Transformations and symmetries in quantum mechanics

These notes give a brief and basic introduction to some central aspects concerning transformations and symmetries in quantum mechanics. Examples discussed include translations in space and time, as well as rotations.

Example 1: Translations in space

Translations in space are also called spatial translations, and sometimes even just "translations" for short, with "spatial" left implicit. To introduce the concept, let us consider the simplest example of a single particle in one spatial dimension. The state $|x\rangle$ is the eigenstate of the position operator $\hat{x}$ with eigenvalue x:

$$\hat{x}|x\rangle = x|x\rangle. \tag{1}$$

The eigenstates of $\hat{x}$ obey $\langle x|x'\rangle = \delta(x - x')$. Now consider the state $\exp(-i\hat{p}_x\Delta x/\hbar)|x\rangle$, where $\hat{p}_x$ is the momentum operator and Δx is some arbitrary spatial displacement. As the commutator $[\hat{x}, \hat{p}_x] = i\hbar$ is just a c-number, the simplified version (39) of the Baker-Hausdorff theorem holds, which gives

$$\begin{aligned} \hat{x}\exp(-i\hat{p}_x\Delta x/\hbar)|x\rangle &= \exp(-i\hat{p}_x\Delta x/\hbar)\underbrace{\exp(+i\hat{p}_x\Delta x/\hbar)\hat{x}\exp(-i\hat{p}_x\Delta x/\hbar)}_{\hat{x}+\Delta x}|x\rangle \\ &= \exp(-i\hat{p}_x\Delta x/\hbar)(x + \Delta x)|x\rangle = (x + \Delta x)\exp(-i\hat{p}_x\Delta x/\hbar)|x\rangle. \end{aligned} \tag{2}$$

This shows that $\exp(-i\hat{p}_x\Delta x/\hbar)|x\rangle$ is an eigenstate of $\hat{x}$ with eigenvalue $x + \Delta x$. Also, because the operator $\exp(-i\hat{p}_x\Delta x/\hbar)$ is unitary, the appropriate normalization is preserved:

$$\langle x|\underbrace{\exp(+i\hat{p}_x\Delta x/\hbar)\exp(-i\hat{p}_x\Delta x/\hbar)}_{\hat{I}}|x'\rangle = \delta(x - x') = \delta(x - \Delta x - (x' - \Delta x)). \tag{3}$$

Therefore we conclude that

$$\exp(-i\hat{p}_x\Delta x/\hbar)|x\rangle = |x + \Delta x\rangle. \tag{4}$$

Thus the operator $\exp(-i\hat{p}_x\Delta x/\hbar)$ transforms $|x\rangle$ into $|x+\Delta x\rangle$, i.e. it produces translations in space by Δx. If the displacement Δx is infinitesimal, i.e. $\Delta x \to dx$, the operator $\exp(-i\hat{p}_x dx/\hbar)$ producing the transformation can be approximated as

$$\exp(-i\hat{p}_x dx/\hbar) \approx 1 - \frac{i}{\hbar}\hat{p}_x dx. \tag{5}$$

The generalization of these results to three dimensions is straightforward since components of the momentum operator in different directions commute with each other, so translations along different directions commute. Thus

$$\exp\left(-i\hat{\boldsymbol{p}}\cdot\Delta\boldsymbol{r}/\hbar\right)|\boldsymbol{r}\rangle = |\boldsymbol{r}+\Delta\boldsymbol{r}\rangle, \tag{6}$$

where $\boldsymbol{r} = (x, y, z)$, $\Delta\boldsymbol{r} = (\Delta x, \Delta y, \Delta z)$, and $\hat{\boldsymbol{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$. The generalization of (5) becomes

$$\exp\left(-i\hat{\boldsymbol{p}}\cdot d\boldsymbol{r}/\hbar\right) \approx 1 - \frac{i}{\hbar}\hat{\boldsymbol{p}}\cdot d\boldsymbol{r}. \tag{7}$$

One says that the momentum operator $\hat{\boldsymbol{p}}$ is the generator of infinitesimal translations in space.

Example 2: Translations in time

Translations in time are also sometimes called temporal translations. To study these we start from the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial|\Psi(t)\rangle}{\partial t} = \hat{H}|\Psi(t)\rangle. \tag{8}$$

We will limit our discussion to systems for which the Hamiltonian $\hat{H}$ does not depend explicitly on time. Then the solution of the Schrödinger equation can be written $|\Psi(t)\rangle = \exp\left(-i\hat{H}(t-t')/\hbar\right)|\Psi(t')\rangle$ (you can verify that differentiation of this equation leads back to (8)). Here $\exp(-iH(t-t')/\hbar)$ is called the time evolution operator. Its unitarity implies that the norm of the state vector is preserved in time: $\langle\Psi(t)|\Psi(t)\rangle = \langle\Psi(t')|\Psi(t')\rangle$, thus if the state is normalized at one time it will remain normalized at all times ("conservation of total probability"). Clearly one can write

$$\exp(-i\hat{H}\Delta t/\hbar)|\Psi(t)\rangle = |\Psi(t+\Delta t)\rangle, \tag{9}$$

where Δt is an arbitrary temporal displacement. In words, $\exp(-i\hat{H}\Delta t/\hbar)$ produces translations in time by Δt. An infinitesimal translation by time dt is produced by

$$\exp\left(-i\hat{H}dt/\hbar\right) \approx 1 - \frac{i}{\hbar}\hat{H}dt. \tag{10}$$

One says that the Hamiltonian operator $\hat{H}$ is the generator of infinitesimal translations in time.

Unitary transformations. Continuous vs. discrete transformations

Note the similar structure of the results for the transformations discussed so far: Both spatial and temporal translations are produced by an operator of the form

$$\exp(-i\hat{Q}\Delta a/\hbar) \tag{11}$$

where $\hat{Q}$ is a Hermitian operator ($\hat{Q}$ is often referred to as a *generator* in this context) and Δa is the displacement, which is a real-valued quantity. Thus (11) is a unitary operator, so it preserves inner products[1] and therefore also normalizations. For translations in space, Δa is the spatial displacement along some arbitrary direction in space, and $\hat{Q}$ is the component of the momentum operator along that direction, while for translations in time, Δa is the temporal displacement and $\hat{Q}$ is the Hamiltonian operator. For infinitesimal translations we can write

$$\exp(-i\hat{Q}\,da/\hbar) \approx 1 - \frac{i}{\hbar}\hat{Q}da. \tag{12}$$

Furthermore, using the mathematical identity

$$e^x = \lim_{N\to\infty}\left(1 + \frac{x}{N}\right)^N, \tag{13}$$

we see that for both types of transformations one also has the very natural result that a finite translation can be made by concatenating an infinite number of infinitesimal translations:

$$\exp(-i\hat{Q}\Delta a/\hbar) = \lim_{N\to\infty}\left(1 - \frac{i}{\hbar}\hat{Q}\frac{\Delta a}{N}\right)^N \tag{14}$$

As will be discussed in the next section, another very important class of transformations, namely that of rotations, is described by the same theoretical structure. In fact, the structure discussed here is very general, as almost all transformations in quantum mechanics are produced by unitary operators. The only exception is the time reversal transformation, which is produced by an *anti*-unitary operator. Also note that translations in space (along some direction) or time, as well as rotations (see the next section), can be carried out for arbitrary values of the appropriate displacement parameter Δa, which thus take values in a continuous set. For this reason such transformations are called *continuous*. Furthermore, as the continuous set also includes $\Delta a = 0$, which corresponds to the "do-nothing" transformation (i.e. the identity operator $\hat{I}$), it follows that the set of continuous transformations also includes infinitesimal transformations, which are described by (12). Infinitesimal transformations do however not exist for those kinds of transformations which are *discrete* (as opposed to continuous), whose displacement parameters can only take values in a discrete set (examples of discrete transformations include the time reversal and parity transformations).

Example 3: Rotations

A rotation is specified by a rotation axis and a rotation angle about that axis. Rotations around a given axis can be regarded as translations of an angular variable; in this limited sense, rotations are "angular translations". And like spatial and temporal translations, rotations around the same axis commute with each other. However, rotations around different

[1] In other words, if $\hat{U}$ is a unitary operator and $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are two arbitrary state vectors, the inner product of $U|\Psi_1\rangle$ and $U|\Psi_2\rangle$ is the same as that of $|\Psi_1\rangle$ and $|\Psi_2\rangle$: $\langle\Psi_1|\hat{U}^\dagger\hat{U}|\Psi_2\rangle = \langle\Psi_1|\Psi_2\rangle$.

axes do not.[2]

The rotation axis will be specified by the unit vector $\boldsymbol{n}$ pointing in the direction of the axis. Denoting the rotation angle by $\Delta\phi$, we claim that the operator producing the rotation is

$$\exp\left(-i\hat{\boldsymbol{J}}\cdot\boldsymbol{n}\,\Delta\phi/\hbar\right), \tag{15}$$

where $\hat{\boldsymbol{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ is the angular momentum operator. Thus rotation operators are of the general form discussed in Sec. with $\hat{Q} = \hat{\boldsymbol{J}}\cdot\boldsymbol{n}$ (i.e. the component of $\hat{\boldsymbol{J}}$ along the rotation axis) and $\Delta a = \Delta\phi$. One says that the angular momentum operator $\hat{\boldsymbol{J}}$ is the generator of infinitesimal rotations. The components of $\hat{\boldsymbol{J}}$ satisfy the angular momentum commutation relations[3]

$$[J_x, J_y] = i\hbar J_z, \quad [J_y, J_z] = i\hbar J_x, \quad [J_z, J_x] = i\hbar J_y. \tag{16}$$

The noncommutativity of rotations around different axes is related to the noncommutativity of these generators.

To illustrate that (15) is the operator producing rotations, let us consider a concrete example: the rotation by a finite angle $\Delta\phi$ about the z axis. We will show that the expectation value of $\hat{\boldsymbol{J}}$ transforms as a classical vector under this rotation, which is what one would expect. Denoting the state before the rotation by $|\Psi\rangle$, the rotated state is

$$\exp(-i\hat{J}_z\Delta\phi/\hbar)|\Psi\rangle \equiv |\Psi'\rangle. \tag{17}$$

Thus the expectation value of $\hat{J}_i$ $(i = x, y, z)$ in the rotated state is

$$\langle\Psi'|\hat{J}_i|\Psi'\rangle = \langle\Psi|\exp(i\hat{J}_z\Delta\phi/\hbar)\hat{J}_i\exp(-i\hat{J}_z\Delta\phi/\hbar)|\Psi\rangle. \tag{18}$$

We use the Baker-Hausdorff formula (36) to calculate the operator product here. For $i = x$ we get

$$\begin{aligned}
&\exp\left(i\hat{J}_z\Delta\phi/\hbar\right)\hat{J}_x\exp\left(-i\hat{J}_z\Delta\phi/\hbar\right)\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}[\hat{J}_x, -i\hat{J}_z\Delta\phi/\hbar]_n = \hat{J}_x + \left(\frac{i\Delta\phi}{\hbar}\right)\underbrace{[\hat{J}_z, \hat{J}_x]}_{i\hbar\hat{J}_y}\\
&+ \frac{1}{2!}\left(\frac{i\Delta\phi}{\hbar}\right)^2\underbrace{[\hat{J}_z, \underbrace{[\hat{J}_z, \hat{J}_x]}_{i\hbar\hat{J}_y}]}_{\hbar^2\hat{J}_x} + \frac{1}{3!}\left(\frac{i\Delta\phi}{\hbar}\right)^3\underbrace{[\hat{J}_z, \underbrace{[\hat{J}_z, [\hat{J}_z, \hat{J}_x]]}_{\hbar^2\hat{J}_x}]}_{i\hbar^3\hat{J}_y} + \cdots\\
&= \hat{J}_x\left(1 - \frac{\Delta\phi^2}{2!} + \cdots\right) - \hat{J}_y\left(\Delta\phi - \frac{\Delta\phi^3}{3!} + \cdots\right).
\end{aligned} \tag{19}$$

[2] This is a well-known fact to everyone who has tried to get a big piece of furniture up a narrow non-straight stairway.

[3] Also known as the SU(2) commutation relations.

The two series inside the parentheses are the Taylor expansions of $\cos\Delta\phi$ and $\sin\Delta\phi$, respectively. Thus

$$\exp{(i\hat{J}_z\Delta\phi/\hbar)}\hat{J}_x\exp{(-i\hat{J}_z\Delta\phi/\hbar)} = \hat{J}_x\cos\Delta\phi - \hat{J}_y\sin\Delta\phi. \tag{20}$$

An entirely analogous calculation shows that

$$\exp{(i\hat{J}_z\Delta\phi/\hbar)}\hat{J}_y\exp{(-i\hat{J}_z\Delta\phi/\hbar)} = \hat{J}_x\sin\Delta\phi + \hat{J}_y\cos\Delta\phi. \tag{21}$$

Finally, since $\exp{(\pm i\hat{J}_z\Delta\phi/\hbar)}$ commutes with $\hat{J}_z$ it follows immediately that

$$\exp{(i\hat{J}_z\Delta\phi/\hbar)}\hat{J}_z\exp{(-i\hat{J}_z\Delta\phi/\hbar)} = \hat{J}_z. \tag{22}$$

Introducing the simplified notation $\langle\Psi|\hat{J}_i|\Psi\rangle \equiv \langle\hat{J}_i\rangle$ and $\langle\Psi'|\hat{J}_i|\Psi'\rangle \equiv \langle\hat{J}_i\rangle'$, the results above imply that

$$\langle\hat{J}_x\rangle' = \langle\hat{J}_x\rangle\cos\Delta\phi - \langle\hat{J}_y\rangle\sin\Delta\phi, \tag{23}$$

$$\langle\hat{J}_y\rangle' = \langle\hat{J}_x\rangle\sin\Delta\phi + \langle\hat{J}_y\rangle\cos\Delta\phi, \tag{24}$$

$$\langle\hat{J}_z\rangle' = \langle\hat{J}_z\rangle. \tag{25}$$

These transformation rules are identical to those for a classical vector being rotated an angle $\Delta\phi$ around the z-axis. This illustrates that the expectation value $\langle\hat{\boldsymbol{J}}\rangle$ of the angular momentum operator behaves as a classical vector under rotations.

Transformed states vs. transformed operators (active vs. passive view)

Consider a transformation produced by some unitary operator $\hat{U}$. The transformation changes a state $|\Psi\rangle$ into

$$|\Psi'\rangle = \hat{U}|\Psi\rangle. \tag{26}$$

Let us now consider the expectation value of some operator $\hat{A}$ in the transformed state:

$$\langle\Psi'|\hat{A}|\Psi'\rangle = \langle\Psi|\hat{U}^\dagger\hat{A}\hat{U}|\Psi\rangle = \langle\Psi|\hat{A}'|\Psi\rangle, \tag{27}$$

where we defined

$$\hat{A}' = \hat{U}^\dagger\hat{A}\hat{U}. \tag{28}$$

Thus one can think of the expectation value in (27) in two alternative ways: either as the expectation value of the original operator $\hat{A}$ in the transformed state $|\Psi'\rangle$, or as the expectation value of the transformed operator $\hat{A}'$ in the original state $|\Psi\rangle$. The first interpretation, in which the transformation affects the state, is the *active* view of the transformation, while the second interpretation, in which the transformation affects the operator, is the *passive* view of the transformation. The two views are completely equivalent mathematically and neither of them is “more correct” than the other.

Although we did not emphasize it, we have already encountered some examples of transformed operators. From (2) and its generalization to three dimensions it follows that the position operator $\hat{\boldsymbol{r}}$ transforms like a classical vector under spatial translations ($\hat{\boldsymbol{r}} \to \hat{\boldsymbol{r}}' = \hat{\boldsymbol{r}} + \Delta\boldsymbol{r}$). And (20)-(22) show that the angular momentum operator $\hat{\boldsymbol{J}}$ transforms like a classical vector under rotations, as the lhs's of these equations are $\hat{J}'_x$, $\hat{J}'_y$, and $\hat{J}'_z$, respectively.

An example of the active vs. passive view is the Schrödinger vs. Heisenberg picture of time evolution produced by the operator $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$. In the Schrödinger picture the time evolution affects the states, not the operators. In the Heisenberg picture it is the other way around.

Invariants and symmetries

Suppose that the expectation value of an operator $\hat{A}$ is invariant under some transformation produced by a unitary operator $\hat{U}$. That is, for any state $|\Psi\rangle$,

$$\langle\Psi|\hat{A}|\Psi\rangle = \langle\Psi|\hat{A}'|\Psi\rangle. \tag{29}$$

Since $|\Psi\rangle$ is arbitrary, we conclude that $\hat{A}' = \hat{A}$. We say that $\hat{A}$ is invariant under the transformation. Acting with $\hat{U}$ from the left on $\hat{U}^\dagger\hat{A}\hat{U} = \hat{A}$ and using $\hat{U}\hat{U}^\dagger = \hat{I}$ gives

$$[\hat{A}, \hat{U}] = 0. \tag{30}$$

If the transformation in question is continuous, we can consider infinitesimal transformations given by $\hat{U} = 1 - i\hat{Q}\Delta a/\hbar + \hat{C}$ where $\hat{C}$ is an operator representing terms of order $(\Delta a)^2$ and higher. Inserting this into (30) gives $0 = -i[\hat{A}, \hat{Q}]\Delta a/\hbar + [\hat{A}, \hat{C}]$. Dividing this equation by Δa and then taking the limit $\Delta a \to 0$, the term $[\hat{A}, \hat{C}]/\Delta a \to 0$, which implies

$$[\hat{A}, \hat{Q}] = 0. \tag{31}$$

If $\hat{A} = \hat{H}$, i.e. if the Hamiltonian is invariant under the transformation, the transformation is said to be a *symmetry* of the Hamiltonian/theory/system. If the transformation is continuous, this is called a continuous symmetry; otherwise it is called a discrete symmetry. For a continuous symmetry, (31) holds, i.e.

$$[\hat{H}, \hat{Q}] = 0. \tag{32}$$

Using the Heisenberg equation of motion for the generator, i.e.[4] $d\hat{Q}(t)/dt = \frac{i}{\hbar}[\hat{H}, \hat{Q}(t)]$, one finds $d\hat{Q}(t)/dt = \frac{i}{\hbar}[\hat{H}, \hat{Q}(t)] = \frac{i}{\hbar}(\hat{H}e^{i\hat{H}t/\hbar}\hat{Q}e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar}\hat{Q}e^{-i\hat{H}t/\hbar}\hat{H}] = \frac{i}{\hbar}e^{i\hat{H}t/\hbar}[\hat{H}, \hat{Q}]e^{-i\hat{H}t/\hbar}$, which upon using (32) gives

$$\frac{d\hat{Q}(t)}{dt} = 0, \tag{33}$$

[4]We assume that $\hat{Q}$ has no explicit time dependence (thus there is no term $\partial\hat{Q}/\partial t$ on the rhs of its Heisenberg equation of motion).

i.e. $\hat{Q}(t) = \hat{Q}$, i.e. time-independent, i.e. the generator is a conserved quantity. This result, that *a continuous symmetry implies a conserved quantity*, is often referred to as *Noether's theorem*.[5] Applying this theorem to spatial translations and rotations, we see that

$$\text{If the Hamiltonian is invariant under spatial translations, } \hat{\boldsymbol{P}} \text{ is conserved.} \quad (34)$$

$$\text{If the Hamiltonian is invariant under rotations, } \hat{\boldsymbol{J}} \text{ is conserved.} \quad (35)$$

Also, note that we assumed from the outset in our discussion that $\hat{H}$ is not explicitly time-dependent, and then it follows automatically from Heisenberg's equation of motion for $\hat{H}$ that $\hat{H}$ is conserved. Applying Noether's theorem to temporal translations just reproduces the result that $\hat{H}$ is conserved.

A The Baker-Hausdorff theorem

The Baker-Hausdorff theorem is extremely useful for evaluating expressions of the form $\exp(-\hat{B})\hat{A}\exp(\hat{B})$ which are ubiquitous when considering transformations in quantum mechanics. The theorem states that

$$e^{-\hat{B}}\hat{A}\,e^{\hat{B}} = \hat{A} + [\hat{A},\hat{B}] + \frac{1}{2!}[[\hat{A},\hat{B}],\hat{B}] + \cdots . \quad (36)$$

The rhs can be written more succinctly as $\sum_{n=0}^{\infty}\frac{1}{n!}[\hat{A},\hat{B}]_n$ where $[\hat{A},\hat{B}]_n$ is a *nested* commutator, defined recursively as

$$[\hat{A},\hat{B}]_0 \equiv \hat{A}, \quad (37)$$

$$[\hat{A},\hat{B}]_n \equiv [[\hat{A},\hat{B}]_{n-1},\hat{B}], \qquad (n = 1, 2, \ldots). \quad (38)$$

$[\hat{A},\hat{B}]_1$ is just the regular commutator $[\hat{A},\hat{B}]$. Note that if $[\hat{A},\hat{B}]$ commutes with $\hat{B}$ (which is the case, e.g., if this commutator is just a number, not an operator), the Baker-Hausdorff formula reduces to

$$e^{-\hat{B}}\hat{A}\,e^{\hat{B}} = \hat{A} + [\hat{A},\hat{B}]. \quad (39)$$

To prove (36) we consider

$$\hat{Y}(s) \equiv e^{-s\hat{B}}\hat{A}\,e^{s\hat{B}} \quad (40)$$

where s is a real number. The lhs of (36) is then $\hat{Y}(1)$. Let us calculate $\hat{Y}(s)$ using its Taylor series expansion around $s = 0$:

$$\hat{Y}(s) = \sum_{n=0}^{\infty}\frac{1}{n!}\frac{d^n\hat{Y}(s)}{ds^n}\Big|_{s=0} s^n. \quad (41)$$

[5] Referring to our result as Noether's theorem may be a slight abuse of terminology, since the original Noether's theorem is a theorem in classical mechanics/classical field theory. The result derived here should therefore perhaps be referred to as a quantum version/analogue of Noether's theorem. Note also that the classical Noether's theorem makes use of the Lagrangian formulation, while in our discussion for the quantum case it is the invariance of the Hamiltonian that enters.

Here

$$\frac{d^n\hat{Y}(s)}{ds^n} = e^{-s\hat{B}}[\hat{A},\hat{B}]_n e^{s\hat{B}} \tag{42}$$

which can be proved by induction. It is true for $n = 0$, since $d^0\hat{Y}(s)/ds^0 = \hat{Y}(s) = e^{-s\hat{B}}\hat{A}\,e^{s\hat{B}} = e^{-s\hat{B}}[\hat{A},\hat{B}]_0 e^{s\hat{B}}$. Assuming it is true for an arbitrary nonnegative integer n, we get

$$\begin{aligned}\frac{d^{n+1}\hat{Y}(s)}{ds^{n+1}} &= \frac{d}{ds}\frac{d^n\hat{Y}(s)}{ds^n} = \frac{d}{ds}e^{-s\hat{B}}[\hat{A},\hat{B}]_n e^{s\hat{B}} \\ &= e^{-s\hat{B}}(-\hat{B}[\hat{A},\hat{B}]_n + [\hat{A},\hat{B}]_n\hat{B})e^{s\hat{B}} = e^{-s\hat{B}}[\hat{A},\hat{B}]_{n+1}e^{s\hat{B}},\end{aligned} \tag{43}$$

which is indeed of the form (42) with n replaced by $n+1$. This concludes the induction proof. Inserting (42) into (41) gives $\hat{Y}(s) = \sum_{n=0}^{\infty}\frac{1}{n!}[\hat{A},\hat{B}]_n s^n$. Setting $s = 1$ then gives (36). QED.