

4 Transformations and Symmetries

While classical physics is firmly rooted in space-time, quantum mechanics takes place in the more abstract Hilbert space. Actions such as translations or rotations around a given origin have a straightforward space-time interpretation, but how can these affect a quantum particle living in Hilbert space?

In this chapter, we'll return to the general formalism of QM. We'll understand how transformations such as spatial translations or rotations affect states in Hilbert space, linking together the world of quantum mechanics with our familiar experience in 'normal' space \mathbb{R}^3 . In so doing, we'll obtain a deeper appreciation of the origin of many of the most common commutation relations, including most of the ones you're familiar with from 1B. We'll also understand *why* the momentum operator $\hat{P} = -i\hbar \partial/\partial x$, a fact which probably seemed rather mysterious in 1B QM.

4.1 Transformations of States and Operators

If we apply a transformation such as a translation or rotation to our physical system, we'd expect states to be different before and after the transformation. For example, if a particle is prepared in a state $|\psi\rangle$ that is strongly peaked around the origin $\mathbf{0} \in \mathbb{R}^3$ and we translate our system through a vector \mathbf{a} , after this transformation we should expect to find our particle in a new state $|\psi'\rangle$ whose wavefunction is strongly peaked around $\mathbf{x} = \mathbf{a}$. This suggests that the effect of any given spatial transformation should be represented on Hilbert space by a linear operator

$$U : \mathcal{H} \rightarrow \mathcal{H} \quad U : |\psi\rangle \mapsto |\psi'\rangle = U|\psi\rangle. \quad (4.1)$$

where the operator in question depends on the specific type of transformation.

Whatever state we start with, after applying the transformation, we still expect to find it *somewhere* in space. Thus, properly normalised states should remain so after applying U , or in other words

$$1 = \langle\psi|\psi\rangle = \langle\psi'|\psi'\rangle = \langle\psi|U^\dagger U|\psi\rangle \quad \text{for all } |\psi\rangle \in \mathcal{H}. \quad (4.2)$$

We'll now show that the requirement that U preserves the norm of our state fixes it to be a *unitary* operator¹⁶. The required trick is sometimes called the *polarization identity*. We first write $|\psi\rangle$ as $|\psi\rangle = |\phi\rangle + \lambda|\chi\rangle$ for some states $|\phi\rangle, |\chi\rangle \in \mathcal{H}$ and some $\lambda \in \mathbb{C}$. Equation (4.2) becomes

$$\langle\phi|\phi\rangle + \bar{\lambda}\langle\chi|\phi\rangle + \lambda\langle\phi|\chi\rangle + |\lambda|^2\langle\chi|\chi\rangle = \langle\phi|U^\dagger U|\phi\rangle + \bar{\lambda}\langle\chi|U^\dagger U|\phi\rangle + \lambda\langle\phi|U^\dagger U|\chi\rangle + |\lambda|^2\langle\chi|U^\dagger U|\chi\rangle \quad (4.3)$$

¹⁶Because physical states are represented by rays in \mathcal{H} , rather than actual vectors, there is one other possibility, transformations may be represented by operators that are both *anti-linear*, in the sense that $U(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = \bar{c}_1 U|\psi_1\rangle + \bar{c}_2 U|\psi_2\rangle$, and *anti-unitary*, meaning that $\langle\chi|U^\dagger U|\phi\rangle = \langle\phi|\chi\rangle = \overline{\langle\chi|\phi\rangle}$ for all $|\phi\rangle, |\chi\rangle \in \mathcal{H}$. Such antiunitary operators turn out to be related to reversing the direction of time; we will not discuss them further in this course. Wigner showed that all transformations are represented by either linear, unitary operators or anti-linear, anti-unitary ones.

where we recall that the inner product is antilinear in its left entry. By assumption $\langle \phi|U^\dagger U|\phi\rangle = \langle \phi|\phi\rangle$ and $\langle \chi|U^\dagger U|\chi\rangle = \langle \chi|\chi\rangle$, so this simplifies to

$$\lambda \left(\langle \phi|\chi\rangle - \langle \phi|U^\dagger U|\chi\rangle \right) = \bar{\lambda} \left(\langle \chi|U^\dagger U|\phi\rangle - \langle \chi|\phi\rangle \right). \quad (4.4)$$

In order for this to hold for *every* $|\psi\rangle \in \mathcal{H}$, it must continue to hold as we vary λ . In particular, as the phase of λ varies, the only way for (4.4) to be satisfied is if

$$\langle \phi|\chi\rangle = \langle \phi|U^\dagger U|\chi\rangle. \quad (4.5)$$

Finally, for this hold for arbitrary states $|\phi\rangle$ and $|\chi\rangle$ we must have $U^\dagger U = 1$. Multiplying through on the right by U^{-1} shows that

$$U^\dagger = U^{-1} \quad (4.6)$$

which says that the adjoint of U is equal to its inverse. Operators with this property are said to be *unitary*¹⁷.

There's one further condition on our operators U . Transformations of space such as translations and rotations form a *group*. For example, the composition of two rotations is again a rotation, the trivial rotation forms the identity, and the inverse of a rotation is a rotation of the same amount around the same axis, but in the opposite sense (*e.g.* clockwise instead of anticlockwise). Let's suppose that our spatial transformations form a group G . To reflect this group structure, our operators should¹⁸ provide a *homomorphism* from G to the group of unitary operators in the sense that, for all $g_1, g_2 \in G$,

$$U(g_2) \circ U(g_1) = U(g_2 \cdot g_1) \quad \text{and} \quad U(1) = 1_{\mathcal{H}} \quad (4.7)$$

where \cdot denotes the group multiplication in G and \circ denotes the composition of linear operators acting on \mathcal{H} . (We'll typically suppress these composition symbols from now on.) Note that if U_1 and U_2 are each unitary operators, then $(U_2 U_1)^{-1} = U_1^{-1} U_2^{-1} = U_1^\dagger U_2^\dagger = (U_2 U_1)^\dagger$ and so the composite operator $U_2 U_1$ is also unitary. Note also that the identity operator $1_{\mathcal{H}}$ is trivially unitary.

A particularly important class of transformations are those that depend smoothly on some parameter θ . For example, we can smoothly vary both the axis about which and the angle through which we rotate, or the magnitude and direction of a translation vector. If $\theta = 0$ is trivial transformation, represented on \mathcal{H} by the identity operator, then for infinitesimal transformation we have

$$U(\delta\theta) = 1 - i\delta\theta T + O(\delta\theta^2) \quad (4.8)$$

where T is some operator that is independent of θ . (The factor of $-i$ in this equation is just a convention for later convenience.) T is called the *generator* of the transformation. For such infinitesimal transformations, the condition (4.6) that U is unitary becomes

$$1 + i\delta\theta T^\dagger + O(\delta\theta^2) = 1 + i\delta\theta T + O(\delta\theta^2), \quad (4.9)$$

¹⁷Note that unitary operators are certainly bounded, and in fact have unit norm in the operator topology.

¹⁸This statement is not quite accurate – there's an important refinement that we'll return to later in the course.

which to first order in $\delta\theta$ gives

$$T = T^\dagger. \quad (4.10)$$

Thus the generator T is Hermitian and hence a good candidate for an observable quantity. A finite transformation can be generated by repeatedly performing an infinitesimal one. Specifically, if we set $\delta\theta = \theta/N$ and transform N times with $U(\delta\theta)$, then in the limit $N \rightarrow \infty$ we have

$$U(\theta) = \lim_{N \rightarrow \infty} \left(1 - i \frac{\theta}{N} T \right)^N = e^{-i\theta T}, \quad (4.11)$$

where the exponential of an operator may be defined by (2.42), or equivalently by its power series expansion. This form of the unitary operator $U(\theta)$ is especially useful when states are expressed in terms of the basis of eigenstates of the Hermitian generator T .

We obtain an important equation by using (4.8) to evaluate $|\psi'\rangle = U(\delta\theta)|\psi\rangle$. Subtracting $|\psi\rangle$ from both sides and dividing through by $\delta\theta$, in the limit $\delta\theta \rightarrow 0$ we obtain

$$i \frac{\partial |\psi\rangle}{\partial \theta} = T |\psi\rangle \quad (4.12)$$

There is no assumption here that the wavefunction $\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle$ should be differentiable as a function of space. We merely say that as our transformation varies smoothly away from the identity, so too does the state $|\psi\rangle$ vary smoothly inside \mathcal{H} . The rate at which $|\psi\rangle$ varies is governed by the generator.

Having described how transformations act on states in \mathcal{H} , we should now understand how they act on operators. Suppose A is some operator whose matrix elements we are interested in. If our transformation maps $|\psi\rangle \rightarrow |\psi'\rangle$, then the expectation value of A will be mapped as

$$\langle \psi | A | \psi \rangle \mapsto \langle \psi' | A | \psi' \rangle = \langle \psi | U^\dagger(\theta) A U(\theta) | \psi \rangle. \quad (4.13)$$

Consequently, we can find the expectation value A will take after the transformation by working with the original states, but instead transforming the operator as

$$A \mapsto A' := U^\dagger(\theta) A U(\theta) = U^{-1}(\theta) A U(\theta), \quad (4.14)$$

using the fact that U is unitary. This is known as a *similarity transform*. Note that

$$[A', B'] = [U^{-1} A U, U^{-1} B U] = U^{-1} [A, B] U, \quad (4.15)$$

so the commutator of two similarity transforms is the similarity transform of the commutator. Also, similarity transforms do not change the spectrum of any operator: if $|\psi\rangle$ is an eigenstate of A with eigenvalue a , then $U^{-1}|\psi\rangle$ is an eigenstate of A' with the same eigenvalue.

Finally, for an infinitesimal transformation, we have

$$U^\dagger(\delta\theta) A U(\delta\theta) = A + i \delta\theta [T, A] + O(\delta\theta^2), \quad (4.16)$$

so $\delta A = i \delta\theta [T, A]$. Thus, while the rate of change of states themselves is given by the action of the generator, the rate of change of operators under a transformation is determined by

the *commutator* of the generator with the operator. We'll see that most of the commutation relations we meet in quantum mechanics (and certainly all the familiar ones) arise this way. Furthermore, because Hermitian operators represent observable quantities with which we are familiar, in practice it's often easier to understand how a transformation should act on an operator, rather than on the more abstract notion of a state in Hilbert space.

The above discussion has been rather abstract. Let's ground it by looking at a few of the most important examples of transformations and their corresponding generators.

4.2 Translations

When we move an object around, we expect to find it in a new place. Specifically, suppose $\langle \psi | \mathbf{X} | \psi \rangle = \mathbf{x}_0$ for some normalised state $|\psi\rangle$. Since \mathbf{x}_0 just labels a spatial point, it must behave under translations and rotations like any vector. Thus, after the translation we expect our object to be described by a new state $|\psi'\rangle$ with $\langle \psi' | \mathbf{X} | \psi' \rangle = \mathbf{x}_0 + \mathbf{a}$. The general theory of the previous section asserts that this translation is represented on Hilbert space by a unitary operator¹⁹ $U(\mathbf{a})$ so that $|\psi'\rangle = U(\mathbf{a})|\psi\rangle$. Therefore we can write the expected location of the translated state as

$$\begin{aligned} \langle \psi' | \mathbf{X} | \psi' \rangle &= \langle \psi | \mathbf{X} | \psi \rangle + \mathbf{a} = \langle \psi | (\mathbf{X} + 1_{\mathcal{H}} \mathbf{a}) | \psi \rangle \\ &= \langle \psi | U^{-1}(\mathbf{a}) \mathbf{X} U(\mathbf{a}) | \psi \rangle, \end{aligned} \quad (4.17)$$

where $1_{\mathcal{H}}$ represents the identity operator on \mathcal{H} . (We'll often drop the identity operator where it's unambiguous.) Since this is true for any state $|\psi\rangle$, the same argument as above shows that

$$U^{-1}(\mathbf{a}) \mathbf{X} U(\mathbf{a}) = \mathbf{X} + \mathbf{a} 1_{\mathcal{H}}. \quad (4.18)$$

Let's be clear about what this equation means. In this course, we're interested in Quantum Mechanics on \mathbb{R}^3 , so the (boldface) position operator $\mathbf{X} : \mathcal{H} \rightarrow \mathcal{H}$ is really a collection of three operators corresponding to the three coordinates of \mathbb{R}^3 . Taking the standard Cartesian basis of \mathbb{R}^3 , in more detail (4.18) says

$$\begin{pmatrix} U^{-1}(\mathbf{a}) X U(\mathbf{a}) \\ U^{-1}(\mathbf{a}) Y U(\mathbf{a}) \\ U^{-1}(\mathbf{a}) Z U(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} X + a_x 1_{\mathcal{H}} \\ Y + a_y 1_{\mathcal{H}} \\ Z + a_z 1_{\mathcal{H}} \end{pmatrix}. \quad (4.19)$$

In other words, conjugating the X (component!) operator by $U(\mathbf{a})$ gives a new operator $U^{-1}(\mathbf{a}) X U(\mathbf{a}) : \mathcal{H} \rightarrow \mathcal{H}$ that we identify with $X + a_x$, and similarly for Y and Z . If $U(\mathbf{a})$ acted on the position operator in any other way, it could not represent a translation, so equation (4.18) may be taken as the defining property of the translation operator, distinguishing it from other unitary transformations.

¹⁹Strictly, we should label this as $U_{\text{trans}}(\mathbf{a})$ to emphasise that it is the unitary operator generating *translations*, as opposed to anything else we might associate to \mathbf{a} . We typically drop such additional labels. Exactly which unitary operator we're talking about should hopefully be clear from the context.

For an infinitesimal translation $\delta\mathbf{a}$, equation (4.8) states that the translation operator can be written as²⁰

$$U(\delta\mathbf{a}) = 1 - \frac{i}{\hbar} \delta\mathbf{a} \cdot \mathbf{P} + O(\delta\mathbf{a}^2) \quad (4.20)$$

where \mathbf{P} is Hermitian, I've named this operator \mathbf{P} , so I'm not going to be able to fool any of you that it will turn out to be anything other than the momentum operator. But that's not how I want you to think of it yet. *By definition*, \mathbf{P}/\hbar is the generator of infinitesimal translations.

The translation generator \mathbf{P}/\hbar must have units 1/(length) so that it makes sense to add the second term of (4.20) to the first. Because we've included a factor of \hbar , our \mathbf{P} must have dimensions of (mass \times velocity). In fact, this is just a convention. We could choose to measure masses in units of 1/(length \times velocity), using \hbar as a conversion factor for our units. In relativity, the speed of light c provides a further natural conversion factor between lengths and times, so in units of \hbar/c , masses are equivalent to inverse lengths. In *natural units*, we work with length and time scales adapted to Nature rather than our civilisation, so we set $\hbar = 1$ and $c = 1$. In these units, the translation generator is precisely \mathbf{P} . Incidentally, you may be worried about setting $\hbar = 1$ when it is such a 'small' number; $\hbar \approx 1.0547 \times 10^{-34}$ Js/rad. But this is quite wrong. Every atom in the Universe knows that \hbar is not at all small or insignificant. It only appears small when measured in units such as Joules and seconds that are relevant for steam engines and pocketwatches. The real, deep question is not why \hbar appears small, but why humans are so big.

Plugging (4.20) into the defining equation (4.18) for the translation operator we find

$$\frac{i}{\hbar} [\delta\mathbf{a} \cdot \mathbf{P}, \mathbf{X}] = \delta\mathbf{a} \quad (4.21)$$

and since this holds for any infinitesimal translation,

$$[X_i, P_j] = i\hbar \delta_{ij}. \quad (4.22)$$

These, as I'm sure you recognise, are the fundamental commutation relations of quantum mechanics. I'm prepared to bet that the first time you met them (and perhaps even up until now) you thought they were a weird and mysterious feature of quantum mechanics. Here they're revealed as little more than common sense: asking where you are and then translating is not the same as first translating and then asking where you are. What *is* weird and quantum about this equation is *not* that \mathbf{X} and \mathbf{P} don't commute, but the fact that they are operators acting on a Hilbert space.

By repeatedly performing the same infinitesimal translation, we can write

$$U(\mathbf{a}) = \exp\left(-\frac{i}{\hbar} \mathbf{a} \cdot \mathbf{P}\right) \quad (4.23)$$

for a translation through the finite vector \mathbf{a} . Since $U(\mathbf{a})$ provides a homomorphism of this group of translations to the group of linear operators on \mathcal{H} , we have

$$U(\mathbf{b})U(\mathbf{a}) = U(\mathbf{b} + \mathbf{a}) = U(\mathbf{a} + \mathbf{b}) = U(\mathbf{a})U(\mathbf{b}), \quad (4.24)$$

²⁰Here, $\delta\mathbf{a} \cdot \mathbf{P} = \sum_{i=1}^3 \delta a_i P_i$ involves the standard dot product in \mathbb{R}^3 . Note that the result is still an operator on \mathcal{H} , a linear combination of the three operators \mathbf{P} .

where in the second equality we used the fact that $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$, stating that the order in which we perform two translations doesn't matter²¹. Using (4.20) to expand the far left & far right of this equation to lowest non-trivial order in $a_i b_j$ shows that the generators \mathbf{P} obey

$$[P_i, P_j] = 0 \quad \text{for all } i, j \quad (4.25)$$

and so commute with each other. Again, the fundamental meaning of this equation is simply that translations form an Abelian group — the order of translations doesn't matter. It only says anything about our ability to simultaneously know all three components of a particle's momentum once we identify \mathbf{P} as corresponding to momentum, which we haven't yet done.

Now let's think about what the translation operator does to states. Let $|\mathbf{x}\rangle$ represent an eigenstate of the position operator with eigenvalue \mathbf{x} , so that $\mathbf{X}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle$, with normalisation condition

$$\langle \mathbf{x}' | \mathbf{x} \rangle = \delta^3(\mathbf{x}' - \mathbf{x}) \quad (4.26)$$

as usual for continuum states²². This state represents a particle which is definitely located at $\mathbf{x} \in \mathbb{R}^3$. Applying the translation operator,

$$\mathbf{X}(U(\mathbf{a})|\mathbf{x}\rangle) = ([\mathbf{X}, U(\mathbf{a})] + U(\mathbf{a})\mathbf{X})|\mathbf{x}\rangle. \quad (4.27)$$

We can evaluate the commutator by multiplying (4.18) through by $U(\mathbf{a})$ on the left, finding $[\mathbf{X}, U(\mathbf{a})] = U(\mathbf{a})\mathbf{a}$. Since \mathbf{a} is just a constant vector (not an operator on \mathcal{H}) it commutes with the translation operator, so

$$\mathbf{X}(U(\mathbf{a})|\mathbf{x}\rangle) = (\mathbf{x} + \mathbf{a})(U(\mathbf{a})|\mathbf{x}\rangle). \quad (4.28)$$

As anticipated, our new state $U(\mathbf{a})|\mathbf{x}\rangle$ is certainly located at $\mathbf{x} + \mathbf{a}$. Consequently, $U(\mathbf{a})|\mathbf{x}\rangle$ must be proportional to the state $|\mathbf{x} + \mathbf{a}\rangle$, the normalised state that is definitely located at $\mathbf{x} + \mathbf{a}$. Setting $U(\mathbf{a})|\mathbf{x}\rangle = c|\mathbf{x} + \mathbf{a}\rangle$ for some $c \in \mathbb{C}$ and taking the inner product with another position eigenstate $|\mathbf{x}'\rangle$ we have²³

$$\begin{aligned} c\delta^3(\mathbf{x}' - \mathbf{x} - \mathbf{a}) &= c\langle \mathbf{x}' | \mathbf{x} + \mathbf{a} \rangle = \langle \mathbf{x}' | U(\mathbf{a})|\mathbf{x}\rangle = (U^{-1}(\mathbf{a})|\mathbf{x}'\rangle)^\dagger |\mathbf{x}\rangle \\ &= \left(\frac{1}{c}|\mathbf{x}' - \mathbf{a}\rangle\right)^\dagger |\mathbf{x}\rangle = \frac{1}{c}\delta^3(\mathbf{x}' - \mathbf{a} - \mathbf{x}) \end{aligned} \quad (4.29)$$

²¹Here this is of course the statement that the group of spatial translations is $(\mathbb{R}^3, +)$, where the group operation '+' is commutative. Groups for which the group operation is commutative are often called *Abelian*.

²²Technically, $|\mathbf{x}\rangle$ is a non-normalizable state, which really means it doesn't live in the Hilbert space $L^2(\mathbb{R}^3, d^3x)$. This is related to the fact that the position operator \mathbf{X} is *unbounded* — there is no constant M such that $\|\mathbf{X}|\psi\rangle\| \leq M\|\psi\rangle\|$ for all $|\psi\rangle \in \mathcal{H}$. Naïvely, this means that 'the eigenvalues of \mathbf{X} can be arbitrarily large', which is physically reasonable since our particle may be located very far away. However, in general, unbounded operators do not have any eigenstates in \mathcal{H} , so do not really have eigenvalues. The distinction between bounded and unbounded operators is tremendously important in functional analysis, but will not play a role in this course.

²³Note the slight awkwardness of the Dirac notation here. Because operators act to the right, we have to write $\langle \chi | U | \phi \rangle = (U^{-1}|\chi\rangle)^\dagger |\phi\rangle$ for any unitary operator U , as we used in the third equality. In mathematical notation we would simply write $(\chi, U\phi) = (U^{-1}\chi, \phi)$.

where we used the fact that $U(\mathbf{a})$ is unitary and the fact that if $U(\mathbf{a})|\mathbf{x}\rangle = c|\mathbf{x} + \mathbf{a}\rangle$ then $U^{-1}(\mathbf{a})|\mathbf{x} + \mathbf{a}\rangle = c^{-1}|\mathbf{x}\rangle$. Comparing both sides shows that $|c|^2 = 1$, so c is a pure phase and without loss of generality we can set $c = 1$.

Now suppose we consider translating an arbitrary state $|\psi\rangle$. The position space wavefunction of this state is simply its coefficient

$$\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle \quad (4.30)$$

in the position basis. Thus the wavefunction $\psi_{\text{trans}}(\mathbf{x})$ of the translated state $|\psi_{\text{trans}}\rangle = U(\mathbf{a})|\psi\rangle$ is

$$\begin{aligned} \psi_{\text{trans}}(\mathbf{x}) &= \langle \mathbf{x} | \psi_{\text{trans}} \rangle = \langle \mathbf{x} | U(\mathbf{a}) | \psi \rangle = (U^{-1}(\mathbf{a}) | \mathbf{x} \rangle)^\dagger | \psi \rangle \\ &= \langle \mathbf{x} - \mathbf{a} | \psi \rangle = \psi(\mathbf{x} - \mathbf{a}). \end{aligned} \quad (4.31)$$

Eq. (4.31) says that the wavefunction of the translated state takes the same value at \mathbf{x} as the original wavefunction took at $\mathbf{x} - \mathbf{a}$. This is completely natural as we've translated our state through \mathbf{a} !

In particular, for an infinitesimal translation $\delta\mathbf{a}$, on the one hand we can Taylor expand the translated wavefunction to find²⁴

$$\psi_{\text{trans}}(\mathbf{x}) - \psi(\mathbf{x}) = -\delta\mathbf{a} \cdot \nabla \psi(\mathbf{x}), \quad (4.32)$$

while on the other we expand the operator $U(\delta\mathbf{a})$ to find

$$\psi_{\text{trans}}(\mathbf{x}) - \psi(\mathbf{x}) = \left\langle \mathbf{x} \left| 1 - \frac{i}{\hbar} \delta\mathbf{a} \cdot \mathbf{P} \right| \psi \right\rangle - \langle \mathbf{x} | \psi \rangle = -\frac{i}{\hbar} \delta\mathbf{a} \cdot \langle \mathbf{x} | \mathbf{P} | \psi \rangle \quad (4.33)$$

using (4.20) to lowest order in $\delta\mathbf{a}$. Consequently, for any state $|\psi\rangle$ the momentum operator \mathbf{P} acts in the position representation as

$$\langle \mathbf{x} | \mathbf{P} | \psi \rangle = -i\hbar \nabla \psi(\mathbf{x}). \quad (4.34)$$

This is just what you would have said in IB QM last year, here derived from the point of view of the effect a translation of \mathbb{R}^3 has on states in Hilbert space.

As a special case, let's apply this argument to the state $|\mathbf{p}\rangle$ that obeys²⁵ $\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle$, representing a particle whose momentum is certainly \mathbf{p} . In this case the above argument becomes

$$\begin{aligned} \psi_{\mathbf{p}}(\mathbf{x} - \mathbf{a}) &= \langle \mathbf{x} - \mathbf{a} | \mathbf{p} \rangle = \langle \mathbf{x} | U(\mathbf{a}) | \mathbf{p} \rangle = \langle \mathbf{x} | e^{-i\mathbf{a} \cdot \mathbf{P} / \hbar} | \mathbf{p} \rangle \\ &= e^{-i\mathbf{a} \cdot \mathbf{p} / \hbar} \langle \mathbf{x} | \mathbf{p} \rangle = e^{-i\mathbf{a} \cdot \mathbf{p} / \hbar} \psi_{\mathbf{p}}(\mathbf{x}), \end{aligned} \quad (4.35)$$

where in going to the second line we used the fact that $|\mathbf{p}\rangle$ is an eigenstate of \mathbf{P} . Comparing the first and last expression, we deduce that the position space wavefunction of a momentum eigenstate must take the form

$$\psi_{\mathbf{p}}(\mathbf{x}) = C e^{i\mathbf{p} \cdot \mathbf{x} / \hbar} \quad (4.36)$$

²⁴At least if the wavefunction is sufficiently smooth.

²⁵As with the position operator, the momentum operator \mathbf{P} is unbounded, so the 'momentum eigenstates' $|\mathbf{p}\rangle$ are necessarily non-normalizable, as we'll soon see explicitly.



Figure 4: An infinitesimal rotation transforms $\mathbf{v} \mapsto \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}$.

for some \mathbf{x} -independent factor C . It's convenient to choose C to be $(2\pi\hbar)^{-3/2}$ since in this case we have

$$\begin{aligned} \langle \mathbf{p}' | \mathbf{p} \rangle &= \int d^3y d^3x \langle \mathbf{p}' | \mathbf{y} \rangle \langle \mathbf{y} | \mathbf{x} \rangle \langle \mathbf{x} | \mathbf{p} \rangle = \int \frac{d^3x d^3y}{(2\pi\hbar)^3} e^{-i\mathbf{p}' \cdot \mathbf{y} / \hbar} \delta^3(\mathbf{y} - \mathbf{x}) e^{i\mathbf{p} \cdot \mathbf{x} / \hbar} \\ &= \int \frac{d^3x}{(2\pi\hbar)^3} e^{-i(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{x} / \hbar} = \delta^3(\mathbf{p}' - \mathbf{p}) \end{aligned} \quad (4.37)$$

so that the momentum eigenstates are normalised in the same way as the position eigenstates (and in the same way as for any continuum states). Again, the form

$$\psi_{\mathbf{p}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p} \cdot \mathbf{x} / \hbar} \quad (4.38)$$

of the wavefunction for momentum eigenstates is familiar, but here we've derived it from first principles rather than using the position space representation (4.34) of the momentum operator. (We assumed this result earlier when showing that the position and momentum space wavefunctions are each other's Fourier transforms.) Note also that $|\psi_{\mathbf{p}}(\mathbf{x})|^2 = 1/(2\pi\hbar)^3$, so $\int_{\mathbb{R}^3} |\psi_{\mathbf{p}}(\mathbf{x})|^2 d^3x$ diverges. Like position eigenstates, momentum eigenstates are also non-normalizable.

4.3 Rotations

We now consider the effect rotating \mathbb{R}^3 has on a quantum state $|\psi\rangle \in \mathcal{H}$. For a vector $\mathbf{v} \in \mathbb{R}^3$ a rotation anticlockwise around the axis $\hat{\boldsymbol{\alpha}}$ by an amount $|\boldsymbol{\alpha}|$ is a linear transformation

$$\mathbf{R}(\boldsymbol{\alpha}) : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad \mathbf{R}(\boldsymbol{\alpha}) : \mathbf{v} \mapsto \mathbf{v}' = \mathbf{R}(\boldsymbol{\alpha})\mathbf{v} \quad (4.39)$$

that obeys

$$\mathbf{v}' \cdot \mathbf{v}' = \mathbf{v} \cdot \mathbf{v} \quad \text{and} \quad \det(\mathbf{R}(\boldsymbol{\alpha})) = +1. \quad (4.40)$$

The first of these conditions says that rotations preserve lengths, and implies that $\mathbf{R}(\boldsymbol{\alpha})$ must be an *orthogonal* transformation. The second ensure that $\mathbf{R}(\boldsymbol{\alpha})$ preserves the orientation. For infinitesimal rotations, figure 4 shows that

$$\mathbf{v}' = \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v} + O(\delta\boldsymbol{\alpha}^2). \quad (4.41)$$

(Note that this obeys (4.40).) Given two orthogonal transformations $\mathbf{R}(\boldsymbol{\alpha})$ and $\mathbf{R}(\boldsymbol{\beta})$ the composite $\mathbf{R}(\boldsymbol{\beta})\mathbf{R}(\boldsymbol{\alpha})$ is again an orthogonal transformation with unit determinant. Spatial rotations form a group, known as $\text{SO}(3)$, with the identity being the trivial rotation and the inverse of $\mathbf{R}(\boldsymbol{\alpha})$ being a rotation of the same amount around the same axis $\boldsymbol{\alpha}$, but now clockwise. However, in general $\mathbf{R}(\boldsymbol{\beta})\mathbf{R}(\boldsymbol{\alpha}) \neq \mathbf{R}(\boldsymbol{\alpha})\mathbf{R}(\boldsymbol{\beta})$, so the order in which we apply rotations is important and the rotation group is *non-Abelian*.

As for any group of transformations, in quantum mechanics the group of rotations is represented on \mathcal{H} by unitary operators. We denote the unitary operator corresponding to $\mathbf{R}(\boldsymbol{\alpha})$ as $U(\boldsymbol{\alpha})$. $U(\boldsymbol{\alpha})$ acts on the position operator as

$$U^{-1}(\boldsymbol{\alpha})\mathbf{X}U(\boldsymbol{\alpha}) = \mathbf{R}(\boldsymbol{\alpha})\mathbf{X} \quad (4.42)$$

where the *lhs* involves composition of operators in \mathcal{H} , while the *rhs* here is the usual action of a rotation on the three components of \mathbf{X} . For example, if $\boldsymbol{\alpha} = \alpha\hat{\mathbf{z}}$ so that the rotation is around the z -axis in \mathbb{R}^3 , then in more detail (4.42) says

$$\begin{pmatrix} U^{-1}(\boldsymbol{\alpha})XU(\boldsymbol{\alpha}) \\ U^{-1}(\boldsymbol{\alpha})YU(\boldsymbol{\alpha}) \\ U^{-1}(\boldsymbol{\alpha})ZU(\boldsymbol{\alpha}) \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (4.43)$$

Thus, conjugating the operator X by this rotation operator changes it to a new operator $U^{-1}(\boldsymbol{\alpha})XU(\boldsymbol{\alpha})$ on Hilbert space, which we identify with the usual linear combination $X\cos\alpha - Y\sin\alpha$. Again, $U(\boldsymbol{\alpha})$ must act on \mathbf{X} this way if it is indeed to correspond to a rotation.

For an infinitesimal transformation, following (4.8) we can write

$$U(\delta\boldsymbol{\alpha}) = 1 - \frac{i}{\hbar}\delta\boldsymbol{\alpha} \cdot \mathbf{J} + O(\delta\boldsymbol{\alpha}^2) \quad (4.44)$$

for some Hermitian generators \mathbf{J}/\hbar . Since angles are dimensionless, \mathbf{J} , like \hbar , must have dimensions of (length \times momentum). Later we will see that \mathbf{J} corresponds to the angular momentum operator, though *by definition* we have that \mathbf{J}/\hbar is the generator of rotations. Using this and (4.41) in (4.42) shows that we have

$$\frac{i}{\hbar}[\delta\boldsymbol{\alpha} \cdot \mathbf{J}, \mathbf{X}] = \delta\boldsymbol{\alpha} \times \mathbf{X}, \quad (4.45)$$

and since this is true for any axis of rotation, we have the commutation relations

$$[J_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k. \quad (4.46)$$

These relations are nothing more than the statement that the operator \mathbf{X} transforms as a vector under rotations. They are the infinitesimal version of (4.42) – they *must* hold if \mathbf{J} is indeed to generate rotations.

Just as for translations, the mutual commutation relations among the components of \mathbf{J} themselves follow from the fact that the $U(\boldsymbol{\alpha})$ s provide a homomorphism from the group²⁶

²⁶Again, later we'll see that this statement needs to be slightly refined.

$SO(3)$ of rotations to the space of unitary operators on \mathcal{H} . However, the non-Abelian nature of the rotation group means we should expect these commutators to be non-trivial in general. Performing two infinitesimal rotations through angles $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\beta}$, for any vector \mathbf{v} we have

$$\begin{aligned}\mathbf{R}(\delta\boldsymbol{\beta})\mathbf{R}(\delta\boldsymbol{\alpha})\mathbf{v} &= \mathbf{R}(\delta\boldsymbol{\beta})(\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) = (\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) + \delta\boldsymbol{\beta} \times (\mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v}) \\ &= \mathbf{v} + \delta\boldsymbol{\alpha} \times \mathbf{v} + \delta\boldsymbol{\beta} \times \mathbf{v} + \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v})\end{aligned}\quad (4.47)$$

to lowest non-trivial order in $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\beta}$. Consequently, the difference between these rotations and the two rotations performed in the opposite order is

$$\begin{aligned}[\mathbf{R}(\delta\boldsymbol{\beta})\mathbf{R}(\delta\boldsymbol{\alpha}) - \mathbf{R}(\delta\boldsymbol{\alpha})\mathbf{R}(\delta\boldsymbol{\beta})]\mathbf{v} &= \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v}) - \delta\boldsymbol{\alpha} \times (\delta\boldsymbol{\beta} \times \mathbf{v}) \\ &= (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \times \mathbf{v} \\ &= \mathbf{R}(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha})\mathbf{v} - \mathbf{v}\end{aligned}\quad (4.48)$$

using standard properties of the vector triple product. The *rhs* again involves a rotation, through the angle $\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}$. Applying the homomorphism U we obtain

$$[U(\delta\boldsymbol{\beta}), U(\delta\boldsymbol{\alpha})] = U(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) - I \quad (4.49)$$

for our operators in Hilbert space. Using (4.44), this is

$$-\frac{1}{\hbar^2}[\delta\boldsymbol{\beta} \cdot \mathbf{J}, \delta\boldsymbol{\alpha} \cdot \mathbf{J}] = -\frac{i}{\hbar}(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \cdot \mathbf{J} \quad (4.50)$$

and since this must hold for arbitrary successive infinitesimal rotations, we have finally

$$[J_i, J_j] = i\hbar \sum_k \epsilon_{ijk} J_k \quad (4.51)$$

as the commutation relations among the rotation generators.

Again, there's nothing 'weird' or 'quantum' about (4.51), beyond the fact that they involve operators on Hilbert space. In particular, their form just reflects the fact that the order of rotations around different axes matters, and that the difference between the two orderings is itself a rotation around the axis perpendicular to the original two. Compared to the relations $[P_i, P_j] = 0$, the non-triviality of the commutation relations (4.51) arises purely because $SO(3)$ is a non-Abelian group, unlike the group of translations. These non-trivial commutation relations do not prevent us from exponentiating (4.44) to write $U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar}$ for a finite rotation around a fixed axis $\boldsymbol{\alpha}$, because the exponentiation always involved the same component $\hat{\boldsymbol{\alpha}} \cdot \mathbf{J}$ of the rotation generator, which certainly commutes with itself.

The combined group of translations *and* rotations of Euclidean \mathbb{R}^3 is sometimes known as $E(3)$, or else $ISO(3)$. For translations through a_1 and a_2 , and rotations R_1 and R_2 , $E(3)$ has the group composition law

$$(a_2, R_2) \cdot (a_1, R_1) = (a_2 + R_2 a_1, R_2 R_1), \quad (4.52)$$

where we note that the second rotation also acts on the first translation. Clearly, both \mathbb{R}^3 and $SO(3)$ are subgroups of $E(3)$, but the fact that rotations act non-trivially on previous translations means that $E(3)$ is the *semi-direct* product $E(3) \cong \mathbb{R}^3 \rtimes SO(3)$. This group law shows that rotations and translations do not commute: translating any vector \mathbf{v} through \mathbf{a} and then rotating around $\hat{\alpha}$ gives us $\mathbf{R}(\alpha)(\mathbf{v} + \mathbf{a})$, whereas first rotating then translating gives $\mathbf{a} + \mathbf{R}(\alpha)\mathbf{v}$ instead. In particular, for infinitesimal translations and rotations we have

$$\mathbf{R}(\delta\alpha)(\mathbf{v} + \delta\mathbf{a}) - (\mathbf{R}(\delta\alpha)\mathbf{v} + \delta\mathbf{a}) = \delta\alpha \times \delta\mathbf{a} \quad (4.53)$$

independent of \mathbf{v} . Applying the homomorphism U , in Hilbert space we obtain commutation relations²⁷

$$[U_R(\delta\alpha), U_T(\delta\mathbf{a})] = U_T(\delta\alpha \times \delta\mathbf{a}) - I \quad (4.54)$$

or equivalently

$$[J_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k \quad (4.55)$$

in terms of the rotation and translation generators.

More generally, any 3-component linear operator $\mathbf{V} : \mathcal{H} \rightarrow \mathcal{H}$ is said to *transform as a vector under rotations* if it obeys $U^{-1}(\alpha)\mathbf{V}U(\alpha) = \mathbf{R}(\alpha)\mathbf{V}$ for any α . (Depending on its behaviour under parity transformations, discussed in section 4.6, such a \mathbf{V} could either be a vector or pseudovector operator.) Just as above, for infinitesimal rotations this implies the commutation relations

$$[J_i, V_j] = i\hbar \sum_k \epsilon_{ijk} V_k \quad (4.56)$$

among the components of \mathbf{V} and \mathbf{J} . We see that \mathbf{X} , \mathbf{P} and \mathbf{J} itself²⁸ each transform as vectors under rotations. This of course is just what we'd expect for position, momentum and angular momentum in classical mechanics.

On the other hand, if an operator S obeys

$$U^{-1}(\alpha) S U(\alpha) = S \quad (4.57)$$

for any α , so that it is unchanged by the rotation operator, we say S is a *scalar operator*²⁹. The corresponding infinitesimal version is

$$[\mathbf{J}, S] = 0. \quad (4.58)$$

Just as we can form a scalar by taking the Euclidean inner product $\mathbf{v} \cdot \mathbf{w}$ of the two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, so we can form scalar operators from the Euclidean inner product of two vector

²⁷Here, for the sake of clarity, we have labelled the translations operator by U_T and rotation operator by U_R . In particular, the operator on the *rhs* of (4.54) is a translation.

²⁸This is always true of \mathbf{X} and \mathbf{P} , but turns out to be a three-dimensional coincidence for \mathbf{J} : the rotation group $SO(d)$ of \mathbb{R}^d has dimension $d(d-1)/2$ and the generators are generically represented by antisymmetric matrices \mathcal{J}_{ij} . When $d=3$, a 3×3 antisymmetric matrix has 3 independent components, so we can equivalently package the \mathcal{J}_{ij} as vectors through $J_i = \epsilon_{ijk} \mathcal{J}_{jk}$.

²⁹More precisely, we say that S transforms as a scalar under rotations — depending on its behaviour under parity transformations it could be either a scalar or pseudoscalar operator.

operators. Indeed, if $U^{-1}(\boldsymbol{\alpha})\mathbf{V}U(\boldsymbol{\alpha}) = \mathbf{R}(\boldsymbol{\alpha})\mathbf{V}$ and similarly for \mathbf{W} , then

$$\begin{aligned} U^{-1}(\boldsymbol{\alpha})(\mathbf{V} \cdot \mathbf{W})U(\boldsymbol{\alpha}) &= (U^{-1}(\boldsymbol{\alpha})\mathbf{V}U(\boldsymbol{\alpha})) \cdot (U^{-1}(\boldsymbol{\alpha})\mathbf{W}U(\boldsymbol{\alpha})) \\ &= (\mathbf{R}(\boldsymbol{\alpha})\mathbf{V}) \cdot (\mathbf{R}(\boldsymbol{\alpha})\mathbf{W}) \\ &= \mathbf{V} \cdot \mathbf{W} \end{aligned} \quad (4.59)$$

by the standard rotational invariance of the dot product. As always, we can write this in terms of commutators by considering infinitesimal rotations:

$$\begin{aligned} \left[J_i, \sum_j V_j W_j \right] &= \sum_j [J_i, V_j] W_j + \sum_j V_j [J_i, W_j] \\ &= i\hbar \sum_{jk} \epsilon_{ijk} V_k W_j + i\hbar \sum_{jk} V_j \epsilon_{ijk} W_k \\ &= i\hbar \sum_{jk} \epsilon_{ijk} (-V_j W_k + V_j W_k) = 0, \end{aligned} \quad (4.60)$$

where in going to the final line we relabelled dummy indices $j \leftrightarrow k$ in the first term and used the antisymmetry of ϵ_{ijk} . Note that our calculations always preserved the order of \mathbf{V} and \mathbf{W} — our result holds irrespective of whether \mathbf{V} and \mathbf{W} commute.

An important special case of (4.58) is to take $S = \mathbf{J}^2 = \mathbf{J} \cdot \mathbf{J}$. The fact that the rotation generators obey the commutation relations

$$[J_i, J_j] = i\hbar \sum_k \epsilon_{ijk} J_k \quad \text{and} \quad [J_i, \mathbf{J}^2] = 0 \quad (4.61)$$

means that we cannot find a complete set of simultaneous eigenstates of all of the J_i s, but we *can* find a complete set of simultaneous eigenstates of \mathbf{J}^2 and any one component of \mathbf{J} . Looking ahead to our interpretation of \mathbf{J} as the angular momentum operator, Born's 2nd postulate of QM tells us that we cannot say a particle has a definite angular momentum vector, but we can know the magnitude of the angular momentum and the amount aligned along any given axis.

4.3.1 Translations Around a Circle

In 1B QM, you defined the orbital angular momentum operator $\mathbf{L} = \mathbf{X} \times \mathbf{P}$ which also has the commutation relations (4.56) with itself and with \mathbf{X} and \mathbf{P} . We'll understand the relation between \mathbf{J} and \mathbf{L} later, but it's important to understand that in general $\mathbf{J} \neq \mathbf{L}$. We can understand \mathbf{L} from the present perspective as follows.

When a system is displaced through the vector \mathbf{a} , its state is transformed by the unitary operator $U(\mathbf{a}) = e^{-i\mathbf{a} \cdot \mathbf{P}/\hbar}$. We now imagine successively performing n translations successively through the set of vectors $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$. Each translation is represented by a unitary operator $U(\mathbf{a}_i)$, so the final state will be

$$U(\mathbf{a}_n) \cdots U(\mathbf{a}_2)U(\mathbf{a}_1)|\psi\rangle = U(\mathbf{b})|\psi\rangle \quad (4.62)$$

where $\mathbf{b} = \mathbf{a}_1 + \cdots + \mathbf{a}_n$ is the total displacement vector. Since the net result depends only on the total translation \mathbf{b} , the change in $|\psi\rangle$ is independent of the particular path that the

system takes. In particular, if the path is closed so that $\mathbf{a}_1 + \dots + \mathbf{a}_n = \mathbf{0}$, then $|\psi\rangle$ is unchanged.

Now consider the effect of translating the system around an arc of a circle centred on the origin. We can approximate this circle by an N -sided regular polygon, with the approximation improving as $N \rightarrow \infty$, so we move around a circle by applying a succession of small translations, each in a slightly different direction. Specifically, if the system initially lies at some location \mathbf{x} , making an angle α with some axis, then when move it in the plane normal to \mathbf{n} to lie at an angle $\alpha + \delta\alpha$, we translate it through $\delta\mathbf{a} = \delta\alpha \mathbf{n} \times \mathbf{x}$. Thus the associated unitary translation operator obeys

$$U^{-1}(\delta\mathbf{a}) \mathbf{X} U(\delta\mathbf{a}) = \mathbf{X} + \delta\alpha (\mathbf{n} \times \mathbf{X}) \quad (4.63)$$

and so can be written as

$$U(\delta\mathbf{a}) = 1 - \frac{i}{\hbar} \delta\alpha (\mathbf{n} \times \mathbf{X}) \cdot \mathbf{P} + O(\delta\alpha^2) = 1 - \frac{i}{\hbar} \delta\alpha \mathbf{n} \cdot \mathbf{L} + O(\delta\alpha^2), \quad (4.64)$$

where

$$\mathbf{L}/\hbar = \mathbf{X} \times \mathbf{P}/\hbar \quad (4.65)$$

is Hermitian operator. We now see that \mathbf{L} is the *generator of circular transformations*. (4.64) contains only one operator, $\mathbf{n} \cdot \mathbf{L}$, so it inevitably commutes with itself and we can exponentiate to find $U(\mathbf{a}_{\text{circ}}) = e^{-i\alpha \mathbf{n} \cdot \mathbf{L}/\hbar}$ for finite translations around our circle.

4.3.2 Spin

The commutation relations $[L_i, X_j]$, $[L_i, P_j]$ and $[L_i, L_j]$ of the composite operator \mathbf{L} all follow from the more primitive commutation relations $[X_i, P_j]$, $[X_i, X_j]$ and $[P_i, P_j]$. If we're only concerned with these operators, then $U_{\text{circ}} = e^{-i\alpha \mathbf{n} \cdot \mathbf{L}/\hbar}$ is indistinguishable from the rotation operator $U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar}$. However, if our system has any internal structure, circular translations using the centre of mass position operator \mathbf{X} and centre of mass



Figure 5: The rotation operator $U(\boldsymbol{\alpha})$ swings a system around the origin and also rotates its orientation in \mathbb{R}^3 , while a circular translation merely moves the system around a circular path, without affecting its orientation. The difference is a rotation of the body around its own centre of mass, reorientating it without changing its location.

momentum \mathbf{P} are not the same as rotations: the difference is best explained by a picture, which you can find in figure 5.

We define the *spin operator* \mathbf{S} to be the difference

$$\mathbf{S} = \mathbf{J} - \mathbf{L} \quad (4.66)$$

so that $\mathbf{J} = \mathbf{L} + \mathbf{S}$. From figure 5 we see that difference between rotating an object around some fixed origin and a translating it's centre of mass along a circular path is a rotation around the object's centre of mass. generates a rotation of the body *around its own centre of mass*. We thus expect (and will confirm below) that \mathbf{S} generates rotations of a body around its own centre of mass, reorienting it in space. This is why \mathbf{S} is called the spin operator.

In the case of a macroscopic body, made up of many constituent particles, it's reasonable to suppose that the spin operator just account for the difference between translating the body as a whole and translating each individual particle around their own arcs, with slightly different radii according to where in the body the particle is located. That is,

$$\mathbf{S} \stackrel{?}{=} \left(\sum_a \mathbf{X}_a \times \mathbf{P}_a \right) - \mathbf{X} \times \mathbf{P} = \sum_a \mathbf{x}_a \times \mathbf{p}_a \quad (4.67)$$

where \mathbf{X}_a and \mathbf{P}_a are position and translation operators for each individual particle and \mathbf{x}_a and \mathbf{p}_a are their positions and momenta *relative* to the centre of mass³⁰. However, for an object consisting of many particles such a description is clearly going to be very cumbersome. More fundamentally, *we do not know what the 'fundamental' constituents of our object really are*. (For example, if I sit on a merry-go-round, is the 'right' quantum description of my motion given in terms of my cells, or my atoms, or protons and neutrons, or quarks and gluons, or bits of string, or ...?) It's thus crucial that we can consider objects *as a whole* in quantum mechanics, just as we can classically. For rotations, understanding \mathbf{S} will allow us to do this. In addition, as we'll see in section 5.3.1, one of the surprises of quantum mechanics is that even fundamental particles such as electrons and photons may have an 'intrinsic' spin that (as far as we know) is not related to any composite structure.

Fortunately, commutation relations involving the spin operator \mathbf{S} are easy to obtain from the ones we already have for \mathbf{J} and \mathbf{L} . Firstly, the fundamental relations

$$[J_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k \quad \text{and} \quad [J_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k \quad (4.68)$$

show that

$$[J_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k, \quad (4.69)$$

³⁰We'll understand how to properly describe the quantum mechanics of a system with many degrees of freedom (*e.g.* many particles) in chapter 7.

so that the angular momentum operator \mathbf{L} also transforms as a vector under rotations. Recalling from 1B QM that³¹ $[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k$, we have

$$\begin{aligned} [S_i, S_j] &= [J_i - L_i, J_j - L_j] = [J_i, J_j] - [J_i, L_j] - [L_i, J_j] + [L_i, L_j] \\ &= i\hbar \sum_k \epsilon_{ijk} (J_k - L_k) = i\hbar \sum_k \epsilon_{ijk} S_k. \end{aligned} \quad (4.70)$$

It then immediately follows that $[\mathbf{S}, \mathbf{S}^2] = 0$. This algebra – the same as that obeyed by the components of \mathbf{J} and \mathbf{L} – confirms that the spin operators \mathbf{S} generate some form of rotation. On the other hand, since $[L_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k$ and $[L_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k$, we have

$$\begin{aligned} [S_i, X_j] &= [J_i, X_j] - [L_i, X_j] = 0 \\ [S_i, P_j] &= [J_i, P_j] - [L_i, P_j] = 0 \end{aligned} \quad (4.71)$$

These commutation relations confirm that the spin operator has nothing to do with an object's location in or motion through space, but is purely to do with rotating its *intrinsic* orientation. Finally, since \mathbf{S} commutes with both \mathbf{X} and \mathbf{P} , it also commutes with \mathbf{L} :

$$[S_i, L_j] = 0 \quad \forall i, j. \quad (4.72)$$

This allows us to factorize the operator $U(\boldsymbol{\alpha})$ describing finite rotations as

$$U(\boldsymbol{\alpha}) = e^{-i\boldsymbol{\alpha} \cdot \mathbf{J}/\hbar} = e^{-i\boldsymbol{\alpha} \cdot (\mathbf{L} + \mathbf{S})/\hbar} = e^{-i\boldsymbol{\alpha} \cdot \mathbf{L}/\hbar} e^{-i\boldsymbol{\alpha} \cdot \mathbf{S}/\hbar} = e^{-i\boldsymbol{\alpha} \cdot \mathbf{S}/\hbar} e^{-i\boldsymbol{\alpha} \cdot \mathbf{L}/\hbar}. \quad (4.73)$$

As in figure 5, these equations confirm that we can think of a quantum rotation as consisting of a translation of a body's centre of mass along an arc centred on the origin together with a simultaneous rotation of the body around its own centre of mass by the same amount. The order in which we perform these two operations makes no difference.

4.4 Time Translations

It's not only transformations of space that are represented on \mathcal{H} by unitary operators. Consider translations in time. These again form an Abelian group, since sending t_0 to $t_0 + t$ and then to $t_0 + t + t'$ gives the same end result as does $t_0 \rightarrow t_0 + t' \rightarrow t_0 + t' + t$. If we want the total probability of our particle being found *somewhere* in \mathbb{R}^3 at *all* times, then time translation must also be represented by a unitary operator $U(t) : \mathcal{H} \rightarrow \mathcal{H}$. Since we can translate through an arbitrarily small time, the time translation operator takes the form

$$U(t) = \exp\left(-\frac{i}{\hbar} H t\right) \quad (4.74)$$

where H/\hbar must be Hermitian and is, by definition, the generator of time translations. This generator must have dimensions 1/(time), so the conventional factor of \hbar means that H itself has dimensions of energy. Of course, H will turn out to be the Hamiltonian, but for now I'd like you to think of it more abstractly just as the generator of time translations.

³¹Check you're able to derive this statement from the definition of \mathbf{L} and the fundamental commutation relations $[X_i, P_j] = i\hbar \delta_{ij}$.

The fact that $U(t)$ represents the effect of a time translation on \mathcal{H} means that if a particle is in some state $|\psi(0)\rangle$ at time $t_0 = 0$, translating forward to time t we will find it in the state

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle. \quad (4.75)$$

In particular, the difference between $|\psi(t)\rangle$ and the state we find a short time later is

$$|\psi(t + \delta t)\rangle - |\psi(t)\rangle = -\frac{i}{\hbar} \delta t H |\psi(t)\rangle + O(\delta t^2) \quad (4.76)$$

or, taking the limit $\delta t \rightarrow 0$,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (4.77)$$

This, famously, is the *Time Dependent Schrödinger Equation*, which I'll often abbreviate to *TDSE* in what follows. It's just the infinitesimal version of the statement that all states in \mathcal{H} evolve in time according to the action of a unitary operator $U(t)$.

4.4.1 The Heisenberg Picture

In 1B QM, we were used to the idea that states evolve in time according to the TDSE. However, just as we did for spatial transformations, we can instead work in a picture where the states are time independent and instead the *operators* evolve in time. To see how this works, suppose O_S is some operator that contains no explicit time dependence. Then the amplitude for the state $O_S|\psi(t)\rangle$ to agree at time t with the state $|\chi(t)\rangle$ is

$$\langle \chi(t) | O_S | \psi(t) \rangle = \langle \chi(0) | U^{-1}(t) O_S U(t) | \psi(0) \rangle \quad (4.78)$$

The right hand side here only makes explicit reference to the initial values of the states $|\psi\rangle$ and $|\chi\rangle$. Since it holds true for any pair of states, just as above we can obtain the same results by always working with the initial states and instead evolving the operators as

$$O_H(t) = U^{-1}(t) O_S U(t). \quad (4.79)$$

The version of quantum mechanics where we use the TDSE to evolve the states in time, leaving operators unaltered, is known as the *Schrödinger picture*, whereas the version where the states are fixed at their initial values and instead the operators evolve in time is called the *Heisenberg picture*.

In fact, all this has a precise analogue in classical mechanics. Classically, there are also two ways of thinking about time evolution. On the one hand, we can think of a particle moving in some way through phase space M . If we know it's location $(\mathbf{x}(t), \mathbf{p}(t)) \in M$ for every time t we can compute any quantity we wish, represented by some function $f : M \rightarrow \mathbb{R}$, by evaluating f at the location of our particle, obtaining the value $f(\mathbf{x}(t), \mathbf{p}(t))$. (This is the perspective we took in the Introduction.) However, Newton's Laws are deterministic, so, given a force, the entire trajectory is determined by the initial conditions $(\mathbf{x}_0, \mathbf{p}_0)$. This suggests a perspective in which the 'state' of our particle is simply a choice of initial conditions. These initial conditions do not themselves evolve, rather, it is the quantities we measure that vary in time. Thus, instead of thinking of a physical quantity f as a map from

phase space, we treat it just as a map from time, so $f : [t_0, \infty) \rightarrow \mathbb{R}$. You'll examine these classical pictures further if you're taking the Classical Dynamics course, in the context of Hamiltonian mechanics. (It's really for this reformulation of classical mechanics, not the particular H , that Hamilton is famous.)

Differentiating (4.79) wrt t shows that

$$\begin{aligned} \frac{d}{dt}O_{\text{H}}(t) &= \frac{d}{dt}(U^{-1}(t)O_{\text{S}}U(t)) = \frac{i}{\hbar}(U^{-1}(t)HO_{\text{S}}U(t) - U^{-1}O_{\text{S}}HU(t)) \\ &= \frac{i}{\hbar}[H, O_{\text{H}}(t)], \end{aligned} \quad (4.80)$$

where in the last step we used the fact that $[U(t), H] = 0$, since $U(t)$ depends only on H . This is *Heisenberg's equation of motion*. It's completely equivalent to Schrödinger's equation (4.77), and is also just a (very important!) special case of our general argument of how operators transform under the action of unitary groups of transformations. More generally, if the original operator O_{S} had some explicit time dependence of its own — independent of any particular particle's motion — then we would obtain a further term in this equation, modifying it to

$$\frac{d}{dt}O_{\text{H}}(t) = U^{-1}(t)\frac{\partial O_{\text{S}}}{\partial t}U(t) + \frac{i}{\hbar}[H, O_{\text{H}}(t)]. \quad (4.81)$$

For example, we may wish to understand the behaviour of a charged particle in the presence of an applied electric field. If the electric field is itself changing in time, then the potential in which the particle moves will be time-dependent even in the Schrödinger picture.

4.5 Dynamics

So far, we've simply defined some unitary operators $U(\mathbf{a})$, $U(\boldsymbol{\alpha})$ and $U(t)$ that are responsible for translating, rotating or evolving our system through space and time. The commutation relations of the generators \mathbf{P} and \mathbf{J} , and their commutation relations with \mathbf{X} were determined purely by the properties of the corresponding group of transformations of \mathbb{R}^3 . However, while we've seen that commutators such as $[H, \mathbf{X}]$, $[H, \mathbf{P}]$ and $[H, \mathbf{J}]$ are important in the Heisenberg picture, telling us how these operators change in time, we haven't yet given any way to actually *calculate* what such commutators should be. Similarly, although we've said that in the Schrödinger picture, states evolve in time according to $|\psi(t)\rangle = U(t)|\psi(0)\rangle$, we haven't given any way to work out what the action of the generator H on a state should actually be.

To do so, we must specify the *dynamics*: we must provide a relation $H = H(\mathbf{X}, \mathbf{P})$ giving H in terms of the other operators whose commutators we already understand. The simplest (non-trivial) such relation is³²

$$H = \frac{1}{2m}\mathbf{P}^2 \quad (4.82)$$

³²Note that this is the *first* time we've pinned ourselves down to a non-relativistic theory. Up to this point, everything we've said holds good in relativistic quantum mechanics and, suitably interpreted, also in quantum field theory. In the lectures we studied translations and rotations, but in the **first problem set** you'll extend this to also consider Galilean boosts and will show that the non-relativistic Hamiltonian is compatible with the Galilean algebra. If one instead wishes to study a relativistic version of quantum theory, one starts by finding unitary operators representing the action of the Poincaré group $\mathbb{R}^{3,1} \rtimes SO(3, 1)$

where m is a constant with dimensions of mass. Equation (4.82) relates how a state evolves in time (via H) to how its location is translated through space (via \mathbf{P}); in other words, it's telling us something about how particles *move*! In particular, with this form of Hamiltonian the difference between the expected location of a state $|\psi\rangle$ at an initial time and a short time δt later is

$$\begin{aligned}\langle\psi|U^{-1}(\delta t)\mathbf{X}U(\delta t)|\psi\rangle - \langle\psi|\mathbf{X}|\psi\rangle &= \frac{i}{\hbar}\delta t\langle\psi|[H,\mathbf{X}]|\psi\rangle + O(\delta t^2) \\ &= \frac{\delta t}{m}\langle\psi|\mathbf{P}|\psi\rangle + O(\delta t^2).\end{aligned}\tag{4.83}$$

In the limit $\delta t \rightarrow 0$ we see that the expected *velocity* of the particle is $\langle\psi|\mathbf{P}|\psi\rangle/m$. Equivalently, since this is true for all $|\psi\rangle \in \mathcal{H}$, we can write

$$\frac{d}{dt}\mathbf{X}(t) = \frac{\mathbf{P}(t)}{m}\tag{4.84}$$

in the Heisenberg picture. Furthermore, since $[H,\mathbf{P}] = 0$ for this Hamiltonian, $\mathbf{P}(t) = \mathbf{P}(0)$ and all states travel in uniform motion.

In the real world, we observe that particles do not always travel with constant velocities: they may slow down, speed up, or change direction as they encounter various obstacles. These obstacles are typically located at various different points in space. To allow for this, we generalise our dynamical relation (4.82) to

$$H = \frac{1}{2m}\mathbf{P}^2 + V(\mathbf{X}).\tag{4.85}$$

The first term on the right is the *kinetic term* and is the contribution to the Hamiltonian due to the state's travelling through space. The second, *potential term* is the contribution due to the state's location. This familiar form is still a very special case of the general statement $H = H(\mathbf{X},\mathbf{P})$ and later in the course we'll meet examples of Hamiltonians that don't fit this form.

Repeating the calculation of (4.84) we still find that $d\mathbf{X}(t)/dt = \mathbf{P}(t)/m$, but now $[H,\mathbf{P}] \neq 0$ so the rate of motion through space will not always be the same. Rather, using (4.80), the Heisenberg picture operators obey

$$\frac{d}{dt}\mathbf{P}(t) = \frac{i}{\hbar}[H,\mathbf{P}(t)] = -\nabla V(t),\tag{4.86}$$

where $\nabla V(t) = U^{-1}(t)(\partial V/\partial\mathbf{X})U(t) = -\nabla V(\mathbf{X}(t))$. Thus the motion slows down or speeds up according to the gradient of $V(\mathbf{X})$.

Most of our common intuition about momentum comes from equation (4.86). We 'feel' that a tennis ball travelling with a certain speed has less momentum than a cannonball travelling with the same speed, and less than the same tennis ball travelling faster, because

of special relativity. The relation (4.82) between H and \mathbf{P} does not respect the commutation relations appropriate for Poincaré symmetry, but of course the dynamical relation $H^2 = c^2\mathbf{P}^2 + m^2c^4$ does. The difficulty with relativistic quantum mechanics lies not with symmetries, but with interactions. The proper treatment of these requires Quantum Field Theory, a far deeper subject.

we've known from an early age that it will cause us less damage to stand in the way of the first tennis ball than either of the other two. This is really a statement about what our unfortunate bodies will have to do in order to bring these projectiles to rest (or how much effort we will have to exert to launch them in the first place). In other words, our intuitive notion of momentum is built on a feeling for the energy our body will gain as we slow the projectile down during the impact. This energy then excites the atoms that ultimately make up our nerves, muscles and bones, becoming dissipated through our bodies. Ultimately, it is the dynamical relation between the Hamiltonian and other operators that justifies us identifying the operator \mathbf{P}/\hbar — by definition, the generator of translations — with our pre-existing notion of *momentum* (in units of \hbar). Only after we specify our dynamics do commutation relations such as $[X_i, P_j] = i\hbar \delta_{ij}$ tell us, via Born's 2nd postulate of QM, that a particle cannot simultaneously be in a state of well-defined position and of well-defined momentum.

We'll often find it useful to write the kinetic term in a slightly different form. Keeping careful track of the operator ordering, you'll show in the problem sets that

$$\mathbf{L}^2 = (\mathbf{X} \times \mathbf{P}) \cdot (\mathbf{X} \times \mathbf{P}) = \mathbf{X}^2 (\mathbf{P}^2 - P_r^2), \quad (4.87)$$

where for $\hat{\mathbf{X}} = \mathbf{X}/|\mathbf{X}|$,

$$P_r = \frac{1}{2} (\hat{\mathbf{X}} \cdot \mathbf{P} + \mathbf{P} \cdot \hat{\mathbf{X}}) \quad (4.88)$$

is the radial momentum operator. Consequently, we can write the Hamiltonian (4.85) in terms of the generator of circular translations and the radial momentum operator P_r as

$$H = \frac{P_r^2}{2m} + \frac{1}{2m\mathbf{X}^2} \mathbf{L}^2 + V(\mathbf{X}) \quad (4.89)$$

where we note that $[\mathbf{L}, \mathbf{X}^2] = 0$, so the order of the operators in the second term is irrelevant provided we keep \mathbf{X}^2 together as a composite operator. (In particular, in the position representation, \mathbf{X}^2 acts as multiplication by $\mathbf{x}^2 = r^2$, which is certainly unaffected by translation in a circle.) Most of our intuitive feeling about angular momentum — mine comes from happy childhood hours spent on merry-go-rounds — is encapsulated in this relation between the Hamiltonian and \mathbf{L}^2 . Again, it's the form of the Hamiltonian which is the basis of our identification of the generator \mathbf{L}/\hbar of circular transformations with angular momentum (in units of \hbar).

4.5.1 Symmetries and Conservation Laws

One of the immediate uses of the Heisenberg picture is to deduce a relation between symmetries of the Hamiltonian and conserved quantities³³. If it happens that an operator Q commutes with the Hamiltonian, then it also commutes with $e^{-iHt/\hbar}$, so the Heisenberg picture operator

$$Q(t) = U^{-1}(t) Q U(t) = U^{-1}(t) U(t) Q = Q, \quad (4.90)$$

³³Those of you taking the Classical Dynamics course will see a corresponding relation in Hamilton's approach to classical physics.

coinciding with the Schrödinger operator for all t . Infinitesimally, provided Q has no explicit time dependence, this is

$$\frac{d}{dt}Q(t) = \frac{i}{\hbar}[H, Q(t)] = \frac{i}{\hbar}U^{-1}(t)[H, Q]U(t) = 0. \quad (4.91)$$

Operators that are time independent even in the Heisenberg picture are said to be *conserved*. We've shown that conserved operators are just those that commute with the Hamiltonian.

Suppose we prepare a particle to be in an eigenstate of some conserved operator at time $t = 0$, so $Q|\psi(0)\rangle = q|\psi(0)\rangle$. Then at any later time we have

$$Q|\psi(t)\rangle = QU(t)|\psi(0)\rangle = U(t)Q|\psi(0)\rangle = qU(t)|\psi(0)\rangle = q|\psi(t)\rangle \quad (4.92)$$

so our particle remains in an eigenstate of Q at all subsequent times (provided the state evolves according to the TDSE). For this reason, it's usually sensible to expand our states in a basis of eigenstates of a maximal set of *conserved* operators, rather than a maximal commuting set of any old operators. We label the states in this basis by their corresponding eigenvalues, because the same labelling will remain valid at subsequent times. For example, provided H has no explicit time dependence³⁴ $[H, H] = 0$ trivially, so the Hamiltonian itself is always conserved, and it is often useful to work in a basis of energy eigenstates.

By far the most important source of conserved quantities is *symmetries* of the Hamiltonian. These are transformations that leave the Hamiltonian invariant. We've seen that a generic unitary operator $U(\theta) = e^{-i\theta T}$ representing some transformation of space on \mathcal{H} acts on operators as in equation (4.14). In particular, its action on the Hamiltonian is

$$H \mapsto U^{-1}(\theta)HU(\theta). \quad (4.93)$$

If the Hamiltonian is invariant under this transformation – *i.e.*, if the transformation is a symmetry – then

$$U^{-1}(\theta)HU(\theta) = H, \quad (4.94)$$

or equivalently

$$[T, H] = 0 \quad (4.95)$$

for an infinitesimal symmetry transformation, where T is the Hermitian generator. This is exactly the same condition (4.91) we had for the generator T to be conserved, so *symmetries of the Hamiltonian correspond to conserved quantities*³⁵.

For example, if the Hamiltonian is translationally invariant then $[\mathbf{P}, H] = 0$, so for such Hamiltonians, momentum will be conserved. Note that it's not necessary for the Hamiltonian to be independent of each particle's position \mathbf{x}_a if H is to be translationally

³⁴Even if the Hamiltonian does contain explicit time dependence, for example because it describes the dynamics of a charged particle in a varying electric field, $[H(t), H(t)] = 0$ trivially. However, $[H(t'), H(t)]$ may not vanish as the form of the Hamiltonian itself changes.

³⁵If you're taking the Classical Dynamics or Integrable Systems courses, you'll meet the corresponding statement in classical mechanics in the guise of Nöther's theorem.

invariant, so long as any potential term $V(\mathbf{x}_a - \mathbf{x}_b)$ depends only on the relative positions. Similarly, if the Hamiltonian is rotationally invariant then $[\mathbf{J}, H] = 0$ so angular momentum will be conserved. In this case, J_z , \mathbf{J}^2 and H form a set of mutually commuting operators, so we can expand a general state in a basis $\{|n, \ell, m, \dots\rangle\}$ where the labels (n, ℓ, m) refer to the eigenvalues of H , \mathbf{J}^2 and J_z . We've left open the possibility that our states have further conserved properties, indexed by further labels. These will often contain information about the 'internal' state of our system.

4.6 Parity

Not all transformations are continuous. In physics, the most prominent example of a discrete transformation is the *parity transformation* \mathcal{P} , acting on \mathbb{R}^3 as $\mathcal{P} : \mathbf{x} \mapsto -\mathbf{x}$. Since $\det(\mathcal{P}) = -1$, this is different from a rotation, so may have consequences that cannot be deduced by considering only rotations.

On Hilbert space, parity transformations will still be represented by a unitary operator, but this operator will no longer involve a parameter that may be taken infinitesimally small. Consequently there's no generator associated with parity transformations. Calling the unitary parity operator Π , we have

$$\Pi^{-1} \mathbf{X} \Pi = \mathcal{P} \mathbf{X} = -\mathbf{X} \quad (4.96)$$

as its defining property. Multiplying through on the right by Π , we can also write this as

$$\{\Pi, \mathbf{X}\} = \Pi \mathbf{X} + \mathbf{X} \Pi = 0. \quad (4.97)$$

where the bracket $\{A, B\}$ is sometimes called the *anticommutator* of A and B . Translating through \mathbf{a} and then applying the parity operator $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is the same as first applying \mathcal{P} , then translating through $-\mathbf{a}$, so

$$\Pi^{-1} U(\mathbf{a}) \Pi = U(-\mathbf{a}), \quad (4.98)$$

so that the momentum operator also anticommutes with the parity operator: $\Pi^{-1} \mathbf{P} \Pi = -\mathbf{P}$. More generally, if \mathbf{V} is any operator that transforms as a vector under rotations, we say \mathbf{V} is a *vector operator* if also

$$\Pi^{-1} \mathbf{V} \Pi = -\mathbf{V}, \quad (4.99)$$

so that $\{\Pi, \mathbf{V}\} = 0$.

If instead

$$\Pi^{-1} \mathbf{V} \Pi = +\mathbf{V} \quad (4.100)$$

or equivalently $[\Pi, \mathbf{V}] = 0$, then \mathbf{V} is a *pseudovector operator* (provided it transforms appropriately under rotations). The most prominent example of a pseudovector operator is the rotation generator \mathbf{J} : since the parity transformation \mathcal{P} acts on \mathbb{R}^3 as $-I_{3 \times 3}$, rotations obey $\mathbf{R}(\boldsymbol{\alpha}) \mathcal{P} = \mathcal{P} \mathbf{R}(\boldsymbol{\alpha})$, or $\mathcal{P}^{-1} \mathbf{R}(\boldsymbol{\alpha}) \mathcal{P} = \mathbf{R}(\boldsymbol{\alpha})$. This gives

$$\Pi^{-1} U(\boldsymbol{\alpha}) \Pi = U(\boldsymbol{\alpha}) \quad (4.101)$$

for the parity and rotation operators in Hilbert space, or $\Pi^{-1}\mathbf{J}\Pi = \mathbf{J}$ for the rotation generator. Likewise, the orbital angular momentum operator obeys

$$\Pi^{-1}\mathbf{L}\Pi = \Pi^{-1}(\mathbf{X} \times \mathbf{P})\Pi = (\Pi^{-1}\mathbf{X}\Pi) \times (\Pi^{-1}\mathbf{P}\Pi) = (-)^2 \mathbf{X} \times \mathbf{P} = +\mathbf{L} \quad (4.102)$$

so is a pseudovector as in classical mechanics. It follows that the spin operator $\mathbf{S} = \mathbf{J} - \mathbf{L}$ is also a pseudovector.

Similarly, operators S that are invariant under rotations are *scalar operators* if in addition $\Pi^{-1}S\Pi = S$ so that they are unchanged under parity, and are *pseudoscalar operators* if instead $\Pi^{-1}S\Pi = -S$. (Notice that the signs for scalars and pseudoscalars are the opposite way round to those for vectors and pseudovectors.) As usual, the parity of a system will be conserved if $[H, \Pi] = 0$. As a simple example, if our system is governed by the dynamical relation $H = \mathbf{P}^2/2m + V(\mathbf{X})$ where the potential is an even function, then

$$\begin{aligned} \Pi^{-1}H\Pi &= \frac{1}{2m}(\Pi^{-1}\mathbf{P}\Pi) \cdot (\Pi^{-1}\mathbf{P}\Pi) + \Pi^{-1}V(\mathbf{X})\Pi \\ &= \frac{\mathbf{P}^2}{2m} + V(\Pi^{-1}\mathbf{X}\Pi) \\ &= \frac{\mathbf{P}^2}{2m} + V(\mathbf{X}) = H, \end{aligned} \quad (4.103)$$

so parity is conserved for such Hamiltonians.

Parity transformations form the group $G = \mathbb{Z}_2$, since carrying out a non-trivial parity transformation twice just brings us back to where we were. Thus $\mathcal{P}^2 = 1_{\mathbb{R}^3}$ and applying our homomorphism shows that

$$\Pi^2 = 1_{\mathcal{H}} \quad (4.104)$$

for the parity operator on Hilbert space, too. Now suppose that $|\psi\rangle$ is an eigenstate of Π with eigenvalue η . Then³⁶

$$|\psi\rangle = \Pi^2|\psi\rangle = \eta\Pi|\psi\rangle = \eta^2|\psi\rangle \quad (4.105)$$

so the spectrum of Π is just $\{+1, -1\}$.

If $|\mathbf{v}\rangle$ is an eigenstate of a vector operator \mathbf{V} , with $\mathbf{V}|\mathbf{v}\rangle = \mathbf{v}|\mathbf{v}\rangle$, then from (4.99), the parity reversed state $|\mathbf{v}'\rangle = \Pi|\mathbf{v}\rangle$ obeys

$$\mathbf{V}|\mathbf{v}'\rangle = \mathbf{V}\Pi|\mathbf{v}\rangle = -\Pi\mathbf{V}|\mathbf{v}\rangle = -\mathbf{v}\Pi|\mathbf{v}\rangle = -\mathbf{v}|\mathbf{v}'\rangle \quad (4.106)$$

so is also an eigenstate of \mathbf{V} , but with opposite sign eigenvalue. In particular, $\Pi|\mathbf{x}\rangle$ is an eigenstate of the position operator with eigenvalue $-\mathbf{x}$, so

$$\Pi|\mathbf{x}\rangle = c|-\mathbf{x}\rangle. \quad (4.107)$$

for some constant $c \in \mathbb{C}$. This is not an eigenvalue equation, because the *rhs* involves $|-\mathbf{x}\rangle$ rather than $|\mathbf{x}\rangle$. However, we can immediately identify c with η because applying Π a second time gives³⁷

$$|\mathbf{x}\rangle = \Pi^2|\mathbf{x}\rangle = c^2|\mathbf{x}\rangle, \quad (4.108)$$

³⁶This argument is adequate for our course, and leads to a correct conclusion, but it hides a number of subtleties. See *e.g.* Weinberg *The Quantum Theory of Fields*, vol. 1, section 4.7 for a fuller discussion.

³⁷Note that it is important here that $\Pi|\mathbf{x}\rangle = c|-\mathbf{x}\rangle$ for *all* values of \mathbf{x} .

so again $c^2 = 1$ and $c = \pm 1$. More generally, given any state $|\psi\rangle$, the wavefunction of the parity transformed state $\Pi|\psi\rangle$ is

$$\langle \mathbf{x} | \Pi | \psi \rangle = \eta^{-1} \langle -\mathbf{x} | \psi \rangle = \eta \psi(-\mathbf{x}), \quad (4.109)$$

so, up to a sign, the wavefunction of the new state takes the same value at \mathbf{x} as the original one did at $-\mathbf{x}$.

For example, suppose that for some particle, $|\mathbf{x}\rangle$ obeys $\Pi|\mathbf{x}\rangle = |-\mathbf{x}\rangle$, with $\eta = +1$. Recalling from 1B QM that the spherical harmonics $Y_\ell^m(\mathbf{x})$ obey $Y_\ell^m(-\mathbf{x}) = (-1)^\ell Y_\ell^m(\mathbf{x})$, we see that if this particle is in the state $|n, \ell, m\rangle$ whose wavefunction $\langle \mathbf{x} | n, \ell, m \rangle = R_n(|\mathbf{x}|) Y_\ell^m(\mathbf{x})$ (where the radial part $R_n(r)$ of the wavefunction determines the energy level n), applying the parity operator gives

$$\langle \mathbf{x} | \Pi | n, \ell, m \rangle = \langle -\mathbf{x} | n, \ell, m \rangle = (-1)^\ell \langle \mathbf{x} | n, \ell, m \rangle \quad (4.110)$$

and consequently $\Pi|n, \ell, m\rangle = (-1)^\ell |n, \ell, m\rangle$.

Parity transformations act in a more complicated way on macroscopic objects, or systems with some internal structure. This is clear classically: the parity transformation of a book initially located at \mathbf{x} is not simply a book located at $-\mathbf{x}$, but rather a mirror image (simultaneously left–right, up–down and front–back) of the book. As with the distinction between \mathbf{J} and \mathbf{L} , for a macroscopic body we might hope to account for this by defining a separate parity operator for each constituent particle. However, it turns out that subatomic particles such as the photon or pion have their own ‘intrinsic’ parities. Thus, whether a given Π -eigenstate has $\eta = +1$ or $\eta = -1$ depends not only on details of the spatial wavefunction, but also on the species of particle the state describes.

We’ll explore intrinsic parity further in section 7.4, but as an example of its use, consider transitions between different energy levels of an atom. These are usually mediated by electromagnetic interactions: during such a transition, the atom’s electrons either emit or absorb a photon whose energy $\hbar\omega$ is equal to the difference in energy ΔE between the two levels involved in the transition. As we’ll study in more detail later, when the corresponding wavelength is much larger than the typical size of the atom, the transition rate $\Gamma(i \rightarrow f)$ is given by

$$\Gamma(i \rightarrow f) = \frac{4(\Delta E)^3}{c^3 \hbar^4} |\langle f | \mathbf{D} | i \rangle|^2, \quad (4.111)$$

where $|i\rangle$ and $|f\rangle$ are the initial and final states of the atom, and $\mathbf{D} = \sum e_a \mathbf{X}_a$ is the operator corresponding to the atom’s electric dipole moment. (The sum runs over all the electrons in the atom.) Reversing the parity of space affects all the electrons equally, so the parity operator obeys $\Pi^{-1} \mathbf{X}_a \Pi = -\mathbf{X}_a$ for each \mathbf{X}_a . Thus $\Pi^{-1} \mathbf{D} \Pi = -\mathbf{D}$.

Now suppose that the initial and final states of the atom are eigenstates of Π , with eigenvalues η_i and η_f , respectively. Then³⁸

$$\eta_i \eta_f \langle f | \mathbf{D} | i \rangle = \langle f | \Pi^{-1} \mathbf{D} \Pi | i \rangle = -\langle f | \mathbf{D} | i \rangle \quad (4.112)$$

³⁸Note that since the parity operator Π is unitary its quantum numbers η behave multiplicatively, whilst those of a Hermitian generator of a continuous symmetry (such as \mathbf{P}) behave additively.



Figure 6: *Most atomic transitions are due to the absorption or emission of a photon from a single electron in the atom. The photon's angular frequency $\omega = \Delta E/\hbar$.*

so the amplitude for the transition will vanish unless the initial and final atomic states have opposite parity, $\eta_i \eta_f = -1$. In the most common case, just a single electron is involved in the transition and in this case equation (4.110) shows that $\eta_i \eta_f = (-1)^{\ell_i + \ell_f}$, where $\ell_{i,f}$ are the total orbital angular momentum quantum numbers of the initial and final states of this electron. Thus, in any radiative atomic transition involving just a single electron, the orbital angular momentum quantum number ℓ must change by an odd number. If our initial and final states kept track of the photon as well as the atom, we'd be able to say overall parity is conserved in such transitions if we assign an intrinsic parity -1 to the photon that is emitted or absorbed in the transition. (The intrinsic parity of the photon is related to the fact that the electromagnetic vector potential \mathbf{A} , whose quantum version provides the description of the photon, indeed transforms as a vector, not a pseudovector.) In fact, $[H, \Pi] = 0$ holds in general for particles travelling in the presence of any type of electromagnetic interaction.

Having $[H, \Pi]$ be zero means that if at time $t = 0$ you set up a system to be a mirror image of another system, then their subsequent evolution will be identical in the sense that if you observe them at a later time, they will still be mirror images of one another. Hence, when $[H, \Pi] = 0$ it is impossible to tell whether a system is being observed directly or through a mirror. One of the major surprises of twentieth-century physics was an experiment by Wu *et al.* in 1957, which showed that in fact the Hamiltonian of our Universe has $[H, \Pi] \neq 0$ — you can see things in a mirror that are impossible in our World!