism $\mathrm{Spin}(d) \to \mathrm{SO}(d)$. Without the language of classifying spaces, this is the same thing as saying that as a lifting of the frame bundle $F_{\mathrm{SO}}(TM) \to M$ to a principal $\mathrm{Spin}(d)$-bundle $F_{\mathrm{Spin}}(TM) \to M$. Using the spin representation $S = S_{\frac{1}{2}}$ of $\mathrm{Spin}(d)$ (for $d = 3$, this is $S = \mathbf{C}^2$ with the action of $\mathrm{SU}(2) \simeq \mathrm{Spin}(3)$), we may associate to this principal

bundle the vector bundle

$$\mathbb{S}_M := S \times_{\mathrm{Spin}(n)} F_{\mathrm{Spin}}(TM)$$

and this is the *spinor bundle* on M. To obtain spinors, i.e. spin $\frac{1}{2}$ particles, we then take the Hilbert space $\mathcal{H} := L^2(M; \mathbb{S}_M)$.

But while the previous paragraph is just a drizzle of nonsense on its own (e.g. as commonly explained by mathematicians), it is always driven by the idea of just giving yourself enough structure to be able to act by Pauli matrices on wave-functions of spinors.

I always found the notion of a spin structure very confusing and the mathematical explanation quite impenetrable without the physical picture, which is why I included this digression here. But feel free to forget all about it now, as we move on! :)

6. Particles in Relativistic Quantum Mechaniscs

At last we have come to the point where we mix together all the ingredients!

6.1. **Spin in the Lorenz signature.** So far in the previous chapter, we have seen that $\mathrm{SO}(3)$ needs to be generalized to $\mathrm{Spin}(3)$ to have a real chance at capturing the symmetry of a quantum particle. But this is totally ignoring the lessons we Chapter 4, where we learned that SR forces us to replace $\mathrm{SO}(3)$ with the Lorenz group $\mathrm{SO}^\uparrow(1,3)$.

So what do we do? Well, replace $\mathrm{Spin}(3)$ with $\mathrm{Spin}(1,3)$, the universal cover of $\mathrm{SO}^\uparrow(1,3)$. This actually adds a little bit too much: it is a $4:1$ cover, adding not only the "spin rotation" part, but also a "time reversal" part to it. It kind of depends on what you are doing, but you well might prefer to not include the former, as certain (rather subtle weak-nuclear-force-concerning) experimental observations suggest that it is not satisfied by nature. That is, the weak interaction works depends on the arrow of time - which way is future and which way is past. Before this was discovered, it was commonly believed that time-reversal is a real symmetry on nature!

6.2. **Another exceptional isomorphism.** Anyway, let's ignore this issue for a second and ask ourselves what $\mathrm{Spin}(1,3)$ is like, compared to how we know its Euclidean cousin to be $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$.

The exceptional isomorphism of 5.7 comes to the rescue: modifying the construction 5.7.2 slightly, we define the 0-th Pauli matrix as $\sigma_0 = 1$, the identify 2×2-matrix, and define a map $\mathbf{R}^{1,3} \to \mathbf{C}^{2\times 2}$ via

$$x^\mu \mapsto \sigma_\mu x^\mu = \begin{pmatrix} x^0 + x^3 & x^1 + ix^2 \\ x^1 - ix^2 & x^0 - x^3 \end{pmatrix}.$$

A trivial computation shows that $\det(\sigma_\mu x^\mu) = x_\mu x^\mu$, recovering the spacetime interval. In fact, as in 5.7.2, the matrix $\sigma_\mu x^\mu$ is Hermitian, and counting dimensions, the map $\sigma_\mu \cdot : \mathbf{R}^{1,3} \to \mathrm{Obs}(\mathbf{C}^2)$ is actually an isomorphism of normed spaces.

Now consider the action of $A \in \mathrm{SL}(2;\mathbf{C}) on \mathrm{Obs}(\mathbf{C}^2)$ as $X \mapsto AXA^\dagger$. Clearly it preserves the determinant, thus it induces a group homomorphism $\mathrm{SL}(2;\mathbf{C}) \to \mathrm{SO}(1,3)$, which induces the "exceptional" isomorphism

$$\mathrm{Spin}(1,3) \simeq \mathrm{SL}(2;\mathbf{C}).$$

This is a way in which our favorite Lie algebra enters physics: $\mathfrak{sl}(2) = \mathfrak{spin}(1,3) = \mathfrak{so}(1,3)$, is the infinitesimal Lorenz transforms.

6.3. **The extended Poincare group.** In section 4.10 we saw that the geometry of SR lead us to consider the relevant symmetry group to be the Poincare group $\mathcal{P}_0 = \mathrm{SO}^\uparrow(1,3) \ltimes \mathbf{R}^{1,3}$ of affine isometries of Minkowski spacetime. On the other hand, in light of chapter 5, we must modify the Lorenz group $\mathrm{SO}(1,3)$ and consider the spin group $\mathrm{Spin}(1,3)$ instead. But that neglecting the question of time reversal, so we should only restrict to the subgroup $\mathrm{Spin}^\uparrow(1,3) \subset \mathrm{Spin}(1,3)$ which preserves time-orientation through its action

on $\mathbf{R}^{1,3}$. Thus we have identified the rotational part, and adding the translational one, we obtain the *extended Poincare group* $\mathcal{P} := \mathrm{Spin}^{\uparrow}(1,3) \ltimes \mathbf{R}^{1,3}$.

Thus what we are really after is the unitary irreducible representations of $\mathcal{P}$. (Really what we are calling the "extended" Poincare group is more accurately the "ortochronous extended Poincare group", as we do not allow time-reversal. But we never will, so it's all good.)

6.4. **Poincare group representations, after Wigner.** To determine them, we follow a recipe of Winger.

(1) Find an irrep of the translational part $\mathbf{R}^{1,3}$. This is easy - it is an abelian Lie group, so its only irreps are the characters = 1-dim representations. That is, we wish to determine continuous homomorphisms $\mathbf{R}^{1,3} \to \mathrm{U}(1)$, and it is not very hard to figure out that the only possibility is to send

$$a^\mu \mapsto e^{-ip_\mu a^\mu}$$

for (an arbitrary) fixed 4-vector $p_\mu \in \mathbf{R}^{1,3}$. Thus we fixed what we wish to happen with the translational part, so we should leave it alone. To do that:

(2) Determine the "little group" $H := \mathrm{Stab}_{\mathrm{Spin}^{\uparrow}(1,3)}(p_\mu)$ WRT the "usual" action of $\mathrm{Spin}^{\uparrow}(1,3)$ on $\mathbf{R}^{1,3}$.

(3) Find a unitary irrep of H (hopefully a smaller and more manageable group), extend it via the chosen irrep (corresponding to) p_μ of the translational group, to a unitary irrep of $H \ltimes \mathbf{R}^{1,3}$.

(4) Perform induction along the inclusion $H \ltimes \mathbf{R}^{1,3} \subseteq \mathrm{Spin}^{\uparrow}(1,3) \ltimes \mathbf{R}^{1,3} = \mathcal{P}$ to obtain a unitary irrep of the Poincare group.

It is not immediate that the $\mathcal{P}$-representations that this procedure produces are actually irreducible, but that can be proved by verifying the criteria for Mackey's Criterion for irreducibility of induced reps (a reasonably basic result in rep theory). Also, all unitary irreps are produced this way - that's Wigner's Theorem.

BTW, the way we've set this up, once we're done, playing the Noether's Theorem game for the $\mathbf{R}^{1,3}$-translational symmetry, the Noether current will be (a quantum observable quantizing) p_μ, hence justifying the notation. Indeed, just as ordinary momentum 3-vector $\mathbf{p}$ is the Noether current for a Euclidean $\mathbf{R}^3$-translational symmetry, so is the 4-momentum p_μ the Noether current for Lorenzian $\mathbf{R}^{1,3}$-translational symmetry.

There is also a non-uniqueness in this procedure: different choices for p_μ may produce the same Poincare group reps. In particular, all 4-vectors related by Lorenz transforms will produce an equivalent character and have the same $\mathrm{Spin}(1,3)$-stabilizers. That means that we should only consider different orbits $[p^\mu] \in \mathbf{R}^{1,3}/\mathrm{SO}^{\uparrow}(1,3)$ above.

6.5. **Orbits of the Lorenz group acting on Minkowski spacetime.** We have come to a point where understanding the orbits of the $\mathrm{SO}^{\uparrow}(1,3)$-action on $\mathrm{R}^{1,3}$ would be beneficial. Those are easy to identify:

(1) The upper sheet of the two-sheeted hyperboloids

$$X_m^+ = \{p_\mu \in \mathbf{R}^{1,3} : p_\mu p^\mu = m^2, \quad p^0 > 0\}$$

for any $m > 0$.

(2) The lower sheet of the two-sheeted hyperboloids

$$X_m^- = \{p_\mu \in \mathbf{R}^{1,3} : p_\mu p^\mu = m^2, \quad p^0 < 0\}$$

for any $m > 0$.

(3) The future-pointing light cone

$$X_0^+ = \{p_\mu \in \mathbf{R}^{1,3} : p_\mu p^\mu = 0, \quad p^0 > 0\}.$$

(4) The past-pointing light cone

$$X_0^- = \{p_\mu \in \mathbf{R}^{1,3} : p_\mu p^\mu = 0, \quad p^0 < 0\}.$$

(5) The one-sheeted hyperboloid

$$Y_m = \{p_\mu \in \mathbf{R}^{1,3} : p_\mu p^\mu = -m^2\}$$

for any $m > 0$.

Let us interpret what these orbits mean physically, starting with the orbit X_m^+ from case 1.

As we explained near the end of 6.4, the role of the 4-vector p_μ is to play is as the 4-momentum. Hence recalling 4.9 for the relationship between the 4-momentum and mass, the condition $p_\mu p^m u = m^2$ means that we are considering a particle with mass m. Similarly the condition $p^0 > 0$ means that the 4-momentum $p_m u$, being a timelike 4-vector, is future-pointing. That is, X_m^+ corresponds to a particle with mass m moving forward in time. This orbit is usually called the *(positive) mass shell* in the physics literature.

Analogous consideration identify each of the above orbit types 1.-5. as encoding the following physical situations:

(1) particle with mass m passing forward through time,
(2) particle with mass m passing backward through time,
(3) massless particle passing forward through time,
(4) massless particle passing backward through time,
(5) particle with imaginary mass im.

Of these all matter is of the form 1., antimatter is of the form 2., photons are of the form 3. (though there's no reason why a photon couldn't be going backwards in time, but since photons are a bit bizarre to us anyway, we'd still call something satisfying 4. a photon), and so far as I am aware, 5. doesn't correspond to anything currently known in physics. Though, until relativity explained the photon as a massless particle, neither was 3. (or 4.), and until Dirac's relativistic EM predicted the positron, neither was 2. :)

BTW: note that if we considered $\mathrm{SO}(1,3)$-orbits instead, thus allowing for time reversal, orbit types 1. and 2., as well as 3. and 4. would become identified. This is particularly obvious from the physical interpretations of the orbits. On the other hand, the $\mathrm{SO}^\uparrow(1,3)$-orbits and the $\mathrm{Spin}^\uparrow(1,3)$-orbits do coincide. This will be important in a moment.

6.6. **The little group for $p_\mu \in X_m^+$.** So now we know what the orbit types are. But luckily we don't care about all of them!

We, in the context of the story we are trying to tell (but not as a species!) only care for electrons, which should in particular be massive particles moving forward in time, this means that we may restrict ourselves to the case when $p_\mu \in X_m^+$ for some fixed mass $m > 0$.

Following Wigner's program from 6.4, we must determine the little group H, i.e. the stabilizer of p_μ in $\mathrm{Spin}^\uparrow(1,3)$. Since p_μ is a timelike 4-vector, it is not hard to figure out that we have the freedom to mix around the remaining 3 space-like directions, thus identifying the little group as $H = \mathrm{Spin}(3)$.

Next Wigner tells us to find unitary irreps of the little group. We already know them from chapter 5., and know that they come indexed by a single half-integer $s \in \frac{1}{2}\mathbf{Z}$, the spin. Let S_s be the spin s representation of $\mathrm{Spin}(3)$. To obtain the sought-after irreps of the Poincare group, Wigner says to proceed by forming the representation of $H \ltimes \mathbf{R}^{1,3}$ by combining the H-irrep S_s and the $\mathbf{R}^{1,3}$-irrep corresponding to p_μ. That is to say, we take the underlying vector space of it to be S_s, with H acting on it as it is, and $a^\mu \in \mathbf{R}^{1,3}$ acting on it via

$$S_s \ni x \mapsto e^{-ip_\mu a^\mu} x \in S_s.$$

Thus the translational action of $\mathbf{R}^{1,3}$ is given through the multiplication with scalars $e^{-ip_\mu a^\mu} \in \mathbf{C}$ on S_s - note that the latter is a unitary irrep, and so in particular a complex

vector space. Now the last step of Wigner's plan is to form the induction $\mathrm{Ind}_{H\ltimes\mathbf{R}^{1,3}}^{\mathcal{P}}(S_s)$ of the $\mathbf{R}^{1,3}\ltimes H$-irrep S_s up to the extended Poincare group, through the group inclusion $H\ltimes\mathbf{R}^{1,3}\to\mathrm{Spin}^{\uparrow}(1,3)\ltimes\mathbf{R}^{1,3}=\mathcal{P}$.

6.7. **Intermezzo on induced representations.** Though we, as proud rep-theory-wielding mathematicians already know how induction works, let's brush it up quickly here anyway. This is mostly because we are interested here in induction for unitary representations, while the one we may be best familiar (I'm certainly speaking for myself here!) is induction for representations of finite (or really, more to the point: discrete) groups. In fact, let's brush up on that too!

6.7.1. *Induction for finite group reps, defined via abstract nonsense.* Fix an inclusion of finite groups $H\subseteq G$. Any G-rep V is obviously also a H-rep, with the same vector space and the restriction of the action map to the subgroup, so this G-rep is denoted $\mathrm{Res}_H^G(V)$ and called the *restriction of V from G to H*. This is obviously functorial, so we get the restriction functor $\mathrm{Res}_H^G:\mathrm{Rep}(G)\to\mathrm{Rep}(H)$, and the induction functor $\mathrm{Ind}_H^G:\mathrm{Res}(H)\to\mathrm{Res}(G)$ is most elegantly defined as its left adjoint. Though this makes induction sound like some kind of representation-theoretic (or even, categorical) voodoo, it's really something quite ordinary:

6.7.2. *Induction as a tensor product.* Recall that (let's assume we're working over $\mathbf{C}$ throughout) the group ring $\mathbf{C}[G]$ allow us to identify $\mathrm{Rep}(G)\simeq\mathrm{Mod}_{\mathbf{C}[G]}$, where $\mathbf{C}[G]$ itself corresponds to the regular representation. The subgroup inclusion $H\subseteq G$ induces an algebra map $\mathbf{C}[H]\to\mathbf{C}[G]$, and this ring map induces in the obvious way a forgetful functor on modules $\mathrm{Mod}_{\mathbf{C}[G]}\to\mathrm{Mod}_{\mathbf{C}[H]}$, which is nothing other than Res_H^G. For any ring map $A\to B$ the left adjoint to the forgetful functor $\mathrm{Mod}_B\to\mathrm{Mod}_A$ is given by $M\mapsto B\otimes_A M$. Thus we find that the functor Ind_H^G can be identified with the tensor product $V\mapsto\mathbf{C}[G]\otimes_{\mathbf{C}[H]}V$.

6.7.3. *Inductions as sections of a bundle.* Recall that the group ring $\mathbf{C}[G]$ may also be identifies with the (set-theoretic or continuous, due to discreteness of G it's all the same) functions $\mathrm{Map}(G,\mathbf{C})$. In light of that we get the identification of induction as $\mathrm{Ind}_H^G(V)=\mathrm{Map}(G,\mathbf{C})\otimes_{\mathrm{Map}(H,\mathbf{C})}V$. Explicitly this is spanned by equivalence classes of pairs (f,v) for $f:G\to\mathbf{C}$ and $v\in V$, where we identify $(f\circ h,v)=(f,h\cdot v)$ for all $h\in H$. In this condition h is viewed on the lift as a map $G\to G$ given by $g\mapsto hg$, while on the right the dot stands for the chosen H-action on V (it is an H-rep after all). It's now straightforward to identify this with the vector space $\Gamma(G/H;G\times_H V)$ of the (not-necessarily continuous) sections of the bundle $G\times_H V\to G\times_H 0=G/H$, and to see that the G-action is the one coming from the left action of G on G/H. Thinking of the finite set G/H as a bundle is a little silly, but as we'll see in 6.7.6, it works and is quite enlightening for infinite groups too!

6.7.4. *Induction for continuous groups.* But of course we know from chapters 1 and 2 that in physically-relevant situations, the symmetry groups in question won't be discrete, but continuous (or even more: a Lie group). We wish to take this continuous structure into account, which could be accomplished by considering continuous representations. But we already know since chapter 2 that the symmetries relevant for QM come packaged as unitary representations, hence we consider those instead!

We seemed not to have the unitarity condition in the discrete case, but at least for finite groups, that comes along for the ride as you can always make irreps unitary.

The story of restriction and induction works in the abstract, e.g. for unitary representations, just as in 6.7.1, but in order to proceed with the analogue of 6.7.2 and especially 6.7.3, we need to find an appropriate replacement for the group ring $\mathbf{C}[G]$.

6.7.5. $L^2(G)$ *is the relevant replacement for* $\mathbf{C}[G]$. It turns out that this is $L^2(G)$, the Hilbert space of L^2-functions $G \to \mathbf{C}$. Why this in particular? Well, if we only cared for continuous representations, surely we should do $\mathscr{C}(G, \mathbf{C})$ instead. But instead we wish to consider the theory of unitary representations, so in order to have access to the regular representation, we need to upgrade it to a unitary one. Thus complete $\mathscr{C}(G, \mathbf{C})$ along the friendly L^2-inner product, and voila, we get $L^2(G)$.

Another perhaps more convincing argument, goes by the name Peter-Weyl Theorem: at least for a compact group G (I know that the implicit harmonic analysis works well even for locally-compact groups, so maybe for them too?) the "regular unitary representation" $L^2(G)$ decomposes into a direct sum of unitary G-irreps - analogous to the situation with the usual regular representation of a finite group.

6.7.6. *Induction for unitary reps, continued.* Well, the story now proceeds analogous to 6.7.2 and 6.7.3 (one needs to complete the tensor product, but that's a technicality I'll ignore), and we get

$$\mathrm{Ind}_H^G(V) = L^2(G) \otimes_{L^2(H)} V = L^2(G/H; G \times_H V),$$

where of course the tensor product must be completed, and the last term denotes, in agreement with section 1.2 and everywhere else we used this convention, the L^2-sections of the (associated) vector bundle $G \times_H V$ over the quotient space G/H.

6.8. **Unitary irreps of the Poincare group.** Alright, equipped with understanding of induction, we return to where we left off in section 6.6. Actually, let's review what we have done in chapter 6 so far.

We learned that to determine a unitary irreps of the extended Poincare group, we must first choose a 4-vector $p_\mu \in \mathbf{R}^{1,3}$, which will end up being the momentum 4-vector. We noted near the end of 6.4 that this choice of p_μ only depends on its $\mathrm{Spin}^\uparrow(1,3)$-orbit, and since we are considering massive (vs. massless) and future-travelling particles, the orbit should be $p_\mu \in X_m^+$, a positive mass shell. The positive mass shells are characterized uniquely by the positive real number $m = \sqrt{p_\mu p^\mu}$, which we know from section 4.9 to be the mass of the particle in question.

But the mass does not determine the particle type alone: Wigner tells us to find a unitary irrep of the little group $H \subseteq \mathrm{Spin}^\uparrow(1,3)$. In 6.6 we found the little group in this case to be $H = \mathrm{Spin}(3)$, and so its unitary irreps are the spin s representations S_s for $s \in \frac{1}{2}\mathbf{Z}$. By using p_μ as a character, we extend S_s to a representation of $H \ltimes \mathbf{R}^{1,3}$, and now we must induct up to the (extended) Poincare group $\mathcal{P}$ to obtain the unitary irrep. In 6.7.6 we learned that this is the Hilbert space

$$\mathrm{Ind}_{\mathbf{R}^{1,3}}^{\mathcal{P}}(S_s) = L^2(\mathcal{P}/(H \ltimes \mathbf{R}^{1,3}); \mathcal{P} \times_{H \ltimes \mathbf{R}^{1,3}} S_s)$$

with it evident $\mathcal{P}$-action. This is the L^2-sections of a vector bundle with fiber S_s over the quotient space

$$\mathcal{P}/(H \ltimes \mathbf{R}^{1,3}) = (\mathrm{Spin}^\uparrow(1,3) \ltimes \mathbf{R}^{1,3})/(H \ltimes \mathbf{R}^{1,3}) = \mathrm{Spin}^\uparrow(1,3)/H = X_m^+,$$

where we recalled that H is the stabilizer of the 4-vector p_μ whose $\mathrm{Spin}^\uparrow(1,3)$-orbit is X_m^+, and used the Orbit Stabilizer Theorem for the final identification.

All in all, the unitary irreps of $\mathcal{P}$, which have any chance of encoding massive particles, are of the form

$$\mathcal{H}_{m,s} = L^2(X_m^+; \mathcal{P} \times_{\mathrm{Spin}(3) \ltimes \mathbf{R}^{1,3}} S_s),$$

encoding a particle of mass and spin $m \in \mathbf{R}_{>0}$ and $s \in \frac{1}{2}\mathbf{Z}_{\geq 0}$ respectively.

This is what we mean by: "a particle type is determined by its mass and its spin", the answer that a real-life physicist might give you to the question of "What is a particle?". Even if such physicist were to be married to a mathematician, who works in physics-related math, and whose answer would have to do with representations. So you see, both of their answers are equivalent, and now you understand why. :)

The initial question was for an electron, in which case we must determine the numbers m and s experimentally. It apparently turns out that $m \approx 9.109 \cdot 10^{-31}\,\mathrm{kg}$, and $s = \frac{1}{2}$. And there we are, we answered the title question! :)

6.9. **Comparison with the non-relativistic variant.** Before we go, let's compare the Hilbert space we found to underlie the relativistic QM of a massive particle, with the non-relativistic situation we discussed in section 5.9. There we asserted that the correct Hilbert space for the QM of a spin s particle is $\mathcal{H}_s = L^2(\mathbf{R}^{1,3}; S_s)$. This is quite similar to the relativistic answer $\mathcal{H}_{m,s} = L^2(X_m^+; \mathcal{P} \times_{\mathrm{Spin}(3) \ltimes \mathbf{R}^{1,3}} S_s)$, especially since:

- The mass shell X_m^+, being the upper sheet of the two-sheeted hyperboloid inside Minkowski space $\mathbf{R}^{1,3}$, is clearly diffeomorphic to $\mathbf{R}^3$.
- Both Hilbert spaces $\mathcal{H}_s$ and $\mathcal{H}_{m,s}$ are obtained as L^2-sections of a vector bundle with fiber the spin s-representation S_s of the spin group $\mathrm{Spin}(3)$.

There are also some differences though:

- The vector bundle is no longer trivial in the relativistic case. Instead, its twisting encodes the (infinitesimal) translational $\mathbf{R}^{1,3}$-action being given by the 4-momentum vector p_μ.
- While the non-relativistic version was simply defined to be the L^2-space "because that works", the relativistic version, while also returning a Hilbert space of L^2-sections, does so "from first principles". By that we mean that it is in line with the paradigm outlined in 3.7, of starting out with an expected minimal amount of symmetry required of the physical situation, and considering its unitary representations, irreducible ones for elementary particles. Indeed, since we are considering particles in spacetime, we identified in 4.10 the correct symmetries as the (in light of chapter 5, extended) Poincare group. And from that, following Wigner's construction of the relevant unitary irreps, the L^2-space appeared through the induction technology of 6.7, without "being put there by design".