Notes on quantum mechanics

John O'Neall

Based on

Ballentine + Griffiths + Norsen + Sakurai & Napolitano + Susskind + Teller + Townsend

(Not about QFT, which = applied QM + SR + Lie group theory)

Note to the reader (assuming there is one...).

This document is just an overview of the subject, composed essentially of notes I have made to remind myself of the math behind the main principles of quantum mechanics. I should say, the mathematical and physical concepts which *are* the main principles of QM. No attempt to teach anything or to indicate any real calculations of cross sections or whatever. It presents, for me, the mathematics of physics, the *concepts* you need to know in order to understand the principles presented in simpler books, such as for laymen, no more.

Table of Contents

1. Introduction – Hilbert spaces

It is often said that math is the language of physics. This is true, but math alone is not enough. We also require a way to understand what the math means, what its symbols correspond to in the physical world.

Tiny nutshell version: Quantum mechanics (henceforth QM) is composed of

- basic physical concepts (particles, waves, energy, etc.);
- a mathematical formalism defining objects and rules for their manipulation; and
- a set of rules for identifying the correspondences between physical entities and the mathematical objects.

Only when we possess and understand these tools are we able to draw conclusions based on the theory and devise experiments to test them. The same things can be said of classical physics, but QM is much more abstract, as well as often being little intuitive.

The math of QM considers the world to be made up of objects or groups of objects, each of which is referred to as a ***system*** in a certain ***state***. Each state is associated with a set of properties given by ***dynamical variables***. The mathematical objects representing the states are linear ***vectors*** (or rays). The space of variables wherein states of a physical system exist is a mathematical ***Hilbert space***, an abstract construction of a type of space called a ***vector space***, which we will consider in section 3. A Hilbert space is a complete vector space possessing the structure of an inner product. The dimension of the space depends on the structure of the physical system under study and is often infinite. The rules for manipulation of these vectors are the subject of ***linear algebra***, on which QM depends for manipulating vectors in Hilbert space.

Hilbert spaces are abstract, they only exist in mathematicians' imagination, not in the physical world. (At least, that's my take.) We can manipulate them with the well-defined mathematical methods of linear algebra, just as we can abstractly prove theorems of plane geometry without actually ever measuring a triangle with a ruler or extending straight lines to see if they intersect. But we could draw a triangle on a piece of paper and measure it. In order to do calculations based on vectors and operations in Hilbert space, we must do something analogous. We use a ***representation*** of the state vector, an instantiation of the abstract space in terms of something more concrete, something we can measure, although often indirectly, through its effects on something else (an indicator). A representation is expressed in terms of a certain set of basis vectors, The most common QM representations use coordinates or momenta as bases.

Variables are expressed as ***operators*** on the states. Operators transform the state somehow, into another state or the same state multiplied by a generally complex number. In the latter case, the vector is said to be an ***eigenvector*** and the complex multiplier the ***eigenvalue*** of the operator. We will be interested mainly in so-called ***Hermitian*** operators (defined later) which always have real eigenvalues.

Notice that no mention has been made yet about what is observed or measured – an electron, an EM field or a cat – nor any of its properties. That comes in with the definition and interpretation of the state vector. The math of QM says how to manipulate states, not how to define them. Only when we impose the constraints of certain fairly obvious symmetric transformations will we be able to understand what some of the operators represent in terms of physical quantities. Symmetries are of great importance in modern physics.

A physical entity under study can exist in different states and of course it is interesting to be able to predict how those states change when forces, or equivalently, potential energy fields, are applied -- or simply under the passage of time. Such changes of state due to forces or time are expressed mathematically by an equation, the ***Schrödinger equation***.

We will discuss only non-relativistic QM (NRQM). In other words, we ignore Lorentz invariance and consider only Galilean symmetries. Relativistic QM (RQM) exists also, being better known as quantum field theory (QFT).

But one thing at a time.

2. Quick reviews

We will be using lots of Lagrangian and Hamiltonian theory. It's good to know a bit of group theory.

2.1. Action, the Lagrangian and the Euler-Lagrange equations

QM needs a method of attacking mechanics problems not through coordinates, but through energies. These ideas were developed during the mid-eighteenth century, principally by Euler and Lagrange.

The ***action*** is defined as

$$A = \int_{t_0}^{t_1} L dt$$

where the ***Lagrangian*** is most simply (at least in classical physics) defined as

$$L = T - V,$$

the difference of the kinetic and potential energies of the system in question. The principle of stationary (often, least) action then says:

$$\delta A = \delta \int_{t_0}^{t_1} L(q_i, \dot{q}_i) dt = 0,$$

which leads through use of incremental differences and the product rule to the ***Euler-Lagrange equation***

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) = \frac{\partial L}{\partial q_i}, \tag{2.1}$$

q_i being a coordinate direction in the Hilbert space. The ***canonical momentum*** is defined by

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.2}$$

which, along with the Euler-Lagrange equation, gives

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i}. \tag{2.3}$$

Very Important Result: If the Lagrangian is invariant under translations, i.e., independent of q_i, *canonical* momentum does not change over time; it is *conserved.*

2.2. Energy and the Hamiltonian

The next step in the use of energy was taken by William Rowan Hamilton around 1830.

A different way of approaching the subject is to define the ***Hamiltonian*** as a (negative) Legendre transform of the Lagrangian

$$H = \sum_i (p_i \dot{q}_i) - L. \tag{2.4}$$

Then we have

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}. \tag{2.5}$$

So H varies in time only if the Lagrangian has a specific time dependence. Equivalently, if a system (Lagrangian) is invariant under time translation, then H is conserved.

In the simple case where the Lagrangian is

$$L = \frac{m}{2}\dot{q}^2 - V(q), \tag{2.6}$$

with canonical momentum $p = m\dot{q}$, the Hamiltonian is

$$H = (m\dot{q})\dot{q} - \frac{m}{2}\dot{q}^2 + V(q) = \frac{m}{2}\dot{q}^2 + V(q) = T + V, \tag{2.7}$$

the total energy, the sum of the kinetic and potential energies. Although this equation is only true if the Lagrangian has this simple form, the Hamiltonian is always the energy. It is also the basis for the Hamiltonian formulation of mechanics – and, especially – non-relativistic quantum mechanics. Equation (2.5) tells us that if a system (represented by its Lagrangian) is invariant under time translation, energy is conserved.

The Hamiltonian is generally taken in relation to phase space of positions and momenta and is given by

$$H = \frac{p^2}{2m} + V(q) \tag{2.8}$$

for a single particle in potential energy $V(q)$. Then Newton's second law $F_i = m\ddot{x} = -\frac{dV}{dq_i}$ may be expressed as

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{2.9}$$

and

$$\dot{q}_i = \frac{\partial H}{\partial p_i}. \tag{2.10}$$

These two are ***Hamilton's equations***. They can be derived more rigorously, without supposing the simple form (2.8) for the Hamiltonian, by expanding equation (2.4) to

$$\delta H = \sum_i \Big(p_i \delta\dot{q}_i + \dot{q}_i \delta p_i - \frac{\partial L}{\partial q_i}\delta q_i - \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i \Big).$$

By the definition of the canonical momentum, equation (2.2), the first and last terms cancel. Comparing the result to a standard derivative

$$\delta H(q,p) = \sum_i \Big(\frac{\partial H}{\partial p_i}\delta p_i + \frac{\partial H}{\partial q_i}\delta q_i \Big)$$

and using equation (2.3) leads to Hamilton's equations.

2.3. A brief foray into group theory

A ***group*** is a collection of objects and an operation such that the collection is complete (closed) when the operation acts on the objects. Operating on objects leads to other objects in the group. It must also be associative, and possess both an identity and an inverse.

Physicists employ Lie groups, a ***Lie group*** being a continuous transformation group. Its elements may be transformations such as rotations or translations but not reflections, as these are not continuous. Such groups are abstract mathematical entities, but they may be instantiated for study in ***representations*** with an arbitrary number of dimensions. Since the Lie group and therefore the representation are continuous, the latter can be generated incrementally from the identity operator by using operators called ***generators***. The generators therefore are derivatives of Lie group elements in a representation and form a vector space called a ***Lie algebra***. The generators are especially interesting to physicists because they represent physical observables such as energy or linear or angular momentum. They thereby provide links between transformations and observables, allowing us to describe changes in the physical systems under study. In this way, we can deduce properties of

the world around us from properties of mathematical groups.

Several variables and quantities are associated with Lie groups and their representations. The number of parameters of the group, called its ***order***, is fundamental and Is constant across all representations of the group. The order of an SU(n) group is $n^2 - 1$, 8 for SU(3) or 3 for SU(2); of an SO(n) group $\frac{n(n-1)}{2}$, 3 for SO(2).

The subset of the the generators which commute with each other and are mutually diagonalizable (defined later) are the ***Cartan generators***. Their number, called the ***rank*** of the group, is also constant over all representations.

The eigenvectors of the Cartan generators span the space of a representation and therefore are of the same dimension as the representation. The eigenvalues can be used to label the corresponding states. ***Casimir operators*** commute with every generator and may be used to label representations.

In brief (finally!),

- group: order = dimension = number of parameters, constant across all representations;
- representation = arbitrary dimension, but number of generators = order, so constant;
- Cartan (number = rank) and Casimir operators each in constant number across all representations.

2.4. Translation and canonical momentum

An important example of transformation symmetry is translation. The translation operator $\hat{T}$ has the following action on a state vector, represented by a Dirac ket $|x\rangle$:

$$\hat{T}(\delta x)|x\rangle = |x + \delta x\rangle, \tag{2.11}$$

which for an infinitesimal translation δx is expressed in terms of its generator $\hat{p}_x$ by

$$\hat{T}(\delta x) = 1 - \tfrac{i}{\hbar}\hat{p}_x \delta x. \tag{2.12}$$

We will consider this to be the *definition* of this operator $\hat{p}_x$, which we have not yet identified, although we know it from many sources (and will study in section 5) to be the operator for the momentum in the x direction. Then a finite translation through a can be found by doing an infinite number of infinitesimal operations, to get

$$\hat{T}(a) = e^{-i\hat{p}_x a/\hbar}. \tag{2.13}$$

Going into the position basis (coordinate representation; more later on that) and using (2.11) and (2.12) plus a Taylor expansion of $\psi(x' - \delta x)$ leads to[1]

$$\langle x|\hat{p}_x|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle,$$

identifying the coordinate-representation momentum operator as

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \quad \text{or} \qquad \hat{p} = -i\hbar\nabla. \tag{2.14}$$

The canonical momentum thus is identified as the generator of translations In a system of constant or zero potential and is identified in the coordinate representation by (2.14).

3. Vector spaces, eigenvectors and all that

To partially repeat what was said in the introduction, in QM, the space wherein states of a physical system exist is a mathematical ***Hilbert space***, an abstract construction of a type of space called a ***vector space***. A Hilbert space is a complete vector space possessing the structure of an inner product.[2] It may have any number of dimensions, often an infinite number. The dimension of the space depends on the structure of the physical

1 Townsend, 157.
2 Or, a Hilbert space is an inner-product space which is complete. Jeevanjee, 37.

system under study.

A linear vector space must obey a number of rules. The sum of any two vectors is a vector, and addition is commutative and associative. There exists a zero vector and, for every vector, an inverse vector. The product of a scalar and a vector is a vector and therefore so is a ***linear combination*** of vectors, a sum of such products. Scalar addition is associative and scalar multiplication is distributive and associative.

A vector is ***linearly independent*** of a set of vectors if it can not be represented as a linear combination of vectors in the set. Each vector in a set of linearly independent vectors is independent of all the other members of the set. If every vector in the space can be expressed as a linear combination of vectors in the set, the set Is said to be ***complete*** and to ***span*** the state. Such a set of linearly independent vectors which span the space is called a ***basis***. (Think of $\hat{x}, \hat{y}, \hat{z}$ or $\hat{1}, \hat{j}, \hat{k}$ in 3d Euclidean space.) Since they are linearly independent their number is minimal and is in fact the dimension of the space.

In physics, the basis vectors of such a state may be discrete or continuous. In the latter case, we are interested in functions which are square-integrable over a specified interval, a to b, a space mathematicians call L^2([a,b]).[3]

Vectors need not be intervals or little arrows in the vector space. As soon as a set of functions obeys the above requirements for a vector space, it may be considered a set of vectors. This is true for the functions of physics.

The ***state*** of a system is then represented by its state vector, a unit (normalized) vector in the vector space of states. A physical ***observable*** is described by a linear ***operator***. An operator effects some change on the state vector, even if it is just to multiply it by a number. Not all operators represent observables.

Using Dirac's bra-ket notation, a state vector is a ***ket*** represented by $|A\rangle$. A bra $\langle B|$ is a member of the ***dual space*** of complex conjugates of the bras, used to form inner or outer products, (3.1) or (4.18).

Like good vectors in any Hilbert space, bras and kets can form an ***inner product*** denoted by

$$\langle A|B\rangle, \text{ with } \langle B|A\rangle = \langle A|B\rangle^* \tag{3.1}$$

being its inverted complex conjugate. The Dirac notation is conveniently the same whether the basis be discrete or continuous. In the discrete case, a ket vector may be represented as a sum of basis vectors, in which case

$$|A\rangle = \sum_i a_i|\lambda_i\rangle \text{ and } |B\rangle = \sum_i b_i|\lambda_i\rangle$$

for complex a_i and b_i, assuming the $|\lambda_i\rangle$ are orthonormal basis vectors. Then the inner product is given by

$$\langle A|B\rangle = \sum_i {a_i}^* b_i.$$

The important operators in QM are Hermitian operators. The ***Hermitian conjugat***e $T^\dagger$ of a linear operator (transformation) T is defined by

$$\langle \phi|T\psi\rangle = \langle T^\dagger\phi|\psi\rangle \tag{3.2}$$

which in discrete matrix form is the complex conjugate of the transposed (inverted) matrix

$$T^\dagger_{ij} = T^*_{ji}.$$

A ***Hermitian operato***r is equal to its Hermitian conjugate

$$\langle \phi|\hat{T}\psi\rangle = \langle \hat{T}\phi|\psi\rangle \quad \text{so by (3.2)} \qquad \hat{T}^\dagger = \hat{T}. \tag{3.3}$$

When an observable is continuous, say x, the state is itself a function of continuous variables. The bra $\langle\psi|$ corresponding to a ket $|\psi\rangle$ represented by the function $\psi(x)$ is the complex conjugate $\psi^*(x)$. In these equations, $\psi(x)$is a complex-valued function of an independent real variable x. Then

3 Jeevanjee, 14.

$$\sum_i \rightarrow \int_i dx$$

and the inner product is defined by

$$\langle\psi|\phi\rangle = \int_{-\infty}^{\infty} \psi^*(x)\phi(x)dx.$$

Normalization of probability then requires that

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx \ = 1.$$

Although this integral may be infinite (and so not normalizable), it may nevertheless be useful over a limited range (a,b) of x, so the space is still L^2([a,b]). This means integrating from a to b rather than from $-\infty$ to ∞ in the above equation.

The ψ - ϕ component of an operator T is

$$\langle\psi|T\phi\rangle = \int \psi^*(x)T\phi(x)dx.$$

If an operator M acting on a vector $|\psi\rangle$ just multiplies it by some value

$$M|\psi\rangle = \lambda|\psi\rangle, \tag{3.4}$$

then λ is an ***eigenvalue*** of M and $|\psi\rangle$ is the corresponding ***eigenvector***.

In the discrete case, it is often convenient to represent the bra and the ket as matrices. So rather than looking at all of linear algebra, let's warm up by taking a look at those. Remember, it's all about state vectors and operators for variables.

EXAMPLE. As an example which will illustrate some of the techniques for dealing with these beasts, consider the following operator matrix.[4]

$$M = \begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix}. \tag{3.5}$$

One can equally well represent the eigen-equation (3.4) by $(M - \lambda I)|\psi\rangle = 0$. In order to calculate the eigenvalues of the observable represented by this matrix, we calculate the ***characteristic equation*** (or characteristic polynomial), which is obtained from the determinant

$$det(M - \lambda I) = \begin{vmatrix} (2-\lambda) & 0 & -2 \\ -2i & (i-\lambda) & 2i \\ 1 & 0 & (-1-\lambda) \end{vmatrix} = -\lambda^3 + (1+i)\lambda^2 - i\lambda = 0.$$

This equation has λ in every term, so obviously 0 is one root. It's easy to see that 1 is another and only slightly trickier to see that i is the third. Notice that if we take the negative of the characteristic equation, then it can be factored in such a way as to make clear the eigenvalues.

$$\lambda^3 - (1+i)\lambda^2 + i\lambda = 0 = (\lambda - 0)(\lambda - 1)(\lambda - i).$$

We have three eigenvalues and therefore three corresponding eigenvectors:

$$M|\psi^{(i)}\rangle = \lambda_i|\psi^{(i)}\rangle, i = 1, 2, 3 \qquad \text{for} \qquad \lambda_i = (0, 1, i). \tag{3.6}$$

For $i = 1$ and eigenvalue $\lambda_1 = 0$, assuming a column eigenvector $|\psi_i^{(1)}\rangle = a_i$ this gives three equations

4 Example borrowed from Griffths & Schroeder, 476-479.

$$2a_1 - 2a_3 = 0, \qquad -2ia_1 + ia_2 + 2ia_3 = 0, \qquad a_1 - a_3 = 0.$$

Clearly, $a_1 = a_3$, but neither is completely determined. So let's take $a_1 = 1 = a_3$, which means $a_2 = 0$ and the 1st eigenvector is

$$|\psi^{(1)}\rangle = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \text{ for eigenvalue 0.} \tag{3.7}$$

Note that the arbitrary choice of $a_1 = 1$ just imposed a multiplicative factor on the eigenvector, which amounts to a phase and has no importance for observables, which depend on the absolute square of the eigenvector (section 4.1). Similarly, we can find the other two eigenvectors (again, each one up to a phase factor)

$$|\psi^{(2)}\rangle = \begin{pmatrix} 2 \\ 1-i \\ 1 \end{pmatrix} \quad \text{and} \quad |\psi^{(3)}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{3.8}$$

for eigenvalues 1 and i, respectively.

In this example, the eigenvectors span the space, so we can use them as a basis. Imagine basis vectors $|f_i\rangle = |\psi^{(i)}\rangle$, then $M|f_i\rangle = \lambda_i|f_i\rangle$ and

$$\langle f_j|M|f_i\rangle = \lambda_i\langle f_j|f_i\rangle = \lambda_i\delta_{ij}$$

is diagonal

$$M = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{pmatrix}, \tag{3.9}$$

with normalized eigenvectors

$$|f_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |f_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |f_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.10}$$

Since matrix (3.9) is diagonal, the operator matrix M is said to be ***diagonalizable***. Very Important Point: The diagonalization is possible *if and only if its eigenvectors span the space*.[5]

A ***unitary*** transformation is one whose inverse is equal to its Hermitian conjugate, so that

$$U^\dagger = U^{-1} \quad \text{and} \quad U^\dagger U = I. \tag{3.11}$$

If we transform a state vector by such a ***unitary*** transformation

$$|\psi\rangle \to |\psi'\rangle = U|\psi\rangle, \tag{3.12}$$

then the operators for observables must also transform, so if

$$A|\phi_n\rangle = a_n\phi_n, \qquad \text{then} \quad A'|\phi'_n\rangle = a_n|\phi'_n\rangle.$$

The second equation says

$$A'U|\phi_n\rangle = a_nU|\phi_n\rangle \text{ so} \qquad U^-A'U|\phi_n\rangle = a_n|\phi_n\rangle = A|\phi_n\rangle$$

and the observable operator transforms as

$$A \to A' = UAU^-. \tag{3.13}$$

Two such matrices as A and A' related by a for a non-singular matrix U, are said to be ***similar*** matrices.[6]

5 It is this requirement which leads Ballentine to consider rigged Hilbert spaces for the state vectors.
6 Griffiths and Schroeder, 474.

Similar matrices have two interesting properties: Their *determinants* are the same, as are their *traces*, as can be confirmed in the example by comparing (3.5) and (3.9). Equation (3.13) tells us that matrices which represent the same linear transformation in different bases are similar. Transformation of an operator matrix into a similar, diagonal matrix is possible (if and) only if its eigenvectors span the space and form a ***complete set***.

We can do the diagonalization another way. The similarity matrix to do the job can be constructed from the eigenvectors in the original basis:

$$(S^{-1})_{ij} = |\psi^{(j)})_i \tag{3.14}$$

In our case this gives

$$S^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & (1-i) & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{3.15}$$

In fact, this amounts to using the original column eigenvectors from (3.7) and (3.8) as columns in the operator matrix. The inverse of a matrix with non-zero determinant is given by[7]

$$T^{-1} = \frac{1}{det\ T}\widetilde{C}, \tag{3.16}$$

where $\widetilde{C}$ is the inverse of the ***cofactor matrix***, whose ij[th] element is $(-1)^{i+j}$ times the submatrix obtained by ignoring the i[th] row and j[th] column. With much sweat and pencil usage[8], we can calculate the inverse of this inverse as

$$S = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ (i-1) & 1 & (1-i) \end{pmatrix}. \tag{3.17}$$

Using this, it really is straightforward to show that

$$S|\psi^{(j)}\rangle = |\psi^{(j)\prime}\rangle = |f_i\rangle$$

of (3.10).

But… As is clear from (3.16), the existence of an inverse to a matrix depends on its having a non-zero determinate. Remember that similar (in the math sense) matrices have identical determinants, which is handy. And, as already stated, transformation of an operator matrix into a similar, diagonal matrix is possible if and only if its eigenvectors span the space, i.e.. if they form a complete set. More coming...

4. QM basics

I will try to present a simplified consensus based on five sources, given in the bibliography: Ballentine, Griffiths and Schroeder, Sakurai and Napolitano, Susskind, and the Physics Forums web site.

4.1. Basic principles

Principle 1. A pure ***state*** of a quantum system is described by a ***vector*** $|\psi\rangle$ in Hilbert space.

(An alternative formulation defines the state as an *operator* rather than a vector. See below, section 4.3.)

Principle 2. A physical ***observable*** is represented by a ***Hermitian operator*** on a state vector.[9] The possible

7 Griffiths & Schroeder, 472.

8 My notes J-III, 27,

9 Physics forums, 7 rules. Equivalently, A is ***self-adjoint***. According to Ballentine, a self-adjoint operator is a Hermitian operator whose domain is the same as that of its Hermitian conjugate. He's most liekly thinking in terms of ensembles of

results of measurements of the observable are the ***eigenvalues*** of the Hermitian operator representing the observable, as in (3.4). There may be multiple eigenstates, in which case the set of eigenvalues is called the ***spectrum*** of the operator. Then the possible values of a measurement of an observable are said to be the ***spectral values*** of the corresponding operator A.

The operator's being Hermitian has several consequences.

- The eigenvalues of a Hermitian operator are ***real***.
- Eigenvectors corresponding to distinguishable eigenvalues of a Hermitian operator must be ***orthogonal***. (These two properties, reality of eigenvalues and orthogonality of eigenvectors, constitute Sakurai's Theorem 1.[10])
- In the case of discrete eigenvalues, the eigenvectors of an operator form a ***complete set*** and so span the space of the representation. This means that any vector in the space can be represented as a weighted linear sum of eigenvectors. They can be chosen so as to form an ***orthonormal basis*** of the space.

In the case of continuous (or an infinite number) of eigenvalues/eigenvectors, this becomes problematic. One solution is to extend the notion of Hilbert space to so-called ***rigged Hilbert space***,[11] but we won't go there.[12]

Assuming the eigenvectors of an operator form a complete set, an arbitrary vector $|\alpha\rangle$ can be expressed as a linear sum, or ***superposition***, of eigenvectors

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle, \tag{4.1}$$

where $A|a'\rangle = a'|a'\rangle$, so the complex coefficients are inner products

$$c_{a'} = \langle a'|\alpha\rangle. \tag{4.2}$$

Therefore

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle, \text{ and } \quad \sum_{a'} |a'\rangle\langle a'| = I, \tag{4.3}$$

the identity operator. This equation is the ***completeness relation*** and expresses ***closure***. We will use it a lot in the following pages.

A state which is not pure also may look like equation (4.1) or even something more complicated. We then say it is in a ***mixed*** state.

The set of inner-product coefficients $c_{a'}$ of (4.2) constitutes the ***wave function*** of the system in the basis defined by the set of observables a'.[13] The wave function is the projection of the state vector onto the eigenvectors of the representation, as (4.2) is the projection of state $|\alpha\rangle$ onto the direction $|a'\rangle$. You could consider the set of wave functions to be the coordinates of the state vector in the basis defined by the eigenvectors.

More generally, if a basis of orthonormal state vectors is represented by $|a, b, c, ...\rangle$ for eigenvalues (a,b,c,...), then any state vector may be expanded as

$$|\Psi\rangle = \sum_{a,b,c,\ldots} \psi(a, b, c, ...)|a, b, c, ...\rangle \tag{4.4}$$

where the *set of coefficients* $\psi(a, b, c, ...)$ is called the ***wave function*** of the system in the basis defined by the

events.

10 Sakurai and Napolitano, 16.

11 Ballentine does, 27-28.

12 Leonard Susskind just states completeness as an axiom and Sakurai, 12, assumes it at the outset.

13 Susskind and Friedman, QM, 134. Some authors confuse wave function and state vector. Griffiths, 114, restricts the definition to the position representation.

observables, A, B, C, ….[14] We can expand $|\Psi\rangle$in terms of another set of basis vectors corresponding to different observables and the wave function set will be different even though they describe the same system state. The two sets of basis vectors correspond to different ***representations*** of the system, for the two different observables. We will return to this subject in section 6.2.

The number of eigenvectors in the basis is the dimension of the Hilbert space, or at least that part of it corresponding to this observable. (Think of a fermion's angular momentum in 3-d space and its spin, in 2-d spinor space.) So the dimension of the state is greater than or equal to the number of eigenvalues. If the eigenvalues are continuous, the dimension is infinite and this may apply to many or all observables.

Two or more orthogonal eigenvectors may have the same eigenvalue, in which case they are referred to as ***degenerate*** states. Even then, it is possible to construct a set of orthonormal basis vectors.

In case of degenerate subspaces, let $\{|a,\nu\rangle\}$ be a complete set of (generalized) eigenvectors of A, indexed by ν, where more than one state ν may correspond to the same eigenvalue a . The probability $p_\psi(a)$ to find the measured value a is then given by summing (or integrating) over ν, i.e. over the entire a-subspace, the (degenerate) states with eigenvalue a:

$$p_\psi(a) = \sum_\nu |\langle a,\nu|\psi\rangle|^2.$$

Principle 3. The time evolution of an isolated quantum system represented by the state vector $|\psi(t)\rangle$ is given by

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \tag{4.5}$$

where H is the Hamilton operator and $\hbar$ is Planck's constant. This is the time-dependent ***Schrödinger equation***.

This equation is valid in the formulation of quantum mechanics called the Schrödinger picture. There are other, equivalent formulations of the time evolution, especially the Heisenberg picture and the Dirac (interaction) picture, where time evolution is entirely or partially shifted from the state vector to the operators.

It is equivalent to the Schrödinger equation (4.5) to define the time evolution of an isolated quantum system by

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \tag{4.6}$$

with the unitary time evolution operator

$$U(t) = e^{-iHt/\hbar}. \tag{4.7}$$

The evolution according to (4.7) is therefore also referred to as ***unitary evolution***.

Principle 4. For a normalized state vector $|\psi\rangle$ and observable A, the probability of observing value λ_i by measurement is given by the ***Born rule***

$$P(\lambda_i) = \langle\psi|\lambda_i\rangle\langle\lambda_i|\psi\rangle = |\langle\psi|\lambda_i\rangle|^2. \tag{4.8}$$

In an experiment, say scattering, a particle is prepared (accelerated and collimated) into a particular state of motion, i.e., momentum and position. A subsequent measurement, after scattering, will give a result which is not precisely repeatable, be it for quantum-mechanical or experimental reasons.

Very Important Point: A specific preparation does not determine the result of the following measurement, but only the probability of the results.

A ***preparation*** of a state is just that, a preparation, and is *independent* of whatever measurement may or may not follow. Therefore the ***state*** must determine the ***probability distribution*** of all possible observables[15], an ***observable*** being a measurable (at least in principle) ***dynamical variable***. (This goal will be accomplished by

14 Susskind's notation. Griffiths uses a different definition of the wave function, but this one keeps it separate from the state vector.

15 Ballentine, 45.

using the definition of the state vector as an operator in up-coming section 4.3.)

We can see this calculated probability not as the probable value of a specific event but as the average outcome over an ***ensemble*** of similarly prepared events, the ensemble being "… the virtual unbounded set of similarly prepared systems."[16] According to Ballentine, the chief proponent of this interpretation, this direct identification of a state with a set of probability distributions should make it possible to avoid many objections to QM.[17]

Principle 5. After measurement of an observable in a ***superposition*** of discrete states, as in (4.1), the state vector is no longer a superposition but a single eigenvector $|a\rangle$ of measured eigenvalue a. A second measurement of the state gives the same result as the first. The first observation serves therefore to prepare a state for input to a second observation. An example would be a filter applied to the output of a Stern-Gerlach apparatus.

What happened to all those initial states? It would seem that some process has led from the coherent superposition of macroscopically distinct eigenvectors of the ***indicator*** (the macroscopic variable read off the measurement apparatus) to a single "reduced" state. The standard (Copenhagen) interpretation says that the wave function has "collapsed" to a single state. But what does "collapse" mean? This is the so-called ***measurement problem*** of QM and it's a biggie.

There are many proposals for interpreting what happens when measurement takes place. The ***instrumentalist*** approach, for instance, interprets the wave function as merely an instrument to calculate probabilities, without having any real meaning in itself. This is opposed to the ***realist*** approach, which interprets the wave functions as a thing, a new physical entity or field. There exist nuances of both these ideas. And we haven't mentioned many worlds or pilot waves or other suggestions.

Consider this though. A superposition is not a thing in itself, it is always a superposition of some thing or things. If we have, for instance

$$|\Psi_3\rangle = \tfrac{1}{\sqrt{2}}(|\Psi_1\rangle + |\Psi_2\rangle) \qquad \text{and}$$
$$|\Psi_4\rangle = \tfrac{1}{\sqrt{2}}(|\Psi_1\rangle - |\Psi_2\rangle),$$

then we can also write

$$|\Psi_1\rangle = \tfrac{1}{\sqrt{2}}(|\Psi_3\rangle + |\Psi_4\rangle) \qquad \text{and}$$
$$|\Psi_3\rangle = \tfrac{1}{\sqrt{2}}(|\Psi_3\rangle - |\Psi_4\rangle).$$

It seems difficult to accept all four wave functions as representing real entities.[18] But they do, as shown by succeeding Stern-Gerlach devices with a filter in between. The filtered z-spin up is a superposition of x-spin left or right, either of which is a superposition of z-spin up and down … and so on.

4.2. Multiple observables

If the eigenvectors of an observable are used to span the space, not only are the eigenvalues real and the eigenvectors orthogonal, but the matrix for the hermitian operator of the observable is ***diagonal***, as we saw in equation (3.9) of the example. If two different observables commute, so that

$$[A, B] = 0, \tag{4.9}$$

they are said to be ***compatible***. If the eigenvectors of A are non-degenerate and are used as basis vectors, the matrices for both observables, A and B, are diagonal. (Sakurai's Theorem 2.[19]) The eigenvectors are

16 Ballentine, 47.
17 He says we will get to the point of this in chapter 9.
18 Thanks to Sabine Hossenfelder for this idea. Understanding Quantum Mechanics #2: Superposition and entanglement. https://www.youtube.com/watch?v=j6Mw3_tOcNI
19 Sakurai and Napolitano, 28. Their more nuanced treatment takes degeneracy into account.

simultaneous eigenvectors of A and B. But if two observables, A and B, are not compatible, then they do not have a complete set of simultaneous eigenvectors.

In the case of two non-compatible observables A and B, we may know the eigenvectors and eigenvalues of A

$$A|a'\rangle = a'|a'\rangle, \tag{4.10}$$

as well as the operator B in the a' basis. We may also want to know the eigenvalues and vectors for B such that

$$B|b'\rangle = b'|b'\rangle. \tag{4.11}$$

It is easy to show that if each has a complete set of orthonormal kets, then there exists a unitary operator U which transforms from one basis to the other (Sakurai's Theorem 3.[20]):

$$|b^{(i)}\rangle = U|a^{(i)}\rangle, \tag{4.12}$$

so that the transformation matrix is given simply by the inner product of the original bra and the transformed ket

$$\langle a^{(i)}|U|a^{(j)}\rangle = \langle a^{(i)}|b^{(j)}\rangle. \tag{4.13}$$

With appropriate matrix summations, this becomes

$$\langle a''|B|b'\rangle = \sum_{a'} \langle a''|B|a'\rangle\langle a'|b'\rangle = b'\langle a''|a'\rangle,$$

which we recognize as a matrix equation for the elements of the matrix of B. As in the preceding example (3.5), this can be solved for the eigenvalues by using the characteristic equation

$$det(B - \lambda) = 0.$$

We then can use the eigenvalues b' to solve for the eigenvectors $|b'\rangle$ (up to a multiplicative constant), which is what we wanted.

Since by (4.11) B is diagonalized in $\{|b'\rangle\}$, U is the unitary matrix which diagonalizes it from $\{|a'\rangle\}$. So finding the eigenvalues and eigenvectors of operator B is equivalent to finding the unitary matrix which diagonalizes it.

But there's more. Knowing the unitary operator U which diagonalizes B, we use it to transform A by the unitary transform UAU^{-1}, what we called earlier a ***similar*** matrix. A and UAU^{-1} are called ***unitary equivalent observables***. From the eigenvalue equation (4.10) for A, we then see that

$$UAU^{-1}U|a^{(l)}\rangle = a^{(l)}U|a^{(l)}\rangle = a^{(l)}|b^{(l)}\rangle$$

by (4.12). This is just the eigenvalue equation for UAU^{-1} and tells that operator has the same spectrum of eigenvalues in $\{|b'\rangle\}$ as in $\{|a'\rangle\}$. Unitary equivalent observables have identical spectra. (Sakurai's Theorem 4.[21])

As an example, consider spin components S_x and S_z, spin in the x and z directions. Rotating the latter by an angle of $\pi/2$ about the y-axis gives the former. Rotation is a unitary transformation, so S_x and S_z are unitary equivalent observables and have identical eigenvalues. And they do: $+\hbar/2$ and $-\hbar/2$.

On the other hand, in the case of two operators A and B which do *not* commute, that fact plus some geometry (the Cauchy-Schwarz inequality) leads to the ***generalized uncertainty principle***:[22]

$$\Delta A \Delta B \geq \frac{1}{2}\langle\Psi|[A, B]|\Psi\rangle. \tag{4.14}$$

This result is not a separate postulate, but a consequence of quantum-mechanical non-commutivity. The original ***Ungenauigkeitsprinzip*** would have been better and more accurately translated as "inexactitude principle" or "imprecision principle" and could have avoided overly hasty interpretations, but... The result can also be shown to be

20 Sakurai and Napolitano, 33.
21 Sakurai and Napolitano, 37.
22 Susskind and Friedman, 147.

$$\sigma_a^2 \sigma_b^2 \geq \left(\frac{1}{2i} \langle [A, B] \rangle \right)^2 \tag{4.15}$$

where the standard deviation of an observable is given by

$$\sigma^2 = \langle (A - \langle A \rangle)\Psi | (A - \langle A \rangle)\Psi \rangle. \tag{4.16}$$

This quickly leads to the usual version concerning position and momentum

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

and can be shown to be true also for energy and time

$$\Delta E \Delta t \geq \frac{\hbar}{2}.$$

This is a QM result and has nothing to do with SR four-vectors. “No experiment can ever beat this limitation.”[23]

We can understand this qualitatively. We will see in section 5.4 that the QM operators for observables are the generators for the corresponding symmetries, so a measurement of momentum is equivalent to the action of a translation generator.[24] So a measurement of momentum moves the system some, thus making the position less certain.

4.3. Alternative view -- the state as operator

There is an alternative way of viewing the state -- as an operator. According to this point of view, to each state there corresponds a unique ***state operator***, also called the ***density operator*** , which must be Hermitian

$$\rho = \rho^\dagger, \tag{4.17}$$

nonnegative, and of unit trace.[25] .

For a ***pure state***, the state operator is defined to be

$$\hat{\rho} = |\psi\rangle\langle\psi|, \tag{4.18}$$

where $|\psi\rangle$ is a state vector and $\langle\psi|$ the corresponding dual vector. It is at once state or density operator and ***outer product***, by definition; and ***projection operator***, by its action. It projects an operator A onto the direction of $|\psi\rangle$:

$$\hat{\rho}|A\rangle = |\psi\rangle\langle\psi| \; |A\rangle = \langle\psi|A\rangle \; |\psi\rangle.$$

If we have a sum of states $|\psi_j\rangle$, then as a i projection operator, it projects out the i^{th} state:

$$|\psi_i\rangle\langle\psi_i|\Psi\rangle = |\psi_i\rangle\langle\psi_i| \sum_j a_j |\psi_j\rangle = a_i |\psi_i\rangle.$$

In an orthonormal basis $\{|e_j\rangle\}$, its matrix elements are

$$\rho_{ij} = \langle e_i|\hat{\rho}|e_j\rangle = \langle e_i|\psi\rangle\langle\psi|e_j\rangle,$$

in which form it is indeed a **state** or ***density matrix***.

The state operator has unit trace:

$$Tr(\rho) = \sum_i \langle e_i|\psi\rangle\langle\psi|e_i\rangle = \sum_i \langle\psi|e_i\rangle\langle e_i|\psi\rangle = \langle\psi|\psi\rangle = 1, \tag{4.19}$$

assuming, of course, that the state vector is normalized. If ρ is a non-negative state operator, then for all u

23 Susskind, 270.
24 Schwichtenerg, PS, 197.
25 Ballentine, 46.

$$\langle u|\rho|u\rangle \geq 1. \tag{4.20}$$

In the pure state – and only in the pure state -- it is easily seen to be ***idempotent***:

$$\rho^2 = |\psi\rangle\langle\psi|\ |\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho.$$

The view further stipulates that if a preparation gives rise to a ***virtual ensemble of events*** represented by the operator $\hat{\rho}$, then the average value of multiple measurements of a dynamical variable A, represented by the operator A is

$$\langle A\rangle = Tr(\rho A), \tag{4.21}$$

This interpretation does not deal with the measurement of one event but an ensemble of many events.[26]

In detail,

$$Tr(\rho A) = Tr(|\psi\rangle\langle\psi|A) = \sum_i \langle e_i|\psi\rangle\langle\psi|A|e_i\rangle = \sum_i \langle\psi|A|e_i\rangle\langle e_i|\psi\rangle = \langle\psi|A|\psi\rangle = \langle A\rangle,$$

since A is Hermitian.

We can take the time differential of the definition of a pure-state ρ in (4.18) and use (4.5) to show that

$$i\hbar\frac{d\hat{\rho}}{dt} = [\hat{H}, \hat{\rho}], \tag{4.22}$$

the ***Schrödinger equation*** in terms of the density operator. Now suppose we have a ***superposition*** of states. We have but to generalize the state vector to

$$\hat{\rho} = \sum_k p_k|\psi_k\rangle\langle\psi_k|, \tag{4.23}$$

where p_k is the probability that the system is in the k-th state. This operator is still Hermitian and has unit trace, and the average value of a dynamic variable is given by (4.21). But it is not idempotent. This is one way to distinguish a pure state from a mixed one.

More on all this in chapter 8.

High points - résumé

- A pure state of a physical system is described by a vector in Hilbert space.
- A physical observable is represented by a Hermitian operator on a state vector. Its eigenvalues are real and constitute the spectral values of the operator.
- Possible values of a measurement of an observable are its eigenvalues,. The eigenvectors are taken to form a complete set, capable of forming an orthonormal basis of the space.
- The time evolution of the state vector is given by the Schrödinger equation in the coordinate representation, or, equivalently, by a unitary time-evolution operator whose generator is the Hamiltonian.
- The probability of observing a value of an observable is given by the Born rule.
- A measurement can leave a super-positioned state in a single, or pure, one, which can constitute the prepared state for another measurement.
- An alternative view of the state vector is to consider the state an operator called the state or density operator, a function of state vectors,.

26 This is one of Ballentine's themes, along with the need of rigged Hilbert spaces for completeness of continuous eigenvectors.

5. Galilean symmetry

We approach this subject through a study of transformations. Transformations in space-time may change both the states and the observables. If they are to correspond to what we observe before us, the laws of nature must be invariant under certain ***symmetry*** operations, namely, displacements in space or time, rotations in space, and changes in uniform relative motion of reference frames (inertial frames, to fans of SR).

In this chapter, we consider only pure states (4.18). Nevertheless, it's fairly slow going, so if you don't want to slog through it, you can skip to the résumé in section 5.5.

5.1. Unitarity

Essential point: Transformations on vector spaces of interest in physics can be represented by unitary operators, and a one-parameter unitary transformation can be written as an exponential in terms of a Hermitian generator.

According to a theorem of Wigner: Any mapping of the vector space onto itself that preserves the value of $|\langle\phi|\psi\rangle|$ may be implemented by an operator U which transforms state vectors such that

$$\begin{aligned} |\psi\rangle &\to |\psi'\rangle = U|\psi\rangle, \\ |\phi\rangle &\to |\phi'\rangle = U|\phi\rangle, \end{aligned} \tag{5.1}$$

with U being either unitary or antiunitary.[27] We are interested especially in the unitary operators, for which

$$UU^\dagger = U^\dagger U = 1,$$

because they preserve the complex values of inner products. The transformation of the *state* vectors must be accompanied by a transformation of the operators for *observables*, such that

$$A|\phi_n\rangle = a_n\phi_n \qquad \text{and} \qquad A'|\phi'_n\rangle = a_n|\phi'_n\rangle,$$

Putting the transformation into the second equation gives

$$A'U|\phi_n\rangle = a_nU|\phi_n\rangle \text{ so} \qquad U^-A'U|\phi_n\rangle = a_n|\phi_n\rangle = A|\phi_n\rangle$$

and finally the operator (observable) transforms under a unitary transformation U as

$$A \to A' = UAU^-. \tag{5.2}$$

In the language of linear algebra, two such matrices A' and A are said to be ***similar*** for the transformation matrix U.

It can be shown that matrices representing the same linear transform with respect to different basis vectors are similar.[28] For a family of unitary operators $U(s)$, functions of a continuous parameter s, we may expand $U(s)$ infinitesimally about $s = 0$ as

$$U(s) = I + \frac{dU}{ds}\Big|_{s=0} s + O(s^2).$$

Now expand $UU^\dagger$ similarly, which gives

$$UU^\dagger = I + s\Big[\frac{dU}{ds} + \frac{dU^\dagger}{ds}\Big]\Big|_{s=0} + O(s^2).$$

Since unitarity requires this to be equal to 1, the s term must be equal to zero and we see that

$$\frac{dU}{ds}\Big|_{s=0} = iK \qquad \text{where} \qquad K = K^\dagger. \tag{5.3}$$

27 Ballentine, 64.
28 Griffiths and Schroeder, 474.

So K is the Hermitian ***generator*** of the unitary operator for any transformation, infinitesimal or finite.[29,30] We can see this by taking $U(0) = 1$ and considering

$$U(s_1 + s_2) = U(s_1)U(s_2). \tag{5.4}$$

Differentiate this with respect to s_2, set $s_2 = 0$ and use (5.3) to get:

$$\frac{dU}{ds}\Big|_{s=s_1} = U(s_1)iK,$$

the solution to which is

$$U(s) = e^{iKs}. \tag{5.5}$$

This shows that any one-parameter group of unitary operators can be written in the form of an exponential including its Hermitian generator.[31]

5.2. Active and passive transformations

We can actively transform a function or passively transform the coordinates in terms of which the function is expressed. It's like either you move it ahead or you take a step back.

The symmetry transformations of space-time can be written as[32]

$$\mathbf{x} \to \mathbf{x}' = R\mathbf{x} + \mathbf{a} + \mathbf{v}t$$
$$t \to t' = t + s. \tag{5.6}$$

This is what Ballentine calls the ***Galilei group*** of transformation and has ten parameters, 3 each for rotation (R), displacement (a) and relative velocity (v), and one for time (t). (Be careful, a is a displacement in space, not an acceleration.) In other books, it is referred to as the Galiliean group and I shall call it that. Remember, this is non-relativistic. Nevertheless, the laws of low-energy, non-relativistic QM must be invariant under these transformations.

We usually use these relations in a representation on coordinate space (chapter 6.1), in which case transformations on function space and on coordinate space are inversely related. By definition of the unitary transformation $U(\tau)$, the transformation of the *function* $\Psi(x)$ relative to a fixed set of coordinates is

$$\Psi'(x) = U(\tau)\Psi(x). \tag{5.7}$$

This is not a change of the coordinate system, which would constitute what is called a ***passive*** point of view. This is the ***active*** point of view, in which the object is transformed relative to a fixed coordinate system. Imagine that it is a translation, for instance. The function would look different because it would now be evaluated at a new point in space. The relation of the new point, x', to the old, x, is given by

$$x' = \tau x,$$

where both points are relative to the same coordinate system and the operator τ simply maps one point into the other within that system.[33] Invariance (symmetry) requires the value of the transformed function at the transformed point be the same as the value of the original function at the original point.[34] This essential relationship appears in the two middle terms of the following equation.

$$\Psi'(x') = \Psi'(\tau x) = \Psi(x) = \Psi(\tau^{-1}x'). \tag{5.8),}$$

According to (5.7)

29 See sections 2.3 and 2.3.
30 Ballentine, 65.
31 Ballentine, 68.
32 Example from Ballentine, 66ff.
33 Physics 221A, Rotations in 3-Dimensional Space. https://bohr.physics.berkeley.edu/classes/221/notes/classrot.pdf
34 Ballentine, 67-68.

$$\Psi'(x') = U(\tau)\Psi(x'),$$

so we can combine the last terms of these two equations and drop the prime to get

$$\boxed{U(\tau)\Psi(x) = \Psi(\tau^{-1}x)}, \tag{5.9}$$

This important equation shows clearly the inverse relation between the function and coordinate transformations. In terms of a translation, e.g., moving the object (function) ahead is equivalent to taking a step backward.

Let's consider an example of active and passive rotations by using simple rotations of a scalar function in a 2d plane,[35]

Consider an *active* counter-clockwise rotation of an object, a point, in the 2-d plane through an angle α. The transformation is given by

$$x_i' = \sum_j R_{ij}^{(a)} x_j,$$

where

$$R_{ij}^{(a)}(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{5.10}$$

and the superscript (a) designates the active transformation. Then the transformed function is given by

$$\Psi'(x) = R^{(a)}(\alpha)\Psi(x) \tag{5.11}$$

By invariance, a new (transformed) scalar function at its new point must be equal to the old function at the old point.

$$\Psi'(x') = \Psi(x) = \Psi([R^{(a)}(\alpha)]^{-1}x'),$$

so using (5.11) and changing from primed to non-primed gives

$$R^{(a)}(\alpha)\Psi(x) = \Psi([R^{(a)}(\alpha)]^{-1}x), \tag{5.12}$$

which is equivalent to (5.9) for this case.

On the other hand, we may consider a passive rotation, leaving the point where it is but rotating the coordinate axes through a clockwise angle α, then the transformation is

$$x_i'' = \sum_j R_{ij}^{(p)} x_j,$$

with

$$R_{ij}^{(p)}(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \tag{5.13}$$

by (5.10). Again, invariance tells us that

$$\Psi''(x'') = \Psi(x) = \Psi([R^{(p)}(\alpha)]^{-1}x''),$$

so passing from double-primed to un-primed

$$R^{(p)}(\alpha)\Psi(x) = \Psi([R^{(p)}(\alpha)]^{-1}x), \quad . \tag{5.14}$$

Both (5.12) and (5.14) are equivalent to (5.9).

This example shows that you can rotate an object through an angle or rotate the coordinate system through the inverse angle and get the same result. Notice that $R^{(a)}(\alpha) := R(\alpha)$ means simply that $R^{(p)}(\alpha) = R(-\alpha)$, as we expect intuitively.

35 Ballentine, 176-178. He does it in 3d, I in only 2d.

5.3. Symmetries of the Galilean group

Since the transformations of the Galilean group form just that, a group, two consecutive Galilean transformations form another. Using this fact, we can calculate the Lie brackets for the group generators.

Starting with equations (5.6),

$$\mathbf{x} \to \mathbf{x}' = R\mathbf{x} + \mathbf{a} + \mathbf{v}t$$
$$t \to t' = t + s.$$

we find the result of a pair of transformations:

$$x'' = R_2(R_1\mathbf{x} + a_1 + \mathbf{v}_1 t) + \mathbf{a}_2 + \mathbf{v}_2(t + s_1)$$
$$t'' = t + s_1 + s_2.$$

Remember, $\mathbf{a}$ is displacement, not acceleration, so

$$R_3 = R_2 R_1$$
$$\mathbf{a}_3 = \mathbf{a}_2 + R_2\mathbf{a}_1 + \mathbf{v}_2 s_1$$
$$\mathbf{v}_3 = \mathbf{v}_2 + R_2\mathbf{v}_1$$
$$s_3 = s_2 + s_1. \tag{5.15}$$

Two points, not considered by Ballentine, who does not mention group theory or Lie algebras.

- The Galilean group is not irreducible, as it contains a subgroup, the group of rotations.
- We are talking here about a representation of the Galilean group in a four-dimensional space-time, where the generators are the members of the group's Lie algebra. Now we need to find the commutators (Lie brackets).

Although we can write the unitary operator, from (5.5), as

$$U(\tau) = \prod_{\mu=1}^{10} e^{is_\mu K_\mu},$$

it is convenient to write the generators of this representation of the Galilean group using more recognizable terms. They are illustrated in Table 1.

Space-time transformation	Unitary operator	Generator
Rotation about axis α $\mathbf{x} \to R_\alpha(\theta_\alpha)\mathbf{x}$	$e^{i\theta_\alpha J_\alpha/\hbar}$	J_α
Displacement along axis α $x_\alpha \to x_\alpha + a_\alpha$	$e^{ia_\alpha P_\alpha/\hbar}$	P_α
Velocity along axis α $x_\alpha \to x_\alpha + v_\alpha t$	$e^{iv_\alpha G_\alpha/\hbar}$	G_α
Time displacement $t \to t + s$	$e^{isH/\hbar}$	H

Table 1. Unitary operators of space-time transformations of the Galilean group, after Ballentine[36]

36 Ballentine, 69.

In the table, the factors of $1/\hbar$ in the operators are not yet derived. I include them here just for eventual correctness. For the current derivation of the sense of the generators, consider $\hbar = 1$.

Since Galilean transformations form a group, the product of any two transformations must be another member of the group and by symmetry must give rise to the same state. But any difference of phase between vectors does not affect the state, so

$$U(\tau_2\tau_1) = e^{i\omega(\tau_2,\tau_1)}U(\tau_2)U(\tau_1), \tag{5.16}$$

where the phase difference ω is a function of τ_1 and τ_2.

For the moment, we only know these objects as generators of transformations. They are operators, not scalars, and so, in general, they do not commute. So we must use the Baker-Campbell-Hausdorf formula for non-commuting exponential terms:[37]

$$e^{(A+B)} = e^A e^B e^{-[A,B]/2} \tag{5.17}$$

The commutation terms on the right of the exponent will lead us to commutation rules (Lie brackets) for the generators. Writing the generators all as K_μ, and considering the product of two pairs of transformations and their inverses leads to

$$e^{i\epsilon K_\mu} e^{i\epsilon K_\nu} e^{-i\epsilon K_\mu} e^{-i\epsilon K_\nu} = I + \epsilon^2[K_\nu, K_\mu] + O(\epsilon^3). \tag{5.18}$$

Reasoning from completeness of the group, we can use (5.5) to write for any given transformation

$$U(\tau) = \prod_{\mu=1}^{10} e^{is_\mu K_\mu} = I + i\sum_{\mu=1}^{10} s_\mu K_\mu. \tag{5.19}$$

for infinitesimally small s_μ. But (5.18) also is a member of the group, so by (5.18) and (5.19), the commutators of the group must obey

$$[K_\mu, K_\nu] = i\sum_{\lambda=1}^{10} c_{\mu\nu}^{\lambda} K_\lambda + ib_{\mu\nu} I\ , \tag{5.20}$$

I being the unit matrix. The last term is the difference of phase from (5.16).

We now use several methods to evaluate the commutation relations.[38] For instance, in the absence of rotations or velocity changes, displacement and time changes in equations (5.15) are independent of each other: $a_3 = a_2 = a_1$ and $s_3 = s_2 + s_1$, which are independent of the order of the transformation τ_1 and τ_2 so their commutators must be zero aside from the phase factor (relations (a) and (g) in Table 2). There are similar results for space displacements and velocity changes ($a_3 = a_2 = a_1$ and $v_3 = v_2 = v_1$), and for rotations and time displacements ($R_3 = R_1R_2$ and $s_3 = s_2 + s_1$), so their commutators also must vanish aside from the phase term (relations (f), (b) and (i)).

Moreover, a rotation commutes with a displacement or velocity displacement along the rotation axis, which gives us (g) and (h) for equal indices.

A negative change of velocity along the x-axis followed by a negative time displacement followed by their inverses expressed as in equation (5.6) yields

$$e^{i\epsilon H} e^{i\epsilon G_1} e^{-i\epsilon H} e^{-i\epsilon G_1} = I + \epsilon^2[G_1, H] + O(\epsilon^3).$$

Working out these transformations using equations (5.3) and comparing the two gives us (h). Relation (c) is of course a result of well-known rotation matrices. Similar calculations based on $J_1 + G_2$ and on $J_1 + P_2$ yield relations (e) and (d).

As for the phase terms, all but one can be eliminated by consistency conditions or by suitable choice of phases

37 Griffiths, QM, 121. This version assumes that each of A and B, although not commuting with each other, do commute with their commutator, $[A, B]$.

38 Details in Ballentine, 66-76.

of certain vectors. The Jacobi identity can be used with triplets J_1, G_2, P_1 to show that $[G_\alpha.G_\beta]$ is diagonal and again with J_1, G_2, P_3 to show that the diagonal elements are identical, hence the result (f), where M is a real, as yet unidentified constant.

Galilean group commutation relations	
(a) $[P_\alpha, P_\beta] = 0$	(f) $[G_\alpha, P_\beta] = i\delta_{\alpha,\beta} MI$
(b) $[G_\alpha, G_\beta] = 0$	(g) $[P_\alpha, H] = 0$
(c) $[J_\alpha, J_\beta] = i\epsilon_{\alpha\beta\gamma} J_\gamma$	(h) $[G_\alpha, H] = IP_\alpha$
(d) $[J_\alpha, P_\beta] = i\epsilon_{\alpha\beta\gamma} P_\gamma$	(i) $[J_\alpha, H] = 0$
(e) $[J_\alpha, G_\beta] = i\epsilon_{\alpha\beta\gamma} G_\gamma$	

Table 2. Commutation relations for the Galilean group of transformations

We have now derived the commutation relations (Lie brackets) for the generators of the Galilean symmetry group. But what do they mean?

5.4. Meaning of Galilean generators

Definition of a position operator and derivation of its commutation relations with the Galilean generators, plus some intuition, leads to the identification of the Galilean generators with physical variables: momentum (translation), angular momentum (rotation) and energy (time evolution). The meaning of the generator of velocity change remains obscure (to me).

In order to continue, we begin by defining a position operator in the coordinate representation by

$$Q_\alpha |\mathbf{x}\rangle = x_\alpha |\mathbf{x}\rangle \tag{5.21}$$

and an average velocity

$$\frac{d}{dt}\langle \mathbf{Q}\rangle = \langle \mathbf{V}\rangle. \tag{5.22}$$

We will deduce the physical meanings of the symmetry operators by studying their relation to $\mathbf{Q}$.

By (5.2),

$$|\Psi(t)\rangle \to e^{isH}|\Psi(t)\rangle = |\Psi(t-s)\rangle,$$

so if $s = t$, shifting the exponential to the other side results in $|\Psi(t)\rangle = e^{-itH}|\Psi(0)\rangle$ and

$$\frac{d}{dt}|\Psi(t)\rangle = -iH|\Psi(t)\rangle. \tag{5.23}$$

(This equation is temporary, to derive the meaning of the generators, and differs from the Schrödinger equation (4.5) by a factor of $\hbar$, so you could consider it uses $\hbar = 1$. This will be the case from here up to (5.29).) Then carrying out (5.22) using integration by parts and (5.23) leads to

$$\mathbf{V} = \frac{d}{dt}\langle \Psi(t)|\mathbf{Q}|\Psi(t)\rangle = i[H, \mathbf{Q}]. \tag{5.24}$$

This is the commutator for $\mathbf{Q}$ and H. We can find the other commutators of $\mathbf{Q}$, with P, J and G.[39] The results are:

39 Details in Ballentine, 76-85. Clearly, I am skipping a lot.

$$[Q_\alpha, P_\beta] = i\delta_{\alpha\beta} I,$$
$$[J_\alpha, Q_\beta] = i\epsilon_{\alpha\beta\gamma} Q_\gamma,$$
$$[G_\alpha, Q_\beta] = 0.. \tag{5.25}$$

(Again, factors of $\hbar$ are missing on the right-hand side of these equations.) Consideration of a free particle with no internal degrees of freedom requires

$$G_\alpha = MQ_\alpha \quad \text{and} \quad \mathbf{J} = \mathbf{Q} \times \mathbf{P}. \tag{5.26}$$

(See also section 7.2.1.) The commutator (h) and (5.26) give

$$[Q_\alpha, H] = \frac{iP_\alpha}{M}, \text{ which is satisfied by } \quad H = \frac{\mathbf{P} \cdot \mathbf{P}}{2M} + E_0, \tag{5.27}$$

and from (5.22)

$$\mathbf{V} = \frac{\mathbf{P}}{M}. \tag{5.28}$$

Although M looks a like the mass, we cannot be certain that it is, as equations (5.26) to (5.28) are equally satisfied by a quantity proportional to the mass. We therefore assign the proportionality constant a name. Guess what?

$$\frac{M}{mass} = \frac{P}{momentum} = \frac{H}{energy} = \frac{J}{angular momentum}$$
$$= \text{a fundamental constant, such as } \hbar^{-1}, \tag{5.29}$$

Then $\hbar$ is a number which can be fixed only by experiment. Now we have but to accept M, P, H and J as mass, momentum, energy and angular momentum. Then in all preceding equations, we must replace these symbols with the same divided by $\hbar$, as (5.29) implies $M = \frac{mass}{\hbar}$. So the unitary operators for space displacement, rotation and time evolution are

$$exp(-i\mathbf{a} \cdot \mathbf{P}/\hbar), \qquad exp(-i\theta\hat{\mathbf{n}} \cdot \mathbf{J}/\hbar) \qquad \text{and} \qquad exp(-itH/\hbar).$$

We also need to add $\hbar$ to the non-zero right-hand-side of the commutators in Table 2, especially

$$[J_\alpha, J_\beta] = i\hbar\epsilon_{\alpha\beta\gamma} J_\gamma, \qquad \text{or} \qquad [J_i, J_j] = i\hbar J_k \tag{5.30}$$
$$[J_\alpha, P_\beta] = i\hbar\epsilon_{\alpha\beta\gamma} P_\gamma$$

and the commutation relations of (5.25), written in a more familiar form, the ***canonical commutation relation*** of QM:

$$[x_\alpha, p_\beta] = i\hbar\delta_{\alpha\beta}. \tag{5.31}$$

Identification of the generators proceeded from their commutators with the position operator, the only classical operator to be identified from the outset with a physical observable.

Note that the derivation of these operators started with symmetry under Galilean transformations, their unitarity and hence their generators' being Hermitian, and general QM principles. It was only based on equations of classical physics towards the end of the derivation (e.g., 5.28).

5.5. Résumé of Galilean (NRQM) symmetries

Here are the bare bones of how Galilean symmetries are related to physical observables momentum, angular momentum and energy.

- Essential point: Transformations on vector spaces of interest in physics can be represented by unitary operators.
- Any one-parameter group of unitary operators can be written in the form of an exponential in terms of its

Hermitian generator.

- We can actively transform a function or passively transform the coordinates in terms of which the function is expressed. Either you move it ahead or you take a step back.
- QM states must be invariant – symmetric -- under the ten Galilean transformations, which can be subdivided into three each of rotations, translations and velocity changes, plus one time translation -- three vectors and one scalar.
- Since the transformations of the Galilean group form just that, a group, two consecutive Galilean transformations form another. Using this fact and some obvious physical facts about the order of operations, we can calculate the Lie brackets for the group generators.
- The effects of these transformations then can be used to deduce their commutation relations (Table 2).
- Definition of a position operator and derivation of its commutation relations with the Galilean generators, plus some intuition, leads to the identification of the Galilean generators as momentum (translation), angular momentum (rotation) and energy (time evolution). The meaning of the generator of velocity change remains obscure (to me).
- Identification of the generators proceeded from their commutators with the position operator, the only classical operator to be identified from the outset with a physical observable.

6. Representations

We consider two representations of the state vector, in terms of eigenfunctions of position (coordinates) and momentum.

6.1. Coordinate representation and applications

We use the coordinate representation of an abstract linear vector space $|\psi\rangle$ by taking as basis vectors the set of eigenvectors of the position operator (5.21). An expansion coefficient is commonly called a ***wave function***, which is unfortunate, as they do not represent classical waves:

$$\langle x|\psi\rangle = \psi(x). \tag{6.1}$$

An operator in this representation is given by the rule

$$A\psi(x) = \langle x|A|\psi\rangle. \tag{6.2}$$

The position operator just multiplies by the eigenvalue

$$Q_\alpha\psi(x) = x_\alpha\psi(x). \tag{6.3}$$

The momentum operator can be derived from its role as generator of space displacements

$$e^{(-i\mathbf{a}\cdot\mathbf{P}/\hbar)}|x\rangle = |x+a\rangle.$$

Expansion of this by a series and comparing that to the expansion in a of $\psi(x+a)$ leads to

$$P_\alpha = -i\hbar\frac{\partial}{\partial x_a}. \tag{6.4}$$

Then for a single particle (4.5) becomes ***Schrödinger's equation***

$$\Big[\frac{-\hbar^2}{2M}\nabla^2 + W(x)\Big]\Psi(x) = i\hbar\frac{\partial}{\partial t}\Big]\Psi(x). \tag{6.5}$$

This looks like a wave equation, but that is just because it takes into account only a single particle. Consider the equation for N particles:

$$\Big[\sum_{n=1}^{N}\frac{-\hbar^2}{2M_n}\nabla_n^2+\sum_{n=1}^{N}W(x^n)+V(x^1,...,x^N)\Big]\Psi(x)=i\hbar\frac{\partial}{\partial t}\Psi(x^1,...,x^N) \qquad (6.6)$$

in an abstract N-dimensional configuration space. Here, $V(x^1,...,x^N)$ is the potential for interactions between particles. If each particle was associated with a wave, there would be N waves in ordinary 3d space. But in equation (6.6), there is only one "wave function" for all N particles. So it is impossible to associate a physical wave in 3-d space with each particle. Ballentine insists that the "… correct interpretation of Ψ is as a statistical state function, a function from which probability distributions for all observables may be calculated." Then $|\Psi(x^1,...,x^N)|^2$ is the ***probability density*** in configuration space for particle (1) being at position $x^{(1)}$, particle (2) at position $x^{(2)}$ and so on. In an experiment which separates a beam into two parts by a semi-reflecting mirror, this means that never is a particle found at once in both separated beams. This result has been confirmed experimentally.[40]

All our derivations so far have been based on Galilean symmetry, so the Schrödinger equation must be invariant under those transformations. Consider the example of a transformation only by a uniform velocity displacement:

$$x=x'+vt',\ \ t=t'. \qquad (6.7)$$

Then

$$\frac{\partial}{\partial x'}=\frac{\partial}{\partial x},\qquad \frac{\partial}{\partial t'}=\frac{\partial}{\partial t}+v\frac{\partial}{\partial x}. \qquad (6.8)$$

For the Schrödinger equation to be invariant under these transformations, we must have

$$|\Psi(x,t)|^2=|\Psi'(x',t')|^2 \qquad (6.9)$$

and so, including a phase, f, which does not change any observable,

$$\Psi(x,t)=e^{if}\Psi'(x',t'). \qquad (6.10)$$

Putting (6.8) into (6.5) and assuming $W'(x',t')=W(x,t)$, we ca carry out the (laborious) calculation and show this to be true only if

$$f(x,t)=\frac{Mvx-\frac{1}{2}Mv^2t}{\hbar}. \qquad (6.11)$$

If the potential is zero, then

$$\Psi'(x',t')=e^{i(kx'-\omega t')}, \qquad (6.12)$$

and combining the last three equations gives

$$\begin{aligned}\Psi(x,t)&=e^{if(x,t)}\Psi'(x-vt,t)\\&=exp\Big[\frac{i}{\hbar}(\hbar k+Mv)x-\frac{i}{\hbar}\frac{(\hbar k+Mv)^2}{2M}t\Big],\end{aligned} \qquad (6.13)$$

which agrees with $P=\hbar k\rightarrow\hbar k+Mv$. The so-called wave function does not behave like an ordinary 3d wave because of the extra phase factor in (6.10) and equation (6.13) shows that this term is essential for Galilean symmetry.

For a single particle, the probability of the particle's being within a region Ω is $\int_\Omega|\Psi(x,t)|^2d^3x$. Taking the time derivative of this leads to the ***continuity equation***

$$\frac{\partial}{\partial t}|\Psi(x,t)|^2+\nabla\cdot\vec{J}(x,t)=0$$

with

40 Ballentine, 101.

$$J(x,t) = \frac{-i\hbar}{2M}(\Psi^*\nabla\Psi - \Psi\nabla\Psi*).^{41} \tag{6.14a}$$

$$= \frac{\hbar}{M} Im(\psi^*\nabla\Psi). \tag{6.14b}$$

Interestingly, the velocity operator (5.28)

$$\mathbf{V} = \frac{\mathbf{P}}{M} = \frac{-i\hbar}{M}\nabla$$

and (6.14b) tell us that

$$J(x,t) = Re[\Psi^*(x,t)\mathbf{V}\Psi(x,t)], \tag{6.15}$$

so if $\Psi(x,t)$ is property normalized, the integral of J over all space is the average velocity of the state:

$$\int J(x,t)d^3x = \langle\Psi|\mathbf{V}|\Psi\rangle. \tag{6.16}$$

The continuity equation requires the continuity of J and according to (6.14), it requires the continuity not only of Ψ, but also of $\nabla\Psi$.[42]

Ballentine goes on to discuss again the necessity of rigged Hilbert spaces, as well as tunneling and the path integral method.[43] Although the path-integral method has few practical uses in ordinary QM, it possesses great generality. “It is common to all formulations of quantum mechanics that the probability of a process is given by the squared modulus of a complex amplitude.” Path integrals make clear the importance of interference in QM. The phase of each amplitude is simply related to the action along the path, and this fact makes the classical correspondence logical given that the classically allowed paths often dominate. This requirement is valid for physically realizable states, but not necessarily for eigenfunctions of observables, which are only required to lie in the extended space of the rigged Hilbert space. Such considerations lead to the requirement for $k = p/\hbar$ of a free particle to be real and its energy to be non-negative.[44]

High points

- Momentum operator in coordinate representation/
- Schrödinger equation for single and multiple particles.
- Example of Galilean transformation invariance of Schrödinger equation.
- Probability density, average momentum and continuity equation.

6.2. Momentum representation and applications

Let’s start by defining the ***Fourier transform*** and its inverse in terms of general functions.

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} F(k)e^{ikx}dk \tag{6.17}$$

$$F(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)e^{-ikx}dx \tag{6.18}$$

In these equations, $F(k)$ is the ***Fourier transform*** of $f(x)$; $f(x)$ is the ***inverse Fourier transform*** of $F(k)$.

From (6.4), the momentum eigen-equation in the coordinate representation is

41 I should be putting x in bold-face squarely as a vector in all these equations, but it’s too damn much trouble.
42 Ballentine, 106-107, discusses the nuances of the statement.
43 For the equation(s) of the path integral method, see my notes on symmetry and QFT.
44 Ballentine, 110.

$$\hat{p}\psi(x) = -i\hbar\frac{d\psi(x)}{dx} = p\psi(x) \tag{6.19}$$

of which "normalized" solutions in a single dimension are of the form

$$\psi_p(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{ipx}{\hbar}} = \frac{1}{\sqrt{2\pi}}e^{ikx}, \tag{6.20}$$

where $p = \hbar k$, as usual. What is meant in this case by "normalized" is

$$\langle\psi_{p'}|\psi_p\rangle = \delta(p - p'),$$

which is referred to as ***Dirac normalization***. In fact, the eigenfunctions of a hermitian operator whose spectrum is continuous like this are not normalizable, but a wave packet around this momentum value is. Fortunately, eigenfunctions with real eigenvalues are Dirac ortho-normalizable and complete.

Equation (6.20) is the equation in the x basis of a wave of wavelength

$$\lambda = \frac{2\pi\hbar}{p}, \tag{6.21}$$

which is the ***De Broglie formula***. Equations (6.19) and (6.20) are in the ***x basis*** or ***coordinate representation***. In Dirac notation

$$\psi_p(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{ipx}{\hbar}} = \langle x|p\rangle, \tag{6.22}$$

the momentum wave function in the coordinate representation, or the projection of the p function onto x. In the position representation, the probability for a given value of x (although we should multiply by dx and say for a given increment of x) is

$$P(x) = \psi^*(x)\psi(x) \tag{6.23}$$

and for a given value of momentum

$$P(p) = {\psi_p}^*(x)\psi_p(x).$$

But we can also say that

$$P(p) = \widetilde{\psi}^*(p)\widetilde{\psi}(p) \tag{6.24}$$

where $\widetilde{\psi}(p)$ is the wave function in the ***momentum representation***.

Slipping back into Dirac notation, define the identity operator in both bases

$$\mathbf{I} = \int dx|x\rangle\langle x| \tag{6.25}$$

and

$$\mathbf{I} = \int dp|p\rangle\langle p|. \tag{6.26}$$

Supposing a state vector $|\Psi\rangle$, the wave function in the position-representation is

$$\psi(x) = \langle x|\Psi\rangle \tag{6.27}$$

and in the momentum representation

$$\tilde{\psi}(p) = \langle p|\Psi\rangle. \tag{6.28}$$

Note the order here: *The **wave function** is the **projection** of the **state vector onto** the **eigenvectors** of the **representation**.* This is just the content of equations (6.27) and (6.28).

Using the identity operator in the x basis

$$\tilde{\psi}(p) = \int dx \langle p|x\rangle\langle x|\Psi\rangle,$$

so from the complex conjugate of the momentum eigenvector (6.22) in the x basis and the wave function in the p basis (6.27),

$$\widetilde{\psi}(p) = \frac{1}{\sqrt{2\pi}} \int dx e^{\frac{-ipx}{\hbar}} \psi(x). \tag{6.29}$$

Starting from the x-basis wave function (6.27) and using the identity operator (6.26) leads to the inverse

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int dp e^{\frac{ipx}{\hbar}} \widetilde{\psi}(p) \tag{6.30}$$

These two representations, position and momentum, can be seen from (6.17) and (6.18) to be ***Fourier transforms*** of each other. They are the means for transforming a wave function from the position representation to the momentum representation or vice versa.

Just as the coordinate representation takes as basis vectors the set of eigenvectors of the position operator, so does the momentum representation use as basis vectors the eigenvectors of the momentum operator, Back in 3-d,

$$\hat{p}_\alpha|\mathbf{p}\rangle = p_\alpha|\mathbf{p}\rangle \qquad (\alpha = 1, 2, 3).$$

where continuous eigenvalues require orthogonality and normalization

$$\langle \mathbf{p}|\mathbf{p}'\rangle = \delta(\mathbf{p} - \mathbf{p}').$$

In order to discover the relation between the momentum and position eigenvectors, we calculate the inner product of the two, using the BCH formula (5.17) and the momentum operator in the position representation with $\mathbf{p} = \hbar\mathbf{k}$.

$$-i\hbar\nabla\langle x|\mathbf{p}\rangle = \langle x|\hat{\mathbf{p}}|\mathbf{p}\rangle = \mathbf{p}\langle x|\mathbf{p}\rangle.$$

The solution after normalization is

$$\langle \mathbf{x}|\mathbf{p}\rangle = (2\pi\hbar)^{-3/2} e^{i\mathbf{k}\cdot\mathbf{x}}.$$

Just as we had a "wave function", the coordinate representation of a state vector $\langle x|\psi\rangle = \psi(x)$, so by using $\langle \mathbf{x}|\mathbf{p}\rangle = \langle \mathbf{p}|\mathbf{x}\rangle^*$ can we find the same state vector in the momentum representation.

$$\langle \mathbf{p}|\psi\rangle = \int \langle \mathbf{p}|\mathbf{x}\rangle\langle \mathbf{x}|\psi\rangle d^3x = (2\pi)^{-3/2} \int e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\psi(\mathbf{x})d^3x = \Phi(\mathbf{p}).$$

which is the F***ourier transform*** of $\psi(\mathbf{x})$. We can then find the position operator in the momentum representation.

$$\langle \mathbf{p}|Q_\alpha|\psi\rangle = (2\pi)^{-3/2} \int e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} x_\alpha \psi(\mathbf{x})d^3x = i\hbar\frac{\partial\Phi(\mathbf{p})}{\partial p_\alpha},$$

so the position operator in the momentum representation is

$$Q_\alpha = i\hbar\frac{\partial}{p_\alpha},$$

the form of which is as pleasing as it is unsurprising.

Chapter high points

- The momentum-representation state vector is the Fourier transform of the position one, and vice versa.

- The position operator in the momentum representation is analogous to the momentum operator in the position representation.

7. Useful examples of QM calculation

Angular momentum and harmonic oscillators will pop up all over, for spin or for particles in QFT (which we will not discuss here.)

7.1. The simple harmonic oscillator (SHO)

A harmonic oscillator is a system subject to a quadratic potential which produces a force which tends to restore the system to equilibrium (Hooke's law). The math of harmonic oscillators is extraordinarily valuable because all sorts of systems are subject to such a force, at least infinitesimally.

For a classical 1-d SHO, the potential is

$$V = \tfrac{1}{2}kx^2$$

so Newton's second law is

$$F = -kx = m\frac{dx^2}{dt^2}. \tag{7.1}$$

Solutions of this equation are of the form

$$x(t) = A\sin(\omega t - \phi) + B\cos(\omega t - \phi).$$

If B=0, the kinetic and potential energies are then

$$K(t) = \frac{1}{2}mv^2(t) = \frac{1}{2}\omega^2 A^2\cos^2(\omega t - \phi) \tag{7.2a}$$

and

$$.V(t) = \frac{1}{2}kx^2(t) = \frac{1}{2}kA^2\sin^2(\omega t - \phi). \tag{7.2b}$$

Both of these energies depend on the square of the amplitude, A. The total energy is then

$$E = K + V = \frac{1}{2}kA^2, \tag{7.3}$$

with $k = m\omega^2$, or $\omega = \sqrt{k/m}$.

In QM, a SHO must satisfy a Schrödinger equation of the form

$$\Big(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2\Big)\psi = E\psi. \tag{7.4}$$

The solutions involve Hermite polynomials of degree n and the energy of each state is

$$E_n = \Big(n + \frac{1}{2}\Big)\hbar\omega, \tag{7.5}$$

where again $\omega = \sqrt{k/m}$, or $k = m\omega^2$.

However, there is a more interesting way to solve for the energy eigenvalues. Rewrite the Schrödinger equation in terms of momentum and position operators.

$$\frac{1}{2m}\Big(\hat{p}^2 + (m\omega x)^2\Big)\psi = E\psi.$$

It would be nice to factor this into something like $(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x})$, but that would not work because in

QM the momentum and position operators do not commute. Instead, define two *operators*,[45]

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{x} + \frac{i}{m\omega}\hat{p}\Big) \tag{7.6}$$

and its *Hermitian conjugate*

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{x} - \frac{i}{m\omega}\hat{p}\Big) \tag{7.7}$$

in terms of operators $\hat{x}$ and $\hat{p}$. Since neither one is Hermitian, they are not operators for observables. Because of the position-momentum commutation relations (5.31),

$$[\hat{x}, \hat{p}] = i\hbar,$$

it turns out that

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{7.8}$$

The Hamiltonian can be written

$$\hat{H} = \hbar\omega\Big(\hat{a}^\dagger\hat{a} + \frac{1}{2}\Big), \tag{7.9}$$

where comparison with (7.5) shows that the ***number operator***

$$\hat{n} = \hat{a}^\dagger\hat{a} \tag{7.10}$$

returns the number of entire units $\hbar\omega$ of energy in a given state:

$$\hat{n}|n\rangle = n|n\rangle. \tag{7.11}$$

Then $|n\rangle$ is also an ***eigenstate*** of the Hamiltonian, which may be written as

$$\hat{H} = \hbar\omega\Big(\hat{n} + \frac{1}{2}\Big) \tag{7.12}$$

so that

$$\hat{H}|n\rangle = \Big(n + \frac{1}{2}\Big)\hbar\omega|n\rangle, \tag{7.13}$$

The eigenvalues are equally spaced at intervals of $\hbar\omega$. Since

$$\hat{n}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger\hat{a}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger(1 + \hat{a}^\dagger\hat{a})|n\rangle = (n+1)\hat{a}^\dagger|n\rangle, \tag{7.14}$$

it turns out that $\hat{a}^\dagger$ has the effect of increasing the energy of state $|n\rangle$ by one unit of $\hbar\omega$; it therefore is called a ***raising operator***. Aside from the form of the Hamiltonian (due to the potential), this result depends only on the fact that the position and momentum operators are Hermitian and on their commutation relations.

Similarly,

$$\hat{n}\hat{a}|n\rangle = (n-1)|n\rangle, \tag{7.15}$$

so $\hat{a}$ is considered a ***lowering operator*** and the pair, raising and lowering, are called ***ladder operators***. Energy is positive, so there must be a ground state to which application of $\hat{a}$ gives a zero or negative result, and this can easily be used to find the wave function of that state.[46] Plugging that back into the Schrödinger equation yields the result for the ground-state energy

$$E_0 = \frac{1}{2}\hbar\omega.$$

45 For details, see Townsend, 196-199, or Griffiths and Schroeder, 40-46, or Lancaster and Blundell, 19-23.
46 Griffiths and Schroeter, 43-44.

Since repeated application of the raising operator to the ground state $|0\rangle$ can create any energy state, we can write the normalized state

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{7.16}$$

Note well, all this follows from the (classical) form of the Hamiltonian and the QM commutation relations for the position and momentum operators. In QFT, the increments of energy will be taken to be particles so that the ladder operators become creation and annihilation operators.

7.2. Angular momentum

Angular momentum is an important and multi-faceted topic in QM.

7.2.1. Orbital angular momentum

One way to consider angular momentum in QM is to start with its classical definition

$$\vec{L} = \vec{r} \times \vec{p}$$

and use the canonical commutation relations for position and momentum operators:

$$[x_i, p_j] = i\hbar\delta_{ij}, \qquad [x_i, x_j] = [p_i, p_j] = 0. \tag{7.17}$$

The result (cyclic in I,j,k or x,y,z) is

$$[L_i, L_j] = i\hbar L_k \tag{7.18}$$

along with

$$[L^2, \vec{L}] = 0, \tag{7.19}$$

where

$$L^2 = L_x^2 + L_y^2 + L_z^2.$$

The same equations show that although the three components of L are incompatible as observables, one of them, say L_z, which would be called a Cartan operator in group theory, commutes with L^2, which would then be the Casimir operator, so the two may have simultaneous eigenvalues and common eigenvectors, as we saw in section 4.2.

One can see the commutation relations of equation (7.18) as the basis of angular momentum (and spin) in QM[47]; everything else concerning angular momentum follows from them. We can, nevertheless, look at it the other way around and start with our knowledge of Galilean transformations, in particular, rotations.

From Table 1, the unitary operator for a rotation about an axis parallel to the unit vector $\hat{n}$ is

$$\mathbf{R}_n(\theta) = e^{-i\theta\hat{n}\cdot\mathbf{J}/\hbar}. \tag{7.20}$$

For a one-component state, from (5.9), we have

$$\mathbf{R}\Psi(x) = \Psi(R^{-1}x).$$

Recall that this equation expresses the equivalence between an active transformation of the function, on the left, and a passive transformation of the coordinates, on the right. For a rotation through angle ϵ about the z-axis, it becomes

$$\mathbf{R}_z(\epsilon)\Psi(x, y) = \Psi(x\cos(\epsilon) + y\sin(\epsilon), -x\sin(\epsilon) + y\cos(\epsilon), z).$$

We can consider this a function of a single variable ϵ,

47 Susskind's point of view.

$$\Psi(\epsilon) = \Psi(x(\epsilon), y(\epsilon), z).$$

Then for infinitesimal ϵ it may be expanded in terms of ϵ to give in the lowest order

$$\mathbf{R}_z(\epsilon) = \Psi(x, y, z) + \epsilon\Big(y\frac{\partial\Psi}{\partial x} - x\frac{\partial\Psi}{\partial y}\Big).$$

Comparison of this to the expansion of (7.20) gives

$$J_z = -i\hbar\Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)$$

which is the z component of the orbital angular momentum, $L = Q \times P$, in the coordinate representation.

Suppose the eigenvalue equations for L^2 and L_z to be as follows:

$$L^2 = \lambda f \qquad \text{and} \qquad L_z f = \mu f.$$

Define operators

$$L_\pm = L_x \pm iL_y. \tag{7.21}$$

Because of the commutation relations,

$$[L_z, L_\pm] = \pm\hbar L_\pm,$$

it turns out $L_\pm$ is a ladder (raising/lowering) operator, since

$$L_z(L_\pm f) = (\mu \pm \hbar)(L_\pm f). \tag{7.22}$$

The raising operator raises the L_z value of the eigenstate by $\hbar$. The raised or lowered eigenstate has the same value of L^2. Since L_z has to be less than or equal to L, taking its maximum value to be l (the negative of which turns out to be its minimum), we find that

$$L_z f_t = \hbar l f_t,$$

where f_t is the "top" eigenfunction, as for the case of a SHO. Including both observables in the function notation, this in turn leads to

$$L^2 f_l^m = \hbar^2 l(l+1) f_l^m \tag{7.23}$$

and

$$L_z f_l^m = \hbar m f_l^m, \tag{7.24}$$

where l = 0, ½, 1, 3/2, ... and $m = -l, -l+1, ..., l-1, l$.

These values follow from the requirement that m take on integrally-spaced values from $-l$ to l, so $l = N/2$ and must therefore be an integer or half integer.

Note that the maximum measurable (observable) value of the z component of L is less than the magnitude of L.

7.2.2. Internal angular momentum

If the state contains more than one component, we can use matrices to represent the state and the rotation operator. The most general case, according to Ballentine and equation (5.9), is

$$\mathbf{R}\begin{bmatrix}\Psi_1(x)\\ \Psi_2(x)\\ \cdots\end{bmatrix} = D\begin{bmatrix}\Psi_1(R^{-1}x)\\ \Psi_2(R^{-1}x)\\ \cdots\end{bmatrix}.$$

The operator D may act on internal degrees of freedom and result in linear combinations of the state functions. So we update (7.20) to

$$\mathbf{R}_n(\theta) = e^{-i\theta\hat{n}\cdot\mathbf{L}/\hbar} D_n(\theta). \tag{7.25}$$

Since $D_n(\theta)$ must be unitary, we can write it as

$$D_n(\theta) = e^{-i\theta\hat{n}\cdot\mathbf{S}/\hbar}, \tag{7.26}$$

where the coordinate components of S are Hermitian. Substituting (7.26) into (7.25) and comparing to (7.20) shows that the total angular momentum has the form

$$\mathbf{J} = \mathbf{L} + \mathbf{S}.$$

It's not divulging a major spoiler to say that we will identify $\mathbf{S}$ with the spin angular momentum, with the same commutation elements as L. Helpfully, Ballentine says that the orbital angular momentum is that"...of the motion of the center of mass of the object relative to the origin of coordinates. The spin may be identified as the angular omentum that remains when the center of mass is at rest."[48] If $\mathbf{L}$ and $\mathbf{S}$ are conserved, then so is $\mathbf{J}$. If there is coupling between the two, such as spin-orbit coupling in atoms, then they are not conserved independently, but $\mathbf{J}$ is.

Being an angular momentum, the spin obeys the usual equations (7.18) and (7.19) in the form

$$S^2|s,m\rangle = \hbar^2 s(s+1)|s,m\rangle, \qquad S_z|s,m\rangle = \hbar m|s,m\rangle. \tag{7.27}$$

We normally consider the spin as existing in a state of dimension 2s+1 spanned by the eigenvectors of (7.27).

7.2.3. Spin ½

For the spin ½ case, we write

$$\mathbf{S} = \tfrac{1}{2}\hbar\boldsymbol{\sigma}, \tag{7.28}$$

where $\boldsymbol{\sigma}$ denotes the usual Pauli spin matrices. For group theory fans, this is the two-dimensional representation of SU(2), of rank 3 ($4^2 - 1$), with one Cartan operator, usually J_z, and Casimir element J^2.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.29}$$

The spin operator in an arbitrary direction given by the unit vector $\hat{n}$ is $\hat{n}\cdot\mathbf{S} = \frac{1}{2}\hbar\hat{n}\cdot\sigma$ where

$$\hat{n} = (sin\theta cos\phi, sin\theta sin\phi, cos\theta)$$

from polar coordinates with $r = |\hat{n}| = 1$. Then writing out the components of $\hat{n}\cdot\sigma$ gives

$$\hat{n}\cdot\sigma = \begin{bmatrix} cos\theta & e^{-i\phi}sin\theta \\ e^{i\phi}sin\theta & -cos\theta \end{bmatrix}.$$

That the eigenvalues are ± 1 is easily found from the characteristic equation, and the normalized eigenvectors may be chosen as[49]

$$\begin{bmatrix} e^{-i\phi/2}cos(\theta/2) \\ e^{i\phi/2}sin(\theta/2) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} e^{-i\phi/2}sin(\theta/2) \\ e^{i\phi/2}cos(\theta/2) \end{bmatrix}, \tag{7.30}$$

Note two things:

- those half angles, $\theta/2$, about which more shortly; and
- in each eigenvector the phase is given by one parameter, ϕ , and the magnitude by the other, θ.

The rotation is indeed through an angle θ but its effect on the spin space is that of a rotation by half θ.

We are in a space of two dimensions, so the two angles are adequate for parametrization of both the relative magnitude and the relative phase of the eigenvectors. So if we choose a component of a pure system of spin s=½, usually the z-component, we can orient it in any direction in space in such a way that that is the direction of

48 Ballentine, 166.
49 Ballentine, 172.

the $+\frac{1}{2}\hbar$ eigenvector.[50]

The state operator is a 2x2 matrix and so can be expressed in terms of four linearly independent matrices such as $\mathbf{1}$ and the three Pauli matrices. So any state vector for spin ½ can be written as

$$\rho = \tfrac{1}{2}(\mathbf{1} + \mathbf{a}\cdot\boldsymbol{\sigma}). \tag{7.31}$$

Since the Pauli matrices are traceless and $Tr(\rho) = 1$,

$$\langle\sigma_x\rangle = Tr(\rho\sigma_x) = \frac{1}{2}Tr(\sigma_x\mathbf{1} + a_x\sigma_x\sigma_x + a_y\sigma_y\sigma_x + a_z\sigma_z\sigma_x) = a_x$$

and similarly for the y and z components, so that

$$\langle\boldsymbol{\sigma}\rangle = Tr(\rho\boldsymbol{\sigma}) = \mathbf{a}, \tag{7.32}$$

the ***polarization vector*** of the state.

7.2.4. Rotation by 2 pi

As for the half-angles, from (6.19) we have for a rotation through an angle 2π about unit vector $\hat{n}$

$$\mathbf{R}_n(2\pi)|j,m\rangle = e^{-2\pi i\hat{n}\cdot\mathbf{J}/\hbar}|j,m\rangle = (-1)^{2j}|j,m\rangle. \tag{7.33}$$

(This is fairly easy to show for J along the z-axis, noting that the j in the exponent of the last term is there because it has the same integer or half-integer character as the J_z eigenvalue m.) It can be shown to be true for any direction, so we generally just write it as $\mathbf{R}(2\pi)$.

$\mathbf{R}(2\pi)$ is not a trivial operator and has non-trivial consequences. We are accustomed to think that all dynamical variables are invariant under a rotation of 2π, so, by (3.13),

$$\mathbf{R}(2\pi)A\mathbf{R}^{-1}(2\pi) = A, \qquad \text{so} \qquad [\mathbf{R}(2\pi), A] = 0, \tag{7.34}$$

where A is any physical observable whatsoever. But (7.33) shows that the state vector is not necessarily invariant under a rotation by 2π. This is in agreement with (7.20) and its half angles. So we must distinguish between the invariance of observables and that of states. Indeed, suppose a unitary operator U which leaves a dynamical variable F invariant, so that by (7.34) we know $[U, F] = 0$, but not the state, so

$$|\Psi'\rangle = U|\Psi\rangle \neq \Psi.$$

But in the transformed basis

$$\begin{aligned}\langle F\rangle = \langle\Psi'|F|\Psi'\rangle &= \langle\Psi|U^\dagger FU|\Psi\rangle = \langle\Psi|U^\dagger UF|\Psi\rangle, \qquad \text{by (6.29)}\\ &= \langle\Psi|F|\Psi\rangle,\end{aligned}$$

and the dynamical variable has the same expectation value in both unequal states. This can also be shown for a state vector in format (4.18).

Now, equation (7.33) shows that rotations by 2π divides the vector space into two subspaces. We can denote the integer angular momentum states by $|+\rangle$ and the half-integer ones by $|-\rangle$, so that

$$\mathbf{R}(2\pi)|+\rangle = |+\rangle \qquad \text{and} \qquad \mathbf{R}(2\pi)|-\rangle = -|-\rangle.$$

If A represents any physical observable, then by (7.33)

$$\langle+|\mathbf{R}(2\pi)A|-\rangle = \langle+|A\mathbf{R}(2\pi)|-\rangle \quad \text{so that} \qquad \langle+|A|-\rangle = -\langle+|A|-\rangle,$$

which is impossible, so $\langle+|A|-\rangle = 0$. No physical observable can have non-vanishing matrix elements between integer angular momentum states and half-integer ones. This ***superselection rule*** may be stated as saying that interference of vectors of the $|+\rangle$ and $|-\rangle$ types is not observable. Since they have no observable

50 Ballentine, 172. "Therefore any pure state vector of an s=½ system can be associated with a spatial direction $\hat{n}$ for which it is the $+\frac{1}{2}\hbar$ eigenvector for the component of spin."

consequences, we may consider them to be zero. So the probability of a transition between the two subspaces is zero: Once a half(or whole)-integral, always a half(or whole)-integral.

Such a superselection rule resembles a symmetry in that it represents a transformation which gives rise to a conserved quantity, the eigenvalue ± 1. But whereas there are no observables which do not commute with $\mathbf{R}(2\pi)$, there are such observables in the case of ordinary symmetries, a common example being between position and momentum. Measurement of position permits distinguishing between states that differ only by a displacement, but there is no way to distinguish between states that differ only by a 2π rotation.[51] However, experiments using neutron interferometry have confirmed effects due to the minus sign introduced by a 2π rotation of fermions.[52]

High points

- From the commutation relations, we can define ladder operators which permit deriving the eigenvalues of J^2 and J_z.
- The exponential form of the unitary rotation operators shows that $\mathbf{J}$ is the angular momentum.
- The rotation operator acting on internal degrees of freedom leads to the definition of another angular momentum, which we call the spin. The total angular momentum, orbital plus spin, is conserved.
- Usual equations for the spin in terms of the Pauli matrices. The eigenvalues of spin ½ along a given axis are $\pm\frac{1}{2}\hbar$.
- We must distinguish between invariance of observables and of states. Dynamical variables are invariant under a rotation by 2π. This is not necessarily true for state vectors. $\mathbf{R}(2\pi)$ multiplies a state vector by +1 in the case of integer spin, -1 in the case of half-integer spin, effectively dividing the space into two subspaces.
- No physical observable can have non-vanishing matrix elements between integer angular momentum states and half-integer ones. This superselection rule resembles a symmetry in that it represents a transformation which gives rise to a conserved quantity, the eigenvalue ± 1. There are no observables which do not commute with $\mathbf{R}(2\pi)$, but such is not the case of ordinary symmetries.

7.3. EM in QM – charged particle in EM field

The goal here is to find a Lagrangian from which we can derive the ***Lorentz force*** on a particle of charge q and velocity $\vec{v}$ due to an electric field and a magnetic field:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \tag{7.35}$$

in its classical form. I will go with Leonard Susskind on this one and accept that there is no way one can make a suitable Lagrangian directly involving the magnetic field.[53] We therefore take the Lagrangian for a free particle and add a term which Is a multiple of the vector potential, as well as a term for the scalar potential ϕ .

$$L = \frac{m}{2}v^2 + \frac{e}{c}\vec{v} \cdot \vec{A} - e\phi(x,t). \tag{7.36}$$

We now have adopted e for the charge previously referred to as q (to avoid confusion later on). The other terms of course come from foreknowledge of the desired result. The mechanical or kinetic momentum along the x-axis is still $m\dot{x}$, but the ***canonical momentum***, defined as a derivative of the Lagrangian, is now

$$p_x \equiv \frac{\partial L}{\partial \dot{x}} = m\dot{x} + \frac{e}{c}A_x. \tag{7.37}$$

In this case (presence of magnetic field) the canonical momentum not only is different from the kinetic

51 Ballentine, 185.as a Legendre transform of the Lagrangian

52 Sakurai and Napolitano, 158.

53 Ballentine, 84, shows that only a scalar and a vector potential are compatible with Galilean symmetry. Both must be independent of the canonical momentum.

momentum, $m\dot{x}$, it is gauge-dependent, changing as the gauge field A changes, and therefore *not* an observable. It is the mechanical momentum $m\dot{x}$ which is an observable and a constant of motion.

By (2.3),

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i},$$

so if the Lagrangian is independent of coordinates (translation invariant), the *canonical momentum* is conserved.

So much for the Lagrangian approach, let's consider the Hamiltonian. Using (7.37),

$$H = p \cdot v - L = \frac{1}{2}mv^2 + e\phi(x,t), \tag{7.38}$$

since the A terms cancel. This reflects the fact that the Lorentz force due to a B field is perpendicular to the particle's direction of motion, so no work is done along that direction and no energy is added by this field. The Hamiltonian contains the usual $\frac{(m\dot{x})^2}{2m}$ depending only on the *mechanical* momentum.[54] So the energy is due to the good ol' mechanical momentum, meaning that the kinetic energy calculated is just the usual $\frac{1}{2}mv^2$ and energy is conserved and the magnitude of the velocity is constant.[55]

In terms of momentum and position, the Hamiltonian is

$$H = \sum_i (p_i \dot{q}_i) - L = \sum_i \left\{ \frac{1}{2m}\left[p_i - \frac{e}{c}A_i\right]\left[p_i - \frac{e}{c}A_i\right]\right\} + e\phi(x,t). \tag{7.39}$$

From this equation, we can use Hamilton's equation of motion to calculate

$$\dot{p}_i = \frac{\partial H}{\partial q_i}\,. \tag{7.40}$$

Comparing this to the time derivative of (7.37) and using differentiation by parts, we obtain the Lorentz force due to the B field, which is what we wanted.

We have now seen that both the Lagrangian via (2.3) and the Hamiltonian via (7.40) tell us that if the Lagrangian (or the Hamiltonian) is independent of the coordinates, i.e., invariant under translation, then it is the *canonical momentum which is conserved*. Since it is then the generator of translations [Is it?], it is represented in the coordinate representation by $i\hbar\partial_x$.[56] In any case, that works.

In the case of a change in gauge A, the mechanical momentum $m\dot{x}$, which indicates the trajectory of the particle, is gauge-invariant and so does not change, Therefore, because of (7.37), the canonical momentum must change in order to compensate for the change in A.[57] Different gauge potentials correspond to different state vectors.

- In the absence of a magnetic (or other) force field, the mechanical and canonical momenta are equal.
- In the presence of a magnetic (or, maybe, other) field, the canonical momentum varies between different gauges. The mechanical momentum $m\dot{x}$ still follows the trajectory of the particle and is conserved. The canonical momentum is not gauge-invariant but changes in such a way as to maintain the conservation of mechanical momentum through (7.37).[58]
- The kinetic energy $mv^2/2$ is conserved in both cases.
- The canonical momentum is always the generator of translation. The operator for a translation δx is $exp(ip_x\delta x/\hbar)$, where the p in question is the non-invariant canonical momentum.

54 In the notation of Sakurai and Napolitano.
55 Susskind, TM, 208-209.
56 Heisenberg uncertainty principle and the canonical momentum operator, Physics Forums. www.physicsforums.com/threads/heisenberg-uncertainty-principle-and-the-canonical-momentum-operator.1067394/.
57 Sakurai and Napolitano, 129.
58 QUESTION: Is this still true in GR?

- The canonical-momentum operator in the coordinate representation is $-i\hbar\nabla$ because it's the generator of translations (?).

Put more intuitively, a change of gauge changes properties of the system under translation, so the state vector changes, as does the canonical momentum, since it is the generator of translation. The mechanical momentum and the energy remain invariant.

So, in the case of a charged particle in a magnetic field (ignoring eventual magnetic effects due to the charge).

- The non-invariance of the canonical momentum in this case, plus its being the generator of translation, means simply that translation symmetry does not hold and the canonical momentum is gauge dependent, not conserved. But mechanical momentum mv is conserved, as is the kinetic energy $mv^2/2$, both being gauge-invariant.

8. Composite systems and entanglement

We've been considering single systems with their own wave functions and eigenvalues – what we can call ***pure states***. Then the expectation value of an observable A is

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

A state vector for two different, independent measurements or subsystems may be represented sometimes as a product of the state vectors for each subsystem. For a product state, there exist two sets of normalization equations. For instance, in a product of two independent states each with two eigenvalues, each is represented by two complex components and so four variables, but normalization reduces each by one and a negligible phase factor eliminates another, leaving only four independent components (degrees of freedom) for the product state.

But in general, composite systems are not product states. For a state vector for the same subsystems which is not a product state, there is only one normalization requirement and one phase factor, so there remain six independent parameters, making this a more complicated state. Such a state Is said to be ***entangled***.

8.1. Density operator and entanglement

We will use the tool we learned in section 4.3, the ***density operator***:

$$\hat{\rho} \equiv |\Psi\rangle\langle\Psi|, \tag{8.1}$$

which projects an operator A onto the direction of $|\psi\rangle$:

$$\hat{\rho}|A\rangle = |\psi\rangle\langle\psi|\ |A\rangle = \langle\psi|A\rangle|\psi\rangle$$

In an orthonormal basis $\{|e_j\rangle\}$, its matrix elements are

$$\rho_{ij} = \langle e_i|\hat{\rho}|e_j\rangle = \langle e_i|\Psi\rangle\langle\Psi|e_j\rangle,$$

which constitute the elements of the ***density matrix***.

Case 1. Consider a standard case of spin up or down along the z-axis but write the density matrix for spin up along the x-axis. For this case[59]

$$\Psi = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix},$$

so brute-force insertion of this into (8.1) gives

$$\rho = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

59 Griffiths, QM, 169.

Obviously, ρ is Hermitian, has trace 1 and is idempotent:

$$\rho^2 = \rho.$$

These are properties of density matrices for all pure states. We will see that the last one is not true for entangled states.

As already seen in equation (4.21) , the density matrix allows us to calculate the expectation value of an operator:

$$\langle A \rangle = Tr|\psi\rangle\langle\psi|A = Tr(\rho L). \tag{8.2}$$

What about the case where the system might be in any of a number of states, but we ignore which one? We only know the probability of each state. In this case, the definition of the density operator is expanded as a weighted sum of density operators

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|, \tag{8.3}$$

with p_k being the probability of state ψ_k. A density matrix like this, which is a sum of projection operators, represents a ***mixed state***, one in which multiple states are possible.[60] A single projection operator represents a ***pure state***. Starting from (8.3), the same method of calculation used above leads again to equation (6.2). So knowing the density matrix, i.e., the states and their probabilities, we can calculate expectation values of operators in pure or mixed states.

EXAMPLE

Consider the example of an electron (or any fermion) with equal probabilities of spin up or down[61], represented by eigenvectors

$$|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |d\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad p_u = p_d = 1/2.$$

Then

$$\hat{\rho} = \sum_i p_i |\psi_o\rangle\langle\psi_i| = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}. \tag{8.4}$$

But this non-pure state is not idempotent.

$$\hat{\rho}^2 = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \neq \hat{\rho}.$$

Equation (8.2) for the average is true for such states also. This case represents our *ignorance* of the electron's spin state.

A more interesting is this one which is not a case of ignorance, but of … Suppose that André and Béatrice (hereafter referred to as Andy and Bea) are studying the decay of a π^0 into an electron and a positron, each of which has a spin which can be either up or down along some axis common to both. represented by respective operators (σ_A or σ_B). We use notation where a state is $|Andy\ Bea\rangle$, meaning Andy sees the component on the left in the ket; Bea, that on the right. The π^0 is the singlet state,

$$|sing\rangle = \frac{1}{\sqrt{2}}(|ud\rangle - |du\rangle),$$

where each pair ud (up-down) or du (down-up) represents the spin of the electron and the positron. We find[62]

$$\langle\sigma_{Az}\rangle = \langle\sigma_{Bz}\rangle = 0, \tag{8.5}$$

60 Susskind and Friedman,
61 Susskind calls this a mixed state, but Ballentine deprecates this term.
62 Susskind and Friedman, QM, 173

but[63]

$$\langle \sigma_{Az}\sigma_{Bz}\rangle = -1. \tag{8.6}$$

Result (8.5) means nothing is known about what a measurement of σ_{Az} or σ_{Bz} will find, each being equally likely to return +1 or -1. But because of (8.6), once one σ_z is measured, the other is known. This state is ***maximally entangled***, which means that *although the state is a complete description of the system taken as a whole, nothing is known about either subsystem by itself*. By nothing, we mean that only probabilities for different eigenvalues can be calculated. However, ignorance of individual subsystem states is accompanied by correlation between measurements of the individual subsystems.

It is not possible to isolate the wave functions of one subsystem of an entangled system. "One cannot really speak of the 'the state' of either particle separately."[64] Only in a pure, unmixed state can Bea study her own subsystem without learning anything about Andy's.

The density matrix for the singlet state is

$$\rho = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

showing the 50/50 mixture of states. This makes it clear that

$$\rho^2 = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \neq \rho.$$

It is the same as (8.4), the state vector which represented our *ignorance* of the particle's state, only here it represents the physical *impossibility* of predicting a measured value for either spin direction.

Even if a subsystem changes, it must conserve distinctions (i.e., complex scalar products). Therefore it will change by a unitary transformation U such that

$$U^\dagger U = I,$$

and Andy's density matrix is unchanged under a unitary change of Bea's subsystem[65]:

$$\rho^A_{aa'} = \sum_{b,b',b''} \psi^*(a'b'')(U^\dagger_{b''b}U_{bb'})\psi(ab') = \sum_{b'} \psi^*(a'b')\psi(ab') = \rho^A_{aa'}.$$

This means that Bea cannot influence Andy's statistical results, meaning that "... no influence can propagate faster than the speed of light"[66], which is the ***principle of locality.*** The evolution of Bea's system, even if it is entangled with Andy's, has no influence on Andy's statistical predictions. This is all about expectation values, in the spirit of quantum mechanics. Adding the least little bit of non-unitary evolution to Bea's subsystem would mean that she could influence Andy's subsystem *faster than the speed of light*.

What does this mean for the coexistence of QM and SR? This is a subject of much debate … and discomfort.

8.2. Locality and hidden variables – Bell's theorem

We have defined ***locality*** as the principle that no influence can propagate faster than the speed of light. But in the singlet state for, for instance, the decay of a π^0 into an electron and a positron, we know that if we measure the spin of one particle, then we know in advance of its measurement the spin of the other. But how can this information be passed from one measurement to the other and not disobey locality, i.e., by violating the speed limit c? It was suggested, in particular by Einstein, that QM must be incomplete, that there are so-called "hidden" variables which are unmeasured because unknown, and that these could at the moment of

63 Susskind and Friedman, QM, 177.
64 Grifffiths, QM, 448.
65 Susskind & Friedman, 225.
66 Griffiths, QM, 447.

the decay determine the spin directions of the two particles. The challenge was taken up by John Bell.[67]

Bell considered a slightly different hypothetical experiment. Let the π^0 decay as before, but

1. rotate the axes of spin measurement independently, so they are not always aligned;
2. consider not the individual spin directions +1 or -1, but the average value of their product.

Because of these arbitrary and independent rotations, the product of the spin directions along the different directions is not necessarily -1. Call these directions $\mathbf{a}$ and $\mathbf{b}$ and for given $\mathbf{a}$ and $\mathbf{b}$ define $P(\mathbf{a}, \mathbf{b})$ as the average of the product of the spins. If they are the same (parallel), we have necessarily for the singlet state (ignoring factors of ½)

$$P(\mathbf{a}, \mathbf{a}) = -1, \qquad \text{and if anti-parallel,} \qquad P(\mathbf{a}, -\mathbf{a}) = +1.$$

Then for arbitrary orientations, QM says

$$P(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b}.$$

Now let's add locality in the form of a hidden variable λ, which is the argument to functions $A(\mathbf{a}, \lambda)$ and $B(\mathbf{b}, \lambda)$ which will fix in advance the result of the spin measurements. Evidently,

$$A(\mathbf{a}, \lambda) = \pm 1 \qquad \text{and} \qquad B(\mathbf{b}, \lambda) = \pm 1.$$

Perfect alignment of the detectors forces

$$A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda).$$

Given any arbitrary probability density ρ, the average of the product of the measurements is

$$P(\mathbf{a}, \mathbf{b}) = \int \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) d\lambda.$$

Using only these equations and some simple algebra, one can imagine measurement along a third unit vector $\mathbf{c}$ and calculate the ***Bell inequality***[68]

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq 1 + P(\mathbf{b}, \mathbf{c}),$$

which must therefore hold for any local hidden-variable theory. But it doesn't. For instance, for the three vectors in a plane, $\mathbf{a}$ and $\mathbf{b}$ perpendicular to each other with $\mathbf{c}$ at a 45° angle between them,

$$P(\mathbf{a}, \mathbf{b}) = 0, \quad P(\mathbf{a}, \mathbf{c}) = P(\mathbf{b}, \mathbf{c}) = -0.707,$$

which is clearly inconsistent with Bell's equality. If QM is correct, then Bell's inequality is violated. Other forms of inequality have been calculated, forms more amenable to experiments, which have found it to be violated.

The in fact shocking result is that nature is ***non-local***, and no hidden-variable theory can save it from that. The supposed instantaneous "collapse" of the wave function in the Copenhagen interpretation was not enough, we now have non-locality.

67 Following from Griffiths, QM, 449-452.
68 Griffiths, 450-451.

9. Bibliography

Ballentine, Leslie E. *Quantum mechanics: A modern development*. 2nd edition. World Scientific, 2015. Print.

Griffiths, David J. and Schroeter, Darrell F. *Introduction to quantum mechanics*, 3rd edition. Cambridge: Cambridge University Press, 2018. Print.

Jeevanjee, Nadir. *An introduction to tensors and group theory for physicists*. 2nd edition. Heidelberg: Springer, 2015. Print.

Norsen, Travis. *Foundations of quantum mechanics*. Cham: Springer, 2012. Print.

The 7 basic rules of quantum mechanics, Physics forums, https://www.physicsforums.com/insights/the-7-basic-rules-of-quantum-mechanics/

Sakurai, J. J. and Napolitano, Jim. *Modern quantum mechanics*, 3rd edtition. Cambridg: Cambridge University Press, 2021. Print.

Susskind, Leonard and Friedman, Art. *Quantum mechanics, the theoretical minimum*. New York: Basic Books, 2014. Print.

Susskind, Leonard and Hrabovsky, George. *The theoretical minimum: What you need to know to start doing physics*. New York: Basic Books, 2013. Print.

Teller, Paul. *An interpretive introduction to quantum field theory*. Princeton: Princeton University Press, 1997. Print.

Townsend, John S. *A modern approach to quantum mechanics*. Sausalito: University Science Books, 2012. Print.