Advanced Quantum Mechanics

Notes by P.R. Kaye based on the lectures of A. Kempf

December 18, 2002

Contents

1 Notation

- Quantum mechanical quantities will be denoted with a *hat*: $\hat{}$.
- Vector quantities are denoted using bold face; e.g. $\mathbf{x} \equiv (x_1, x_2, x_3)$.
- Super/sub-scripts: $\hat{x}_i^{(r)}$ denotes the i-component ($i \in \{1, 2, 3\}$) of $\hat{\mathbf{x}}$ for particle r.
- $\hat{f}$ and $\hat{g}$ will always be assumed to be some polynomials in position and momentum variables $\hat{x}$ and $\hat{p}$.
- Time derivatives will sometimes be indicated with a *dot* (e.g. $\frac{dx}{dt} \equiv \dot{\hat{x}}$)
- $[\hat{x}, \hat{p}]$ denotes the commutator $\hat{x}\hat{p} - \hat{p}\hat{x}$.
- The Kronecker and Dirac deltas, $\delta_{m,n}$, and $\delta(x)$, are detailed in the Appendix.

2 From Classical Mechanics to Quantum Mechanics

2.1 Classical Mechanics:

In Classical Mechanics (CM) the equations of motion are derived via *Hamilton's Equation*. To prove this from more basic principles is beyond our scope here, so we just take this as a basic assumption. Suppose f is some physically observable quantity, which is a polynomial function of position x, momentum p (and possibly has an explicit dependence on t).

Physical Assumption 1 *The* equation of motion *governing the time-evolution of f follows Hamilton's equation*

$$\frac{d}{dt}f = \{f, H\} + \frac{\partial f}{\partial t}$$

where $\{,\}$ is the Poisson Bracket*.*

The term $\frac{\partial f}{\partial t}$ on the RHS of Hamilton's equation is the explicit time-dependence of f, that is, not accounting for the time-dependence through x and p (which are ultimately number-valued functions of t). In Hamilton's equation, x and p are treated as abstract *variables*, and f is some polynomial in x and p. In particular, this means that if f is dependent on time *only through* x and p, Hamilton's equation simplifies to

$$\frac{d}{dt}f = \{f, H\} \quad (*).$$

The Poisson bracket is a multiplication rule for polynomials in positions x and momenta p. It is defined by two sets of axioms:

A)
$$\{x_i^{(r)}, p_j^{(s)}\} = \delta_{i,j}\delta_{r,s}$$
$$\{x_i^{(r)}, x_j^{(s)}\} = 0$$
$$\{p_i^{(r)}, p_j^{(s)}\} = 0$$

B)
$$\{f, g\} = -\{g, f\}$$
$$\{cf, g\} = c\{f, g\}$$
$$\{f, g + h\} = \{f, g\} + \{f, h\}$$
$$\{f, gh\} = \{f, g\}h + g\{f, h\}$$

An important consequence of the axioms of $\{,\}$ is the *Jacobi Identity:*

$$\{f,\{g,h\}\}+\{h,\{f,g\}\}+\{g,\{h,f\}\}=0.$$

The choices for f in Hamilton's equation that are most important for classical mechanics are $f=x_i^{(r)}$ and $f=p_i^{(r)}$. For these choices of f Hamilton's equation $(*)$ becomes:

$$\frac{d}{dt}x_i^{(r)}=\{x_i^{(r)},H\} \qquad (1)$$
$$\frac{d}{dt}p_i^{(r)}=\{p_i^{(r)},H\} \qquad (2)$$

As a concrete example illustrating that the equations of motion for a system follow Hamilton's equation, consider a free particle of mass m possessing only kinetic energy. Its Hamiltonian is

$$H=\sum_{j=1}^{3}\frac{p_j^2}{2m}.$$

Using this H in equations (1) and (2) we obtain the following equations of motion:

$$\frac{d}{dt}x_i=\left\{x_i,\sum_{j=1}^{3}\frac{p_j^2}{2m}\right\}=\frac{p_i}{m}$$
$$\frac{d}{dt}x_i=\left\{p_i,\sum_{j=1}^{3}\frac{p_j^2}{2m}\right\}=0.$$

These agree with what is learned in high-school physics: $p_i=m\dot{x}_i$ and $\ddot{x}_i=0$.

As another concrete example, consider a system of two point-masses m_1,m_2 which are connected by a spring with spring-constant k. Its Hamiltonian is

$$H=\frac{\left(\mathbf{p}^{(1)}\right)^2}{2m_1}+\frac{\left(\mathbf{p}^{(2)}\right)^2}{2m_2}+\frac{k}{2}\left(\mathbf{x}^{(1)}-\mathbf{x}^{(1)}\right)^2.$$

Using this H in equations (1) and (2) give the equations of motion

$$\frac{d}{dt}x_i^{(r)}=\{x_i^{(r)},H\}=\frac{p_i^{(r)}}{m_r}$$
$$\frac{d}{dt}p_i^{(1)}=\{p_i^{(1)},H\}=-k\left(x_i^{(1)}-x_i^{(2)}\right)$$
$$\frac{d}{dt}p_i^{(2)}=\{p_i^{(2)},H\}=-k\left(x_i^{(2)}-x_i^{(1)}\right)$$

In CM, x,p are real-valued functions. This means that the Poisson bracket can be evaluated using the following "representation":

$$\{f,g\}=\frac{\partial f}{\partial x}\frac{\partial g}{\partial p}-\frac{\partial f}{\partial p}\frac{\partial g}{\partial x}$$

2.2 What's Different for Quantum Mechanics

Quantum Mechanics (QM) also obeys Hamilton's equation, and all the axioms of the Poisson algebra. What differs is the mathematical representations of position and momentum. The experimentally observed phenomenon of incompatible measurements suggests that position and momentum cannot be number-valued functions in QM. Suppose $\hat{x}$ and $\hat{p}$ denote the mathematical objects we use to represent position and momentum. The condition that $\hat{x}$ and $\hat{p}$ cannot be number valued functions suggests that $\hat{x}\hat{p} \neq \hat{p}\hat{x}$, or equivalently $[\hat{x}, \hat{p}] \neq 0$ (if a representation consistent with the Poisson algebra satisfied $[\hat{x}, \hat{p}] = 0$, it would be isomorphic to a number-valued representation). If $\hat{x}$ and $\hat{p}$ are not number-valued, then the representation of $\{,\}$ in terms of partial derivatives that we had for classical mechanics is no longer valid. So we need a new representation of $\{,\}$ that is both mathematically consistent with the axioms defining it, and is physically consistent with observed data. We begin by examining the constraints given by the requirement for mathematical consistency with the axioms of the Poisson algebra.

We evaluate $\{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\}$ in two ways, and require that the results agree.

$$\begin{aligned}\{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\} &= \hat{u}_1\{\hat{u}_2, \hat{v}_1\hat{v}_2\} + \{\hat{u}_1, \hat{v}_1\hat{v}_2\}\hat{u}_2 \\ &= \hat{u}_1\left(\hat{v}_1\{\hat{u}_2, \hat{v}_2\} + \{\hat{u}_2, \hat{v}_1\}\hat{v}_2\right) + \left(\hat{v}_1\{\hat{u}_1, \hat{v}_2\} + \{u_1, \hat{v}_1\}\hat{v}_2\right)\hat{u}_2 \qquad (1)\end{aligned}$$

and also

$$\begin{aligned}\{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\} &= \hat{v}_1\{\hat{u}_1\hat{u}_2, \hat{v}_2\} + \{\hat{u}_1\hat{u}_2, \hat{v}_1\}\hat{v}_2 \\ &= \hat{v}_1\left(\hat{u}_1\{\hat{u}_2, \hat{v}_2\} + \{\hat{u}_1, \hat{v}_2\}\hat{u}_2\right) + \left(\hat{u}_1\{\hat{u}_2, \hat{v}_1\} + \{u_1, \hat{v}_1\}\hat{u}_2\right)\hat{v}_2 \qquad (2)\end{aligned}$$

Equating (1) and (2) we get the requirement

$$\left(\hat{u}_1\hat{v}_1 - \hat{v}_1\hat{u}_1\right)\{\hat{u}_2, \hat{v}_2\} = \{\hat{u}_1, \hat{v}_1\}\left(\hat{u}_2\hat{v}_2 - \hat{v}_2\hat{u}_2\right).$$

This means we must have

$$\begin{aligned}\left(\hat{u}_1\hat{v}_1 - \hat{v}_1\hat{u}_1\right) &= K\{\hat{u}_1, \hat{v}_1\} \quad \text{and} \\ \left(\hat{u}_2\hat{v}_2 - \hat{v}_2\hat{u}_2\right) &= K\{\hat{u}_2, \hat{v}_2\}\end{aligned}$$

where K is some entity that is independent of the $\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2$. Since $\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2$ were arbitrary, in general we must have

$$K\{\hat{u}, \hat{v}\} = \hat{u}\hat{v} - \hat{v}\hat{u} \qquad (*)$$

for all $\hat{u}$, $\hat{v}$ polynomials in $\hat{x}$, $\hat{p}$.

Classical mechanics is obtained in the case $K = 0$. In quantum mechanics, based on experiments we have the following:

Physical Assumption 2 *Quantum mechanics is consistent with the choice $K = i\hbar$ (where $\hbar$ is Plank's constant) in equation $(*)$.*

So, in quantum mechanics we have

$$[\hat{f}, \hat{g}] = i\hbar\{\hat{f}, \hat{g}\}$$

which leads immediately to Heisenberg's *Canonical Commutation Relations (CCRs):*

$$\begin{aligned}[\hat{x}_i^{(r)}, \hat{p}_j^{(s)}] &= i\hbar\delta_{i,j}\delta_{r,s} \\ [\hat{x}_i^{(r)}, \hat{x}_j^{(s)}] = [\hat{p}_i^{(r)}, \hat{p}_j^{(s)}] &= 0\end{aligned}$$

and a representation of the Poisson bracket for QM:

$$\{\hat{f}, \hat{g}\} = \frac{1}{i\hbar}[\hat{f}, \hat{g}]$$

2.3 Heisenberg Equations

Using the above representation for the Poisson bracket, in QM Hamilton's equation is represented as

$$\frac{d}{dt}\hat{f} = \frac{1}{i\hbar}[\hat{f}, \hat{H}] + \frac{\partial f}{\partial t}.$$

The *Heisenberg equations* are the Hamilton equation for the cases $\hat{f} = \hat{x}$ and $\hat{f} = \hat{p}$:

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

The physics will also require $\hat{x} = \hat{x}^\dagger$, and $\hat{p} = \hat{p}^\dagger$ (*Hermiteicity*, or *reality* conditions). We say $\hat{f}$ is "Hermitean" when $\hat{f} = \hat{f}^\dagger$. We come back to exactly what "†" means when we look at representation theory.

Before proceeding, we record a fundamental physical assumption underlying our development of quantum mechanics.

Physical Assumption 3 *Any physically observable quantity can be expressed as a linear, Hermitean operator $\hat{f}$ which acts on a complex vector space (*Hilbert space*, defined later). Such an $\hat{f}$ is appropriately called an* observable *of the system.*

2.4 A Representation for $\hat{x}$, $\hat{p}$

The $\hat{x}$ and $\hat{p}$ can be represented in a way consistent with the development above, using infinite-dimensional square matrices.

Definition 1

$$a := \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Then let

$$\hat{x}(t_0) = L(a^\dagger + a)$$
$$\hat{p}(t_0) = \frac{i\hbar}{2L}(a^\dagger - a)$$
$$\text{where } L \in \mathbb{R}.$$

This gives $\hat{x}(t_0) = \hat{x}^\dagger(t_0)$, $\hat{p}(t_0) = \hat{p}^\dagger(t_0)$ and $[\hat{x}(t_0), \hat{p}(t_0)] = i\hbar$. As we shall see, it turns out that these properties are preserved under the *time-evolution* of a quantum system, and so they will hold in general for $\hat{x}(t)$ and $\hat{p}(t)$.

2.5 Dynamics

2.5.1 Time-Evolution Preserves the CCRs and Hermiteicity Requirement

Theorem 1 *The CCRs are preserved under time-evolution.*

Proof: We need to show $\frac{d}{dt}[\hat{x},\hat{p}] = 0$.

$$\begin{aligned}
\frac{d}{dt}[\hat{x},\hat{p}] &= [\dot{\hat{x}},\hat{p}] + [\hat{x},\dot{\hat{p}}] \quad \text{product rule} \\
&= \frac{1}{i\hbar}\left[[\hat{x},\hat{H}],\hat{p}\right] + \frac{1}{i\hbar}\left[\hat{x},[\hat{p},\hat{H}]\right] \\
&= \frac{1}{i\hbar}\left[\hat{p},[\hat{H},\hat{x}]\right] + \frac{1}{i\hbar}\left[\hat{x},[\hat{p},\hat{H}]\right] \\
&= -\frac{1}{i\hbar}\left[\hat{H},[\hat{x},\hat{p}]\right] \quad \text{(Jacobi Identity)} \\
&= 0 \quad \square
\end{aligned}$$

Theorem 2 *If $\hat{H}$ satisfies $\hat{H}^\dagger = \hat{H}$, then the Hermiteicity of $\hat{x}$ and $\hat{p}$ is preserved under time-evolution.*

Proof: We show that $\frac{d\hat{x}(t)}{dt}$ is Hermitean.

$$\begin{aligned}
\left(\frac{d\hat{x}(t)}{dt}\right)^\dagger &= \left(\frac{1}{i\hbar}[\hat{x}(t),\hat{H}]\right)^\dagger \\
&= -\frac{1}{i\hbar}[\hat{H},\hat{x}(t)] \quad (\text{use } (AB)^\dagger = B^\dagger \hat{A}^\dagger) \\
&= \frac{1}{i\hbar}[\hat{x}(t),\hat{H}] \\
&= \frac{d\hat{x}(t)}{dt} \qquad \square
\end{aligned}$$

The proof that $\frac{d\hat{p}(t)}{dt}$ is Hermitean is analogous.

2.5.2 The Time-Evolution Operator

Definition 2 *The* time-evolution operator $\hat{U}(t)$ *is the solution to*

$$i\hbar\frac{d}{dt}\hat{U}(t) = \hat{U}(t)\hat{H}(t) \quad , \quad \hat{U}(t_0) = \mathbf{1}$$

which is

$$\hat{U}(t) = Te^{\frac{1}{i\hbar}\int_{t_0}^{t}\hat{H}(t)dt}$$

where T is the time-ordering symbol*, needed to resolve ordering ambiguities arising from noncommutivity.*

Note: $\hat{U}(t)$ is unitary ($\hat{U}^\dagger(t) = \hat{U}^{-1}(t)$).

The time-evolution operator is simpler when $\frac{\partial\hat{H}}{\partial t} = 0$:

$$\hat{U}(t) = e^{\frac{1}{i\hbar}(t-t_0)\hat{H}}.$$

Once we have the time-evolution operator, it gives the solution to the equations of motion.

Theorem 3 *Assume $\hat{x}_i^{(r)}(t_0)$, $\hat{p}_i^{(r)}(t_0)$ obey the CCRs and are Hermitean. Then*

$$\begin{aligned}
\hat{x}_i^{(r)}(t) &= \hat{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\hat{U}(t) \quad (*) \\
\hat{p}_i^{(r)}(t) &= \hat{U}^\dagger(t)\hat{p}_i^{(r)}(t_0)\hat{U}(t)
\end{aligned}$$

are the full solutions to the equations of motion.

Proof: First, we will need the following expressions for $\dot{\hat{U}}$ and $\dot{\hat{U}}^\dagger$, which follow immediately from the definition of $\hat{U}$, and using the fact that $\hat{H}$ is Hermitean, and that $(AB)^\dagger = B^\dagger A^\dagger$:

$$\dot{\hat{U}}(t) = \frac{1}{i\hbar}\hat{U}\hat{H}$$
$$\dot{\hat{U}}^\dagger(t) = -\frac{1}{i\hbar}\hat{H}\hat{U}^\dagger$$

Now, we check that $\hat{x}_i^{(r)}(t)$ obeys the equation of motion.

$$\begin{aligned} i\hbar\frac{d}{dt}\hat{x}_i^{(r)}(t) &= i\hbar\frac{d}{dt}\left(\hat{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\hat{U}(t)\right) \\ &= i\hbar\left(\dot{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\hat{U}(t) + \hat{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\dot{\hat{U}}(t)\right) \\ &= -\hat{H}(t)\hat{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\hat{U}(t) + \hat{U}^\dagger(t)\hat{x}_i^{(r)}(t_0)\hat{U}(t)\hat{H}(t) \\ &= -\hat{H}(t)\hat{x}_i^{(r)}(t) + \hat{x}_i^{(r)}(t)\hat{H}(t) \\ &= [\hat{x}_i^{(r)}(t), \hat{H}(t)] \end{aligned}$$

The proof that $\hat{p}_i^{(r)}(t)$ obeys the equation of motion is similar. □

For any observable $\hat{f}$ (polynomial in $\hat{x}, \hat{p}$) which is Hermitean and obeys the CCRs at time t_0, Theorem 3 generalizes to give the equation for $\hat{f}(t)$:

$$\hat{f}(t) = \hat{U}^\dagger(t)\hat{f}(t_0)\hat{U}(t).$$

We can also recover the Hamiltonian from the time-evolution operator: $\hat{H} = i\hbar\hat{U}^\dagger(t)\frac{d}{dt}\hat{U}(t)$.

2.6 Measurement of Observables

We have used a matrix representation of $\hat{x}$, $\hat{p}$. We would like measurements of these quantities to yield real numbers. How do we compute these? In QM, we can compute the *expected values* $\overline{x}, \overline{p}$ of position and momentum measurements.

Physical Assumption 4 *A unit vector $|\psi\rangle$ (in a complex Hilbert space) can be chosen which contains enough information about the state of a quantum system at time t_0 to predict measurement expecation values. The expectation values are given by*

$$\overline{x}(t) = \langle\psi|\hat{x}(t)|\psi\rangle$$
$$\overline{p}(t) = \langle\psi|\hat{p}(t)|\psi\rangle$$
$$\overline{f}(t) = \langle\psi|\hat{f}(t)|\psi\rangle \text{ for any observable } \hat{f} \text{ satisfying } \hat{f}^\dagger = \hat{f}.$$

Notice that in this picture, the state-vector $|\psi\rangle$ describes the system at the initial time t_0, and it is the *operators* $\hat{f}$ describing *observable quantities* that vary with time.

Note: Observables are required to be Hermitean ($\hat{f}^\dagger = \hat{f}$). The Hermiteicity requirement ensures that measurement outcomes are *real* numbers.

The time-evolution of $\overline{x}(t)$, $\overline{p}(t)$ does *not* in general follow the equations of motion of classical mechanics.

Theorem 4 *(*Ehrenfest Theorem*) If $\hat{H}$ is a polynomial in the $\hat{x}$ and $\hat{p}$ variables of degree ≤ 2, then $\overline{x}(t)$ and $\overline{p}(t)$ obey the classical equations of motion.*

Proof: Suppose $\hat{H}$ is a polynomial in the $\hat{x}$ and $\hat{p}$ variables of degree ≤ 2. Then

$\dot{\hat{x}} = \{\hat{x}, \hat{H}\} =: g(\hat{x}, \hat{p})$ is linear in $\hat{x}, \hat{p}$, and

$\dot{\hat{p}} = \{\hat{p}, \hat{H}\} =: h(\hat{x}, \hat{p})$ is linear in $\hat{x}, \hat{p}$

$$\implies \quad \overline{g(\hat{x}, \hat{p})} = g(\overline{x}, \overline{p}) \text{ and } \overline{h(\hat{x}, \hat{p})} = h(\overline{x}, \overline{p}) \quad \square$$

We also define the variance of $\hat{f}(t)$

$$\nu(\hat{f}(t)) = \overline{\left(\hat{f}(t) - \overline{f}(t)\right)}^2 = \overline{\hat{f}^2(t)} - \left(\overline{\hat{f}(t)}\right)^2 = \langle\psi|\hat{f}^2(t)|\psi\rangle - \langle\psi|\hat{f}(t)|\psi\rangle^2$$

and the standard deviation $\Delta f = \sqrt{\nu(\hat{f}(t))}$.

2.7 The General Uncertainty Principle

Theorem 5 *Assume $\hat{f}, \hat{g}$ are Hermitean, and that a system has state-vector $|\psi\rangle$. Then $\Delta f \Delta g \geq \frac{1}{2}\left|\langle\psi|[\hat{f}, \hat{g}]|\psi\rangle\right|$.*

Proof: Recall from linear algebra: $\langle\psi|\psi\rangle \geq 0$. Choose

$$|\psi\rangle = \left(\hat{f} - \overline{f}1 + i\alpha(\hat{g} - \overline{g}1)\right)|\psi\rangle, \qquad \alpha \in \hat{R}.$$

Then $\langle\psi|\psi\rangle \geq 0$ gives:

$$\begin{aligned}
& \langle\psi|\left(\hat{f} - \overline{f}1 - i\alpha(\hat{g} - \overline{g}1)\right)\left(\hat{f} - \overline{f}1 + i\alpha(\hat{g} - \overline{g}1)\right)|\psi\rangle \geq 0 \\
\implies \quad & \langle\psi|(\hat{f} - \overline{f}1)^2|\psi\rangle + \alpha^2\langle\psi|(\hat{g} - \overline{g}1)^2|\psi\rangle + \alpha\langle\psi|i(\hat{f}\hat{g} - \hat{g}\hat{f})|\psi\rangle \geq 0 \\
\implies \quad & (\Delta f)^2 + \alpha^2(\Delta g)^2 + \alpha\langle\psi|i[\hat{f}, \hat{g}]|\psi\rangle \geq 0 \\
\implies \quad & (\Delta f)^2 + (\Delta g)^2\left(\alpha + \frac{\langle\psi|i[\hat{f}, \hat{g}]|\psi\rangle}{2(\Delta g)^2}\right)^2 - \frac{\left(\langle\psi|i[\hat{f}, \hat{g}]|\psi\rangle\right)^2}{\left(2(\Delta g)^2\right)^2}(\Delta g)^2 \geq 0
\end{aligned}$$

The above inequality is most stringent if we choose $\alpha = -\frac{\langle\psi|i[\hat{f}, \hat{g}]|\psi\rangle}{2(\Delta g)^2}$. Then we have

$$\begin{aligned}
& (\Delta f)^2 - \frac{\left(\langle\psi|i[\hat{f}, \hat{g}]|\psi\rangle\right)^2}{4(\Delta g)^2} \geq 0 \\
\implies \quad & \Delta f \Delta g \geq \frac{1}{2}\left|\langle\psi|[\hat{f}, \hat{g}]|\psi\rangle\right| \quad \square
\end{aligned}$$

Alternative Proof of the Uncertainty Principle: The uncertainty principle is often stated in the (equivalent) form:

$$(\Delta f)^2 (\Delta g)^2 \geq \left(\frac{1}{2i}\overline{[\hat{f}, \hat{g}]}\right)^2$$

which has the advantage that it allows the Cauchy-Schwartz inequality to be exploited, making the proof appear somewhat simpler. We have

$$(\Delta f)^2 = \overline{\left(\hat{f} - \overline{f}\right)}^2 = \langle\psi| \left(\hat{f} - \overline{f}\right)^2 |\psi\rangle = \langle F|F\rangle$$

where $|F\rangle \equiv \left(\hat{f} - \overline{f}\right)|\psi\rangle$ (note that since $\hat{f}$ is Hermitean, so is $\hat{f} - \overline{f}$. Similarly we have $(\Delta g)^2 = \langle G|G\rangle$. Now, apply the Cauchy-Schwartz inequality:

$$(\Delta f)^2 (\Delta g)^2 = \langle F|F\rangle\langle G|G\rangle \geq |\langle f|g\rangle|^2 .$$

For any complex number z,

$$|z|^2 = (\mathrm{Re}(Z))^2 + (\mathrm{Im}(z))^2 \geq (\mathrm{Im}(z))^2 = \left(\frac{1}{2i}(z - z^*)\right)^2$$

So, letting $z = \langle F|G\rangle$, we have

$$(\Delta f)^2 (\Delta g)^2 \geq \left(\frac{1}{2i}(\langle F|G\rangle - \langle G|F\rangle)\right)^2 .$$

Now,

$$\begin{aligned}
\langle F|G\rangle &= \langle\psi| \left(\hat{f} - \overline{f}\right)(\hat{g} - \overline{g})|\psi\rangle \\
&= \langle\psi| \left(\hat{f}\hat{g} - \hat{f}\overline{g} - \hat{g}\overline{f} + \overline{f}\,\overline{g}\right)|\psi\rangle \\
&= \langle\psi|\hat{f}\hat{g}|\psi\rangle - \overline{g}\langle\psi|\hat{f}|\psi\rangle - \overline{f}\langle\psi|\hat{g}|\psi\rangle + \overline{f}\,\overline{g}\langle\psi|\psi\rangle \\
&= \overline{fg} - \overline{f}\,\overline{g} - \overline{g}\overline{f} + \overline{f}\,\overline{g} \\
&= \overline{fg} - \overline{f}\,\overline{g}.
\end{aligned}$$

Similarly,

$$\langle G|F\rangle = \overline{gf} - \overline{f}\,\overline{g}.$$

Therefore

$$\langle F|G\rangle - \langle G|F\rangle = \overline{fg} - \overline{gf} = \overline{[\hat{f}, \hat{g}]}.$$

So we have

$$(\Delta f)^2 (\Delta g)^2 \geq \left(\frac{1}{2i}\overline{[\hat{f}, \hat{g}]}\right)^2 \qquad \square$$

An important special case of the general uncertainty principle is position-momentum uncertainty:

$$\Delta x(t)\Delta p(t) \geq \frac{\hbar}{2} \;\; \forall t.$$

2.8 The Time-Energy Uncertainty Relation

There is a *time-energy uncertainty relation*

$$\Delta t \Delta \hat{H} \geq \frac{\hbar}{2}$$

which is *not* a consequence of the general uncertainty principle above. What does the time-energy uncertainty mean? Time t is not an observable quantity like $\hat{f}$, so how do we define Δt? Any uncertain entity $\hat{f}$ takes

time for a change to be noticeable. A change is not noticeable until it is about one standard deviation Δf. So the time it takes is Δt where $\Delta t \left|\frac{d\overline{f}(t)}{dt}\right| = \Delta f$. That is

$$\Delta t = \frac{\Delta f(t)}{\left|\frac{d\overline{f}(t)}{dt}\right|}.$$

The time-energy uncertainty relation is then derived as follows. For Hermitean $\hat{f}$ we have

$$i\hbar\frac{d}{dt}\hat{f}(t) = [\hat{f}, \hat{H}] \tag{1}$$

$$\Delta f(t)\Delta\hat{H} \geq \frac{1}{2}\left|\langle\psi|[\hat{f}, \hat{H}]|\psi\rangle\right| \tag{2}$$

Sub (1) into (2):

$$\begin{aligned}
\Delta f(t)\Delta\hat{H} &\geq \frac{1}{2}\left|\langle\psi|i\hbar\frac{d}{dt}\hat{f}(t)|\psi\rangle\right| \\
&= \frac{\hbar}{2}\left|\frac{d}{dt}\langle\psi|\hat{f}(t)|\psi\rangle\right| \\
&= \frac{\hbar}{2}\left|\frac{d\overline{f}(t)}{dt}\right| \\
&= \frac{\hbar}{2}\frac{\Delta f(t)}{\Delta t} \\
\implies \Delta t\Delta\hat{H} &\geq \frac{\hbar}{2}
\end{aligned}$$

2.9 "Pictures" of Quantum Mechanics

We have used a matrix representation. Matrix multiplication is noncommutative, but it is associative. Depending on how we "bracket" the fundamental equations, we can get different "pictures". The pictures differ in which different quantities are considered constant, and which are time-dependent.

2.10 The Heisenberg Picture

What we have considered so far is the *Heisenberg Picture*. In this picture, the state vector $|\psi\rangle$ is constant, and the observables $\hat{f}(t)$ are time-dependent. We solve $i\hbar\frac{d}{dt}\hat{f}(t) = [\hat{f}(t), \hat{H}(t)] + \frac{\partial\hat{f}}{\partial t}$ directly and then calculate measurement-expectations $\overline{f}(t) = \langle\psi|\hat{f}(t)|\psi\rangle$.

2.11 The Schrödinger Picture

This is the picture most often used in practise for quantum mechanics. Recall the time-evolution operator method: $\hat{f}(t) = \hat{U}^\dagger(t)\hat{f}(t_0)\hat{U}(t)$. Substituting this into the equation for measurement-expectations gives

$$\overline{f}(t) = \langle\psi|\hat{U}^\dagger(t)\hat{f}(t_0)\hat{U}(t)|\psi\rangle \qquad (*).$$

Recall that in this expression $\hat{f}(t_0)$ is a constant matrix, and $|\psi\rangle$ is the (constant) initial state vector. In the Schrödinger picture we define time-dependent "Schrödinger" states:

$$|\psi_s(t)\rangle = \hat{U}(t)|\psi\rangle$$

which, when substituted in $(*)$ gives:

$$\overline{f}(t) = \langle\psi_s(t)|\hat{f}(t_0)|\psi_s(t)\rangle.$$

An important point is that when $\frac{\partial\hat{H}}{\partial t} = 0$, we can actually calculate $|\psi_s(t)\rangle$ without first having to solve for $\hat{U}(t)$. This is done by solving the *Schrödinger Equation*:

$$i\hbar\frac{d}{dt}|\psi_s(t)\rangle = \hat{H}_s(t)|\psi_s(t)\rangle.$$

where

$$\hat{H}_s(t) = \hat{U}(t)\hat{H}(t)\hat{U}^\dagger(t)$$

is called known as the *Schrödinger Hamiltonian.*

This formulation is especially simple for time-independent Hamiltonians.

Claim 1 *If* $\frac{\partial\hat{H}}{\partial t} = 0$ *then* $\hat{H}_s(t) = \hat{H}(t)$.

Proof: Recall that when $\frac{\partial\hat{H}}{\partial t} = 0$, the time-evolution operator has the simple form

$$\hat{U}(t) = e^{\frac{1}{i\hbar}(t-t_0)\hat{H}} = \sum_{m=0}^{\infty}\frac{\frac{1}{i\hbar}(t-t_0)^m}{m!}\hat{H}^m$$

and so

$$\begin{aligned} & [\hat{U}(t), \hat{H}] = 0 \\ \implies \quad & \hat{H}_s(t) = \hat{U}(t)\hat{H}(t)\hat{U}^\dagger(t) = \hat{H}\hat{U}(t)\hat{U}^\dagger(t) = \hat{H}(t) \qquad \square \end{aligned}$$

Derivation of the Schrödinger equation: Recall the definition of the time-evolution operator $i\hbar\frac{d}{dt}\hat{U}(t) = \hat{U}(t)\hat{H}(t)$. From the definition of the Schrödinger Hamiltonian we have $\hat{H}(t) = \hat{U}^\dagger(t)\hat{H}_s(t)\hat{U}\dagger(t)$. Putting these together we get

$$\begin{aligned} & i\hbar\frac{d}{dt}\hat{U}(t) = \hat{U}(t)\hat{U}^\dagger(t)\hat{H}_s(t)\hat{U}(t) \\ \implies \quad & i\hbar\frac{d}{dt}\hat{U}(t) = \hat{H}_s(t)\hat{U}(t) \qquad \text{(since } U \text{ is unitary)} \\ \implies \quad & i\hbar\frac{d}{dt}\hat{U}(t)|\psi\rangle = \hat{H}_s(t)\hat{U}(t)|\psi\rangle \\ \implies \quad & i\hbar\frac{d}{dt}|\psi_s(t)\rangle = \hat{H}_s(t)|\psi_s(t)\rangle \qquad \square \end{aligned}$$

2.12 The Dirac Picture

This is the picture most often used in practise for quantum field theory. It is sometimes called the "interaction picture". Assume $\hat{H}_s(t)$ is given, and suppose we can write

$$\hat{H}_s(t) = \hat{H}_0(t) + \hat{H}'(t)$$

such that $\hat{H}_0(t)$ is an easily solvable Hamiltonian (for example in quantum computing, $\hat{H}_0(t)$ is that of a quantum gate). By assumption, we can easily find the solution to $i\hbar\frac{d}{dt}\hat{U}_0(t) = \hat{H}_0(t)\hat{U}_0(t)$. So assume we

have $\hat{U}_0(t)$ explicitly. Now, define $\hat{U}'(t) := \hat{U}_0^\dagger(t)\hat{U}(t)$ so we have $\hat{U}(t) = \hat{U}_0(t)\hat{U}'(t)$. Then the equation for measurement-expectations can be written as

$$\begin{aligned}\overline{f}(t) &= \langle\psi|\hat{U}^\dagger(t)\hat{f}(t_0)\hat{U}(t)|\psi\rangle \\ &= \langle\psi|\hat{U}'^\dagger(t)\hat{U}_0^\dagger(t)\hat{f}(t_0)\hat{U}_0(t)\hat{U}'(t)|\psi\rangle \\ &= \left(\langle\psi|\hat{U}'^\dagger(t)\right)\left(\hat{U}_0^\dagger(t)\hat{f}(t_0)\hat{U}_0(t)\right)\left(\hat{U}'(t)|\psi\rangle\right)\end{aligned}$$

So, we define *operators in the Dirac picture* as:

$$\hat{f}_D(t) = \hat{U}_0^\dagger(t)\hat{f}(t_0)\hat{U}_0(t)$$

and *states in the Dirac picture as*:

$$|\psi_D(t)\rangle = \hat{U}'(t)|\psi\rangle$$

So, since $\hat{U}_0(t)$ is assumed easy, if we can calculate the Dirac states $|\psi_D(t)\rangle$ then we get expectation values $\overline{f}(t)$. The dynamics of $|\psi_D(t)\rangle$ is captured in the following theorem.

Theorem 6 *$|\psi_D(t)\rangle$ is the solution to*

$$i\hbar\frac{d}{dt}|\psi_D(t)\rangle = \hat{H}_D(t)|\psi_D(t)\rangle$$

with initial condition $|\psi_D(t_0)\rangle = |\psi\rangle$, where

$$\hat{H}_D(t) = \hat{U}_0^\dagger(t)\hat{H}'(t)\hat{U}_0(t).$$

Proof:

$$\begin{aligned}i\hbar\frac{d}{dt}\hat{U}(t) &= \hat{H}_s(t)\hat{U}(t) \\ i\hbar\frac{d}{dt}\left(\hat{U}_0\hat{U}'\right) &= \left(\hat{H}_0 + \hat{H}\prime\right)\hat{U}_0\hat{U}' \\ i\hbar\dot{\hat{U}}_0\hat{U}' + i\hbar\hat{U}_0\dot{\hat{U}}' &= i\hbar\dot{\hat{U}}_0\hat{U}_0^\dagger\hat{U}_0\hat{U}' + \hat{H}'\hat{U}_0\hat{U}' \\ i\hbar\hat{U}_0\dot{\hat{U}}' &= \hat{H}'\hat{U}_0\hat{U}' \\ i\hbar\dot{\hat{U}}' &= \hat{U}_0^\dagger\hat{H}'\hat{U}_0\hat{U}' \\ i\hbar\dot{\hat{U}}' &= \hat{H}_D\hat{U}' \\ i\hbar\frac{d}{dt}\hat{U}'|\psi\rangle &= \hat{H}_D\hat{U}'|\psi\rangle \\ i\hbar\frac{d}{dt}|\psi_D(t)\rangle &= \hat{H}_D(t)|\psi_D(t)\rangle \qquad \square\end{aligned}$$

Once we have $|\psi_D(t)\rangle$,

$$\overline{f}(t) = \langle\psi_D(t)|\hat{f}_D(t)|\psi_D(t)\rangle.$$

3 Representation Theory

We need to represent $\hat{x}$ and $\hat{p}$ as mathematical objects which are Hermitean and which obey the CCRs. Any mathematical object can be viewed as a map. In Quantum Mechanics the physically successful representations are as maps which are linear. We will look at linear maps in complex and unitary vector spaces in finite dimensions, and then consider what changes for these in infinite dimensions. A number of basic results in linear algebra are recalled; many without proof.

3.1 Dual Spaces

In the following, assume $\mathcal{H}$ is a vector space over the complex numbers $\mathbb{C}$, and vectors in this space are written in the Dirac notation (e.g. $|\psi\rangle$).

Definition 3 *Assume $\mathcal{H}$ is given. $\mathcal{H}^*$ is defined as the set of linear maps $\mathcal{H} \to \mathbb{C}$. We denote elements of $\mathcal{H}^*$ by $\langle\chi|$, where the action of $\langle\chi|$ is:*

$$\langle\chi| : |\psi\rangle \to \langle\chi|\psi\rangle \in \mathbb{C}$$

Theorem 7 *$\mathcal{H}^*$, (with the natural definitions for "+" and scalar multiplication) is a complex vector space, called the* dual vector space *of $\mathcal{H}$.*

Definition 4 *Assume $\mathcal{H}$, $\mathcal{H}^*$ possess a map "†", called "Hermitean conjugation":*

$$\dagger : \mathcal{H} \to \mathcal{H}^* \quad \text{and the inverse operation} \quad \dagger : \mathcal{H}^* \to \mathcal{H}$$

so that

$$|\psi\rangle^\dagger = \langle\psi| \quad , \quad \langle\psi|^\dagger = |\psi\rangle \quad \forall|\psi\rangle \in \mathcal{H}.$$

Assume that "†" obeys the axioms:

$$(|\psi\rangle + |\varphi\rangle)^\dagger = \langle\psi| + \langle\varphi|$$

$$(\lambda|\psi\rangle)^\dagger = \lambda^*\langle\psi|$$

$$\langle\varphi|\psi\rangle^* = \langle\psi|\varphi\rangle$$

$$\langle\psi|\psi\rangle = 0 \text{ iff } |\psi\rangle = \boldsymbol{0}$$

Then we say that $\mathcal{H}$ is a unitary vector space.

3.2 Bases

Assume that $\mathcal{H}$ is finite-dimensional.

Definition 5 *A set of vectors $\{|b_m\rangle\}_{m\in\zeta} \subseteq \mathcal{H}$, where ζ is some index set, is called an* orthonormal basis

$$\langle b_n|b_m\rangle = \delta_{n,m} \quad \forall m, n \in \zeta$$

and every $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = \sum_{n\in\zeta} \psi_n|b_n\rangle \ , \ \psi_n \in \mathbb{C}.$$

$\psi_n = \langle b_n|\psi\rangle$ are called the "coefficients of $|\psi\rangle$ in basis $\{|b_n\rangle\}$".

Theorem 8 *The set $\{\langle b_n|\}$ is an O.N. basis for $\mathcal{H}^*$ called the* dual basis.

Definition 6 *The map $\psi : \zeta \to \mathbb{C}$ which maps $n \to \psi_n$ is called the* wavefunction *of $|\psi\rangle$ in the basis $\{|b_n\rangle\}$.*

Remark: The same $|\psi\rangle$ has a different wavefunction for each choice of O.N. basis.

Theorem 9 *Choose an O.N. basis $\{b_n\}_{n\in\zeta}$. Then every linear map $\hat{T}$ can be written as*

$$\hat{T} = \sum_{n,m\in\zeta} \hat{T}_{n,m}|b_n\rangle\langle b_m|$$

The action of $\hat{T}$ is then

$$\hat{T} : |\psi\rangle \to \sum_{n,m\in\zeta} \hat{T}_{n,m}|b_n\rangle\langle b_m|\psi\rangle \quad \text{where} \quad \hat{T}_{n,m} = \langle b_n|\hat{T}|b_m\rangle.$$

The identity operator can be written as

$$\mathbf{1} = \sum_{n\in\zeta} |b_n\rangle\langle b_n| \quad (*).$$

$(*)$ is called the *resolution of the identity in the basis* $\{|b_n\rangle\}_{n\in\zeta}$.

Consider $\langle\psi|\hat{f}|\psi\rangle$ and insert the resolution of the identity twice:

$$\begin{aligned}\langle\psi|\hat{f}|\psi\rangle &= \sum_{n,m}\langle\psi|b_n\rangle\langle b_n|\hat{f}|b_m\rangle\langle b_m|\psi\rangle \\ &= \sum_{n,m}\psi_n^*\hat{f}_{n,m}\psi_m\end{aligned}$$

Theorem 10 *All matrix representations of the CCRs give the same physics, because they are all obtained in this way, by choosing an O.N. basis.*

3.3 † for Matrix Representations

$\langle\psi|\hat{A}|\psi\rangle$ means $\langle\psi|\hat{A}\,(|\psi\rangle)$.

Definition 7 *$\hat{A}^\dagger : \mathcal{H} \to \mathcal{H}$ is defined as that linear map which, if acting first on the left, will give the same scalar as A acting on the right:*

$$\left(\hat{A}^\dagger|\varphi\rangle\right)^\dagger |\psi\rangle = \langle\varphi|\hat{A}|\psi\rangle \quad , \quad \forall|\psi\rangle, |\varphi\rangle \in \mathcal{H}.$$

That is,

$$\langle\psi|\hat{A}^\dagger|\varphi\rangle^* = \langle\varphi|\hat{A}|\psi\rangle \quad (*).$$

If we insert the identity in $(*)$, we obtain $\hat{A}^\dagger_{m,n} = \hat{A}^*_{n,m}$ (i.e. † is matrix Hermitean conjugation).

3.4 Eigenvectors in Quantum Mechanics: State Collapse

Theorem 11 *If $|\psi\rangle$ is an eigenvector of a Hermitean $\hat{f}$ (an observable) with eigenvalue λ, at time t the measurement of $\hat{f}$ will surely yield λ.*

Suppose $\hat{f}|\psi\rangle = \lambda|\psi\rangle$. Then

$$\begin{aligned}\overline{f} &= \langle\psi|\hat{f}|\psi\rangle \\ &= \langle\psi|\lambda|\psi\rangle \\ &= \lambda\langle\psi|\psi\rangle \\ &= \lambda\end{aligned}$$

and

$$\begin{aligned}\nu(f) &= \langle\psi|\hat{f}^2|\psi\rangle - \langle\psi|\hat{f}|\psi\rangle^2 \\ &= \lambda^2 - \lambda^2 \\ &= 0 \quad \square\end{aligned}$$

The following theorem guarantees that the measurement results will be *real* numbers.

Theorem 12 *If $\hat{f} = \hat{f}^\dagger$ and if $\hat{f}|\psi\rangle = \lambda|\psi\rangle$ then $\lambda \in \mathbb{R}$.*

Suppose we have a system initially in state $|\psi\rangle$ and at time t we perform measurements of several observables $\hat{f}^{(1)}, \hat{f}^{(2)}, \ldots, \hat{f}^{(s)}$ at once. Recall the uncertainty principle

$$\Delta f^{(i)} \Delta f^{(j)} \geq \frac{1}{2}\left|\langle\psi|[\hat{f}^{(i)}, \hat{f}^{(i)}]|\psi\rangle\right|. \quad (*)$$

To allow for reproducibility of measurements we want $\Delta f^{(i)}, \Delta f^{(j)} = 0 \ \forall i, j$. So we need the RHS of $(*)$ to be 0. This means we can only measure simultaneously commuting sets of observables. The maximum number is $s = 3r$, (where r is the number of particles).

Suppose we prepare a state of r particles at time t_0, and perform $3r$ measurements at time t_1 obtaining the results $v_i \in \mathbb{R}$ for $\hat{f}^{(i)}$. We know (from experiment) that an immediate repeat measurement of the $\hat{f}^{(i)}$ at time $t_2 > t_1$ will yield the same numbers v_i. Starting from time t_1, the system has the initial state vector $|\varphi\rangle$ which is a joint eigenvector to all $\hat{f}^{(i)}$ with eigenvalues v_i. The old state $|\psi\rangle$ *collapses* to become the new state $|\varphi\rangle$ at time t_1. We record this fundamental observation here.

Physical Assumption 5 *If a measurement of an observable $\hat{f}$ of a quantum system in state $|\psi\rangle$ is made at time t, then the result of the measurement is some eigenvalue λ of $\hat{f}$, and the new state immediately after the measurement is the eigenvector of $\hat{f}$ corresponding to eigenvalue λ.*

3.5 The Spectral Theorem

The eigenvectors of a Hermitean operator corresponding to different eigenvalues are mutually orthogonal.

Theorem 13 *If $\hat{f} = \hat{f}^\dagger$ and if $\hat{f}|\psi_1\rangle = \lambda_1|\psi_1\rangle$ and $\hat{f}|\psi_2\rangle = \lambda_2|\psi_2\rangle$, and $\lambda_1 \neq \lambda_2$, then $\langle\psi_1|\psi_2\rangle = 0$.*

Hermitean f may possess many non-orthogonal eigenvectors with the same eigenvalue. The set of eigenvectors with a common eigenvalue λ is called the "eigenspace" of $\hat{f}$ corresponding to eigenvalue λ.

Theorem 14 *The eigenvalues of a Hermitean operator are all real.*

Theorem 15 *Any eigenspace is a unitary vector space.*

Definition 8 *Eigenvalues with eigenspaces larger than 1 are called "degenerate" eigenvalues.*

Definition 9 *We denote the set of eigenvectors of a Hermitean $\hat{f}$ by $\{|\hat{f}_n\rangle\}_{n\in G}$, $G \subset \mathbb{Z}$.*

Theorem 16 (Spectral Theorem) *For every Hermitean $\hat{f}$ there is at least one O.N. basis consisting of eigenvectors of $\hat{f}$.*

$\hat{f}$ is diagonal in it's own eigenbasis: $\hat{f} = \sum_n f_n |f_n\rangle\langle f_n|$.

Theorem 17 (Extended Spectral Theorem) *For every maximal set of commuting Hermitean operators $\hat{f}^{(1)}, \hat{f}^{(2)}, \ldots, \hat{f}^{(n)}$ there is a unique O.N. basis of $\mathcal{H}$ consisting of joint eigenvectors of all $\hat{f}^{(1)}, \ldots, \hat{f}^{(n)}$.*

We denote a joint eigenvector by $|k_1, k_2, \ldots k_{3r}\rangle$ so, $\hat{f}^{(i)}|k_1, k_2, \ldots k_{3r}\rangle = k_i|k_1, k_2, \ldots k_{3r}\rangle$.

We have a resolution of the identity:

$$\mathbf{1} = \sum_{n_1, n_2, \ldots, n_{3r}} |k_{n_1}, k_{n_2}, \ldots k_{n_{3r}}\rangle\langle k_{n_1}, k_{n_2}, \ldots k_{n_{3r}}|.$$

3.6 What's Different in ∞-Dimensions

Much of the theory developed above does not apply for infinite dimensional vector spaces. In finite-dimensional spaces, we found that any Hermitean operator has a complete (and orthonormal) set of eigenvectors in the space (spectral theorem). This is no longer true in infinite-dimensional spaces. We will develop a different version of the spectral theorem for infinite-dimensional spaces, in which Hermitean operators will have a complete set of eigenvectors which do not lie in the space, but which can be "orthonormalised" in a different sense.

The reason we investigate infinite-dimensional representations is that they arise naturally when we consider the eigenvectors of the position and momentum operators. We would like to use these sets of eigenvectors as orthonormal bases for our Hilbert space but, as we shall see later, this can only be done in the continuous sense. Also, while matrix equations are very intuitive, they are usually harder to solve than the differential equations that arise in continuous bases.

3.6.1 Vectors $|\psi\rangle$ in ∞-dimensions

For ∞-dimensional spaces, $\sum_{n=1}^{\infty} \psi_n^* \psi_n \not< \infty$ in general. So far we considered the unitary vector space of vectors which contain only a finite number of nonzero components. This is called a *pre-Hilbert space*, denoted $\mathcal{H}_p$.

Definition 10 *The completion $\mathcal{H}$ of a pre-Hilbert space $\mathcal{H}_p$ (obtained by adding the limit of every Cauchy sequence in $\mathcal{H}_p$) is called a* Hilbert Space

Note: In practise we take $\mathcal{H}$ to be the set of all vectors $|\psi\rangle$ such that $\sum_{n=1}^{\infty} \psi_n^* \psi_n < \infty$.

Definition 11 *A Hilbert space $\mathcal{H}$ is called* separable *if it has a countable O.N. basis $\{|b_n\rangle\}_{n\in G}$ (so that every $|\psi\rangle$ has a unique expansion $|\psi\rangle = \sum_{n\in G} \psi_n |b_n\rangle$). Such a basis is called a* Hilbert basis for $\mathcal{H}$.

Note: We know that the CCRs in QM can be represented using a separable Hilbert Space.

Theorem 18 *All separable Hilbert Spaces are isomorphic.*

Theorem 19 *There exist families of non-normalisable vectors* $\{|b_\lambda\rangle\}_{\lambda\in\zeta}$ *parameterized by a parameter* ζ *which takes values in a continuous subset* ζ *of* $\mathbb{R}$ *(e.g.* $\zeta = (1, 4]$*), which form an "improper O.N. basis for* $\mathcal{H}$*", so that* $\langle b_{\lambda'}|b_\lambda\rangle = \delta(\lambda - \lambda')$*, and every* $|\psi\rangle \in \mathcal{H}$ *can be expanded uniquely as:* $|\psi\rangle = \int_\zeta \psi(\lambda)|b_\lambda\rangle d\lambda$*.*

We have a resolution of the identity with respect to an improper O.N. basis: $\mathbf{1} = \int_\zeta |b_\lambda\rangle\langle b_\lambda| d\lambda$. A word about what the above theorem is saying. That the $\{|b_\lambda\rangle\}_{\lambda\in\zeta}$ is an *improper* basis means that the vectors $|b_\lambda\rangle$ are not in the Hilbert space $\mathcal{H}$. That these vectors are non-normalisable means that $\sum_n \langle b_\lambda|b_n\rangle\langle b_n|b_\lambda\rangle$ is divergent. These vectors *can*, however, be "continuously orthonormalised" so that $\langle b_\lambda|b_{\lambda'}\rangle = \delta(\lambda - \lambda')$.

Remark: We can also define a *mixed* O.N. basis $\{|b_n\rangle\}_{n\in G} \cup \{|b_\lambda\rangle\}_{\lambda\in\zeta}$ so that

$$1 = \sum_{n\in G} |b_n\rangle\langle b_n| + \int_\zeta |b_\lambda\rangle\langle b_\lambda| d\lambda.$$

3.6.2 Operators $\hat{A}|\psi\rangle$ in ∞-dimensions

Even if $\sum_{n\in G} \psi_n^* \psi_n < \infty$, it does not necessarily follow that $\psi = \hat{A}|\psi\rangle \in \mathcal{H}$ (because $\sum_{n=1}^\infty \psi_n^* \psi_n$ might be divergent).

Definition 12 *The maximal domain* $D_{\hat{A}}$ *for an operator* $\hat{A}$ *is :*

$$D_{\hat{A}} = \{|\psi\rangle \in \mathcal{H} : \hat{A}|\psi\rangle \in \mathcal{H}\}$$

$\hat{A}$ *is called* unbounded *if* $D_{\hat{A}} \subsetneq \mathcal{H}$

$\hat{A}$ *is called* bounded *if* $D_{\hat{A}} = \mathcal{H}$

3.6.3 Scalars $\langle\varphi|\hat{A}|\psi\rangle$ in ∞-dimensions

In general:

$$\lim_{N\to\infty}\lim_{M\to\infty}\sum_{n=1}^{N}\sum_{m=1}^{M}\varphi_n^*\hat{A}_{n,m}\psi_m \neq \lim_{M\to\infty}\lim_{N\to\infty}\sum_{n=1}^{N}\sum_{m=1}^{M}\varphi_n^*\hat{A}_{n,m}\psi_m$$

So we must be careful with our definition of $\hat{A}^\dagger$.

Definition 13 *Assume* $\hat{A}$ *and* $D_{\hat{A}}$ *are given. Then* $\hat{A}^\dagger$ *and* $D_{\hat{A}^\dagger}$ *are defined so that*

$$\left(\langle\varphi|\hat{A}^\dagger\right)|\psi\rangle = \langle\varphi|\hat{A}^\dagger|\psi\rangle \qquad \forall|\psi\rangle \in D_{\hat{A}}$$

Note: In general $D_{\hat{A}} \neq D_{\hat{A}^\dagger}$.

3.6.4 Spectral Theorem in ∞-dimensions

Definition 14 $\hat{U}$ *is called* unitary *if* $\hat{U}^\dagger = \hat{U}^{-1}$.

Definition 15 $\hat{f}$ *is called* self-adjoint *if* $\hat{f} = \hat{f}^\dagger$.

Definition 16 *$\hat{f}$ is called* symmetric *if* $\langle\psi|\hat{f}|\psi\rangle \in \mathbb{R} \quad \forall|\psi\rangle \in D_{\hat{f}}$.

Theorem 20 *If $\mathcal{H}$ is finite dimensional, then*

$$\hat{f} \text{ symmetric} \quad \Longleftrightarrow \quad \hat{f} \text{ self-adjoint.}$$

Theorem 21 *If $\mathcal{H}$ is ∞-dimensional, then*

$$\hat{f} \text{ self-adjoint} \quad \Longrightarrow \quad \hat{f} \text{ symmetric.}$$

Note: There is no universally agreed definition for "$\hat{f}$ is Hermitean".

Note: In linear algebra, "$\hat{A}$ symmetric" is defined as $\hat{A}_{n,m} = \hat{A}_{m,n}$, rather than as $\hat{A}_{n,m} = \hat{A}^*_{m,n}$.

Note: For $\hat{f}$ to be an observable, we need $\hat{f}$ is symmetric. Most observables are also self-adjoint.

Theorem 22 (Spectral Theorem) *Assume $\hat{f} = \hat{f}^\dagger$. Then $\hat{f}$ possess a (possibly improper) O.N. basis $\{|f_n\rangle\}_{n\in G} \cup \{|f_\lambda\rangle\}_{\lambda\in\zeta}$, $(\zeta \subset \mathbb{R})$ such that*

$$\hat{f}|f_n\rangle = f_n|f_n\rangle \quad \text{and} \quad \hat{f}|f_\lambda\rangle = f_\lambda|f_\lambda\rangle$$

and

$$\langle f_n|f_m\rangle = \delta_{n,m} \quad , \quad \langle f_\lambda|f_{\lambda'}\rangle = \delta(\lambda-\lambda') \quad , \quad \langle f_n|f_\lambda\rangle = 0.$$

Then, we can expand every $|\psi\rangle \in \mathcal{H}$ uniquely as:

$$|\psi\rangle = \sum_{n\in G} \psi_n|f_n\rangle + \int_\zeta \psi_\lambda|f_\lambda\rangle d\lambda$$

and

$$\psi_n = \langle f_n|\psi\rangle \quad , \quad \psi_\lambda = \langle f_\lambda|\psi\rangle.$$

Definition 17 *The set of proper and improper eigenvalues of $\hat{f}$ is called the* spectrum *of $\hat{f}$, denoted spec$(\hat{f})$.*

We can write $\hat{f}$ in it's own eigenbasis as

$$\hat{f} = \sum_{n\in G} f_n|f_n\rangle\langle f_n| + \int_\zeta f_\lambda\langle f_\lambda|f_\lambda\rangle d\lambda$$

and, of course, we have the resolution of the identity

$$\mathbf{1} = \sum_{n\in G} |f_n\rangle\langle f_n| + \int_\zeta \langle f_\lambda|f_\lambda\rangle d\lambda.$$

3.7 Sharpness of Measurement Predictions

If $\hat{f}$ is self-adjoint then by the spectral theorem we can prepare a system so that $\Delta f = 0$ by choosing an *eigenstate* (state vector which is an eigenvector of $\hat{f}$).

For measurement-expectations to be real, however, only requires that $\hat{f}$ be symmetric. There are important examples of observables which are symmetric, but not self-adjoint (e.g. in quantum optics). For such observables, *spec*$(\hat{f})$ may be the empty set. In this case Δf would be bounded below by some positive number Δf_{min}.

3.8 The Position and Momentum Eigenbases

We know from Physical Assumption 5 that if we measure observable $\hat{x}$ and obtain the result x, the state "collapses" to an eigenstate $|x\rangle$ of $\hat{x}$, with corresponding eigenvalue x. A similar thing happens for measurements of $\hat{p}$. If we make the reasonable assumption that a measurement of $\hat{x}$ or $\hat{p}$ can yield any real number, this suggests that every real number is an eigenvalue of $\hat{x}$ and of $\hat{p}$. So

$$spec(\hat{x}) = \mathbb{R} \quad , \quad spec(\hat{p}) = \mathbb{R}.$$

Since $\hat{x}$ and $\hat{p}$ are self-adjoint, the spectral theorem applies. It turns out that the eigenvectors of $\hat{x}$ are improper, and nondegenerate. So there is an improper orthonormal *position basis*

$$\{|x\rangle\}_{x\in\mathbb{R}}$$

so that

$$\langle x|x'\rangle = \delta(x'-x) \quad , \quad \mathbf{1} = \int_{\mathbb{R}} |x\rangle\langle x| dx.$$

Similarly, there is an improper orthonormal *momentum basis*

$$\{|p\rangle\}_{p\in\mathbb{R}}.$$

We can expand every vector in the Hilbert Space $\mathcal{H}$ in the position basis:

$$|\psi\rangle = \int_{\mathbb{R}} \psi(x)|x\rangle,$$

where $\psi(x) = \langle x|\psi\rangle$ is called the *wavefunction of* $|\psi\rangle$ *in the position basis*, or simply the *position wavefunction of* $|\psi\rangle$.

We have $|\psi\rangle \in \mathcal{H} \iff \langle\psi|\psi\rangle < \infty$ which, in the position basis, means

$$\int_{\mathbb{R}} \langle\psi|x\rangle\langle x|\psi\rangle dx = \int_{\mathbb{R}} \psi^*(x)\psi(x)dx < \infty.$$

It is natural to ask how the operator $\hat{x}$ acts on a *position wavefunction* $\psi(x)$. This is easy to calculate, starting with the action of $\hat{x}$ on the state vector ψ:

$$\begin{aligned}
&\hat{x}|\psi\rangle = |\varphi\rangle \\
\implies &\langle x|\hat{x}|\psi\rangle = \langle x|\varphi\rangle \\
&\text{insert the identity on the LHS:} \\
\implies &\int_{\mathbb{R}} \langle x|\hat{x}|x'\rangle\langle x'|\psi\rangle dx' = \langle x|\varphi\rangle \\
\implies &\int_{\mathbb{R}} x'\delta(x-x')\psi(x')dx' = \varphi(x) \\
\implies &x\psi(x) = \varphi(x).
\end{aligned}$$

So we find that $\hat{x}$ acts on position wavefunctions $\psi(x)$ as multiplication by x. What about $\hat{p}$? The computation is analogous, but relies on the following result (the proof of which we omit as it involves distribution theory):

Theorem 23

$$\langle x|\hat{p}|x'\rangle = i\hbar\frac{d}{dx'}\delta(x-x').$$

Note that, even though operators can not be described in an improper basis by a discrete matrix, we will refer, for example, to $\langle x|\hat{p}|x'\rangle$ as a *matrix element* of the $\hat{p}$ operator in the position basis.

Now,

$$\begin{aligned}
&\hat{p}|\psi\rangle = |\varphi\rangle \\
\implies &\langle x|\hat{p}|\psi\rangle = \langle x|\varphi\rangle \\
\implies &\int_{\mathbb{R}} \langle x|\hat{p}|x'\rangle\langle x'|\psi\rangle dx' = \langle x|\varphi\rangle \\
\implies &\int_{\mathbb{R}} i\hbar \frac{d}{dx'}\delta(x-x')\psi(x')dx' = \varphi(x) \\
&\text{integration by parts gives} \\
\implies &-i\hbar\frac{d}{dx}\psi(x) = \varphi(x).
\end{aligned}$$

So $\hat{p}$ acts on position wavefunctions $\psi(x)$ as $-i\hbar\frac{d}{dx}$.

3.9 Hermite and Fourier Transforms

We claimed that every real number x is an eigenvalue of $\hat{x}$, and that the corresponding eigenvectors satisfy $\langle x|x'\rangle = \delta(x-x')$. We can prove that this is the case, by solving $\hat{x}|x\rangle = x|x\rangle$ explicitly in a convenient basis. Choose the orthonormal basis $\{|n\rangle\}_{n\in\mathbb{N}}$ in which

$$\langle n|\hat{x}|m\rangle = L\begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}_{n,m}$$

and

$$\langle n|\hat{x}|m\rangle = L\begin{pmatrix} 0 & -\sqrt{1} & 0 & 0 & \cdots \\ \sqrt{1} & 0 & -\sqrt{2} & 0 & \cdots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}_{n,m}$$

We have $\langle n|\hat{x}|n\rangle = x\langle n|x\rangle$. Inserting the identity,

$$\sum_m \langle n|\hat{x}|m\rangle\langle m|x\rangle = x\langle n|x\rangle.$$

If we define $u_n(x) \equiv \langle n|x\rangle$ (so that $|x\rangle = \sum_n u_n(x)|n\rangle$), this becomes

$$\sum_m \langle n|\hat{x}|m\rangle u_m(x) = xu_n(x).$$

This is a matrix equation, from which we can read-off equations:

$$\begin{aligned}
L\sqrt{1}u_2(x) &= xu_1(x) \\
L\sqrt{1}u_1(x) + L\sqrt{2}u_3(x) &= xu_2(x) \\
&\vdots \\
L\sqrt{n-2}u_{n-2}(x) + L\sqrt{n-1}u_n(x) &= xu_{n-1}(x) \quad \forall n \in \mathbb{N}
\end{aligned}$$

which leads to the recursion relation

$$u_n(x) = \frac{1}{L\sqrt{n-1}}\left(xu_{n-1}(x) - L\sqrt{n-2}u_{n-2}(x)\right).$$

This recursion relation can be solved (we will not do it) to give

$$u_n(x) = \frac{1}{\sqrt{\sqrt{\pi}2^{n-1}(n-1)!}} e^{-\frac{x^2}{4L^2}} H_{n-1}\left(\frac{x}{\sqrt{2}L}\right)$$

where

$$H_n(z) = (-1)^n e^{z^2}\frac{d^n}{dz^n}e^{-z^2}.$$

So, we have solved explicitly for the eigenvectors $|x\rangle = \sum_{n\in\mathbb{N}} u_n(x)|n\rangle$. It can be shown that

$$\langle x|x'\rangle = \sum_{n\in\mathbb{N}}\langle x|n\rangle\langle n|x'\rangle = \sum_{n\in\mathbb{N}} u_n^*(x)u_n(x') = \delta(x-x').$$

A natural question is how we convert the wavefunctions between the position basis and the discrete bases; i.e. how do we convert between $\psi(x)$ and ψ_n? The answer is found (not surprisingly) by the "trick" of inserting the identity, together with the fact that $u_n(x)$ is self-adjoint:

$$\psi(x) = \langle x|\psi\rangle = \sum_n \langle x|n\rangle\langle n|\psi\rangle = \sum_n u_n(x)\psi_n \qquad (A)$$

and conversely

$$\psi_n = \langle n|x\rangle = \int_{\mathbb{R}} \langle n|x\rangle\langle x|\psi\rangle dx = \int_{\mathbb{R}} u_n(x)\psi(x)dx \qquad (B).$$

(A) and (B) are known as the *Hermite Transforms.*

Consider the momentum eigenbasis $\{|p\rangle\}_{p\in\mathbb{R}}$.

Definition 18 $\tilde{\psi}(p) = \langle p|\psi\rangle$ *is known as the* momentum wavefunction of $|\psi\rangle$.

So we have that $|\psi\rangle \in \mathcal{H}$ iff $\int_{\mathbb{R}} \tilde{\psi}^*(p)\tilde{\psi}(p)dp < \infty$, and every state $|\psi\rangle$ can be written as

$$|\psi\rangle = \int_{\mathbb{R}} |p\rangle\langle p|dp = \int_{\mathbb{R}} \tilde{\psi}(p)|p\rangle dp.$$

We have the following result

Theorem 24

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ixp}{\hbar}}.$$

A natural question is how to convert between position wavefunctions and momentum wavefunctions. Again, the answer is calculated by inserting the identity, and using the theorem stated above.

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \int_{\mathbb{R}} \langle p|x\rangle\langle x|\psi\rangle dx = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{ixp}{\hbar}}\psi(x)dx \qquad (C)$$

and conversely

$$\psi(x) = \langle x|\psi\rangle = \int_{\mathbb{R}} \langle x|p\rangle\langle p|\psi\rangle dp = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ixp}{\hbar}}\tilde{\psi}(p)dp \qquad (D)$$

(C) and (D) are known as the *Fourier Transform* and the *Inverse Fourier Transform* respectively.

3.10 Equivalence of Representations: the Stone/Von-Neumann Theorem

Any representation we choose must satisfy the CCRs; i.e. must satisfy $[\hat{x},\hat{p}] = i\hbar$. So, for all $|\psi\rangle$ we need

$$\hat{x}\hat{p}|\psi\rangle - \hat{p}\hat{x}|\psi\rangle = i\hbar|\psi\rangle \qquad (*).$$

In the discrete basis $\{|n\rangle\}$, we have $\hat{x},\hat{p}$ as matrices $L(a^\dagger + a)$ and $\frac{i\hbar}{2L}(a^\dagger - a)$ acting on vectors $(\psi_1,\psi_2,\ldots)^T$, and it is an easy matrix calculation to verify that $(*)$ is satisfied. In the position basis $\{|x\rangle\}$ we verify $(*)$, by first noting that

$$(*) \qquad \Longleftrightarrow \qquad \langle x|\hat{x}\hat{p}|\psi\rangle - \langle x|\hat{p}\hat{x}|\psi\rangle = i\hbar\underbrace{\langle x|\psi\rangle}_{=\psi(x)} \qquad (**)$$

for any position eigenvector $|x\rangle$. So it suffices to check $(**)$. We use the (now familiar) "trick" of inserting the identity a total of 4 times on LHS of $(**)$, and use integration by parts.

$$\begin{aligned}
&\langle x|\hat{x}\hat{p}|\psi\rangle - \langle x|\hat{p}\hat{x}|\psi\rangle \\
&\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} \langle x|\hat{x}|x'\rangle\langle x'|\hat{p}|x''\rangle\langle x''|\psi\rangle dx'dp' - \int_{\mathbb{R}}\int_{\mathbb{R}} \langle x|\hat{p}|x'\rangle\langle x'|\hat{x}|x''\rangle\langle x''|\psi\rangle dx'dp' \\
&\quad = x\left(-i\hbar\frac{d}{dx}\right)\psi(x) + i\hbar\frac{d}{dx}x\psi(x) \\
&\quad = x\left(-i\hbar\frac{d}{dx}\right)\psi(x) + i\hbar\psi(x) + x\left(i\hbar\frac{d}{dx}\right)\psi(x) \\
&\quad = i\hbar\psi(x) \\
&\quad = i\hbar\langle x|\psi\rangle
\end{aligned}$$

So $(**)$ is indeed satisfied.

We can do computations like the above (much) more quickly, without going through the tedious exercise of inserting the identity everywhere, by using the *rule*

> $\hat{x}$ acts on position wavefunctions $\psi(x)$ as multiplication by x, and
>
> $\hat{p}$ acts on position wavefunctions $\psi(x)$ as $-i\hbar\frac{d}{dx}$

which we proved in Section 3.8.

The representations of $[\hat{x},\hat{p}] = i\hbar$ in terms of matrices, and in terms of operators $\hat{x},\hat{p}$, differ only in a choice of basis for the Hilbert space $\mathcal{H}$. The following theorem says that this is so for *all* representations of $[\hat{x},\hat{p}] = i\hbar$.

Theorem 25 (Stone/von-Neumann) *Every representation of $\hat{x},\hat{p}$ obeying $[\hat{x},\hat{p}] = i\hbar$ as linear maps is obtainable as a choice of basis in the Hilbert space $\mathcal{H}$ that we have been working with.*

As another example, consider the representations of $\hat{x}$ and $\hat{p}$ in the momentum basis $\{|p\rangle\}$. Here

> $\hat{x}$ maps momentum wavefunctions $\tilde{\psi}(p)$ to $i\hbar\frac{d}{dp}\tilde{\psi}(p)$, and
>
> $\hat{p}$ maps momentum wavefunctions $\tilde{\psi}(p)$ to $p\tilde{\psi}(p)$.

4 Quantum Mechanics in the Schrödinger Picture, and the Position Basis

4.1 The Schrödinger Equation Revisited

We are now in a position to view the Schrödinger equation as it is often seen in introductory quantum mechanics books as the more general equation expressed in a particular choice of basis for $\mathcal{H}$. We will assume that our Hamiltonians are time-independent (i.e. $\frac{\partial \hat{H}}{\partial t} = 0$) so that the $\hat{H}_s = \hat{H}$, and we will denote the Hamiltonians simply by $\hat{H}$.

Recall the Schrödinger equation

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

For a free particle, with $\hat{H} = \frac{\hat{p}^2}{2m}$, this reads

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \frac{\hat{p}^2}{2m}|\psi(t)\rangle.$$

To write this in the position basis, replace $|\psi(t)\rangle$ with the position wavefunction $\psi(x,t)$, and use the rule for how $\hat{x}$ and $\hat{p}$ acts on position wavefunctions:

$$i\hbar \frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t).$$

Do the same thing for general Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x},\hat{p})$$

where V is a potential energy function. The Schrödinger equation for this Hamiltonian in the position basis is

$$i\hbar \frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2} + V\left(x, -i\hbar\frac{\partial}{\partial x}\right)\right)\psi(x,t).$$

This is the form of quantum mechanics as Schrödinger found it (Heisenberg found quantum mechanics in the discrete $\{|n\rangle\}$ basis.

4.2 Predicting Observables

Recall that in the Schrödinger picture, measurement-expectations obey

$$\overline{f}(t) = \langle\psi(t)|\hat{f}(\hat{x},\hat{p})|\psi(t)\rangle.$$

Using the rule for how $\hat{x}$ and $\hat{p}$ acts on position wavefunctions, we write this in the position basis as

$$\overline{f}(t) = \int_{\mathbb{R}} \psi^*(x,t)\hat{f}\left(x, -i\hbar\frac{\partial}{\partial x}\right)\psi(x,t)dx.$$

Recall that physical states $|\psi\rangle \in \mathcal{H}$ must obey $\langle\psi|\psi\rangle = 1$. In the position basis, this is expressed as $\int_{\mathbb{R}} \psi^*(x)\psi(x)dx = 1$.

Example: Consider a particle in 3 dimensions. We have $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{i,j}$, $i, j \in \{1, 2, 3\}$. We can use x_1, x_2, x_3 as a maximal set of commuting observables and find a joint eigenbasis $\{|x_1, x_2, x_3\rangle\}$. According to the continuous normalization, we have

$$\langle x_1, x_2, x_3 | x_1', x_2', x_3' \rangle = \delta(x_1 - x_1')\delta(x_2 - x_2')\delta(x_3 - x_3')$$
$$1 = \iiint |x_1, x_2, x_3\rangle\langle x_1, x_2, x_3| dx_1 dx_2 dx_3.$$

Any $|\psi\rangle$ can be expanded as

$$\begin{aligned} |\psi\rangle &= \iiint |x_1, x_2, x_3\rangle\langle x_1, x_2, x_3|\psi\rangle dx_1 dx_2 dx_3 \\ &= \iiint \psi(x_1 x_2 x_3) dx_1 dx_2 dx_3 \end{aligned}$$

where $\psi(x_1 x_2 x_3)$ is the position wavefunction. Suppose the particle is an electron orbiting a nucleus. Then the Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{\alpha}{\sqrt{\hat{\mathbf{x}}^2}}.$$

In the Schrödinger picture, and the position basis, $i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$ reads

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_1^2}\right) - \frac{\alpha}{|\hat{\mathbf{x}}|}\right)\psi(x,t). \quad (*)$$

To obtain the dynamics of the electron, we would need to solve $(*)$ with some initial condition $\psi(x, t_0) = \psi_0(x)$. Once $\psi(x_1, x_2, x_3, t)$ is found, we can predict all expectation values! For example

$$\overline{p_2(t)} = \iiint \psi^*(x_1, x_2, x_3, t)(-i\hbar)\frac{\partial}{\partial x_2}\psi(x_1, x_2, x_3, t) dx_1 dx_2 dx_3.$$

4.3 The Green Function Method for solving Schrödinger Equations

The "Green Method" prevents us from having to solve the Schrödinger equation separately for each initial condition $\psi(x, t_0) = \psi_0(x)$. The method is:

1. Solve
$$i\hbar\frac{\partial}{\partial t}G(x, x', t) = \hat{H}\left(x, -i\hbar\frac{\partial}{\partial x}\right)G(x, x', t)$$
with initial condition
$$G(x, x', t_0) = \delta(x - x').$$

2. For any initial condition $\psi_0(x)$ we obtain
$$\psi(x,t) = \int_{\mathbb{R}} G(x, x', t)\psi_0(x')dx'.$$

The function $G(x, x', t)$ is called the *Green function*, and sometimes called the *propagator*.

Proof that Green's Method Works:

$$\begin{aligned} & i\hbar\frac{\partial}{\partial t}\int_{\mathbb{R}} G(x,x',t)\psi_0(x')dx' \\ &= \int \hat{H}\left(x,-i\hbar\frac{\partial}{\partial x}\right)G(x,x',t)\psi_0(x')dx' \\ &= \hat{H}\left(x,-i\hbar\frac{\partial}{\partial x}\right)\int G(x,x',t)\psi_0(x')dx' \\ &= \hat{H}\left(x,-i\hbar\frac{\partial}{\partial x}\right)\psi(x,t) \quad \checkmark \end{aligned}$$

Also,

$$\psi(x,t_0) = \int \underbrace{G(x,x',t_0)}_{\delta(x-x')}\psi_0(x')dx' = \psi_0(x) \quad \checkmark$$

Now, recall the time-evolution operator $\hat{U}(t)$. The problem of solving the Schrödinger equation

$$\begin{aligned} i\hbar\frac{d}{dt}|\psi(t)\rangle &= \hat{H}|\psi(t)\rangle \\ |\psi(t_0)\rangle &= |\psi_0\rangle \end{aligned}$$

becomes that of solving

$$\begin{aligned} i\hbar\frac{d}{dt}\hat{U}(t) &= \hat{H}\hat{U}(t) \\ \hat{U}(t_0) &= 1 \end{aligned}$$

Then we have $|\psi(t)\rangle = \hat{U}(t)|\psi_0\rangle$. Inserting everywhere the identity resolved in the position basis we would obtain

$$G(x,x',t) = \langle x|\hat{U}(t)|x'\rangle.$$

So we see that $G(x,x',t)$ are the matrix elements of the time-evolution operator $\hat{U}(t)$ in the position basis.

It is natural to wonder what is the physical meaning of the propagator. This question takes us to a formulation of quantum mechanics due to Feynman.

5 Feynman's Formulation of Quantum Mechanics

For convenience, choose the Schrödinger picture. At time t_i suppose the initial state of a quantum system is $|\psi\rangle$. Then $|\psi(t)\rangle = \hat{U}(t)|\psi\rangle$.

Suppose at time t_f we use a measurement device which outputs 1 if the system is found in state $|\varphi\rangle$ and outputs 0 otherwise. This device is described by a self-adjoint operator $\hat{P}_{|\varphi\rangle}$ having eigenvalues 0 and 1. The eigenspace corresponding to eigenvalue 1 is spanned by $|\varphi\rangle$. All vectors $|\nu\rangle$ which are orthogonal to $|\varphi\rangle$ are in the eigenspace corresponding to eigenvalue 0. The above observations tell us that $\hat{P}_{|\varphi\rangle}$ is the projector

$$\hat{P}_{|\varphi\rangle} = |\varphi\rangle\langle\varphi|.$$

Decompose the final state as $|\psi(t_f)\rangle = \alpha|\varphi\rangle + |\nu\rangle$ such that $|\nu\rangle\perp|\varphi\rangle$. We need

$$\begin{aligned} \langle\varphi|\nu\rangle = 0 &= \langle\varphi|\psi(t_f)\rangle - \alpha\underbrace{\langle\varphi|\varphi\rangle}_{=1} \\ \implies \quad \alpha &= \langle\varphi|\psi(t_f)\rangle \end{aligned}$$

So

$$|\psi(t_f)\rangle = \alpha|\varphi\rangle + |\nu\rangle$$

Now consider the wavefunction collapse resulting from the measurement. If the measurement outputs 1, then the new state immediately after t_f is $|\varphi\rangle$. If the measurement outputs 0, the state immediately after t_f is $k|\nu\rangle$, where k is chosen so the state is normalised.

Now consider an ensemble of identically prepared systems, and running the experiment separately on each system. In general, we will have many 0's and many 1's in the collection of measurement outputs. The mean of these numbers is equal to the probability of finding $|\varphi\rangle$ at time t_f on one particular run of the experiment. But, from the theory we have developed, we know this probability can be calculated as

$$\overline{P}_{|\varphi\rangle}(t_f) = \langle\psi(t_f)|\hat{P}_{|\varphi\rangle}|\psi(t_f)\rangle = \langle\psi(t_f)||\varphi\rangle\langle\varphi||\psi(t_f)\rangle = |\langle\varphi|\psi(t_f)\rangle|^2 .$$

The quantity $\langle\varphi|\psi(t_f)\rangle$ is called a *probability amplitude* or *transition amplitude*. If $|\varphi\rangle$ is improper (for example $|\varphi\rangle = |x\rangle$), then $\langle x|\varphi(t)\rangle = \varphi(x,t)$ is a *probability density function* and, for example,

$$\begin{aligned}\int_{x_a}^{x_b} |\varphi(x,t)|^2\, dx &= \int_{x_a}^{x_b} \psi^*(x,t)\psi(x,t)dx \\ &= \text{prob. of finding the particle in the interval } (x_a, x_b).\end{aligned}$$

Recall the time-evolution operator $\hat{U}(t)$. We have $\langle\varphi|\psi(t)\rangle = \langle\varphi|U(t)|\psi\rangle$, and so we see that the probability amplitudes are just the "matrix elements" of the time-evolution operator.

Recall the Green function $G(x,x',t) = \langle x|\hat{U}(t)|x'\rangle$. In light of the above, we see that $G(x,x',t)$ is the probability amplitude for a particle to propagate from x' to x.

Now, Feynman compared classical probability theory to the laws of probability dictated by quantum mechanics. Consider a classical system going from state A to state B with some probability $P(A \to B)$. Suppose there are n possible paths from A to B, each passing through one of the n states $C_1, \dots C_n$. Then the (classical) law of conditional probability dictates that

$$P(A \to B) = \sum_{i=1}^{n} P(A \to C_i)P(C_i \to B).$$

Now consider an isolated, unobserved quantum system with Hamiltonian $\hat{H}$, starting in state $|\psi_0\rangle$ at time t_0 and evolving until time t_f. Consider the probability amplitude $\langle\varphi|\psi(t_f)\rangle$ for the state being $|\varphi\rangle$ at time t_f. Suppose that to get to state $|\varphi\rangle$, the system would have to evolve through one of n states $|b_1\rangle, \dots, |b_n\rangle$ at time t_1 ($t_0 < t_1 < t_f$). Then we have

$$\begin{aligned}\langle\varphi|\psi(t_f)\rangle = \langle\varphi|\underbrace{\hat{U}(t_f,t_0)}_{e^{\frac{t_f-t_0}{i\hbar}\hat{H}}}|\psi_0\rangle &= \langle\varphi|\underbrace{\hat{U}(t_f,t_1)}_{e^{\frac{t_f-t_1}{i\hbar}\hat{H}}}\underbrace{\hat{U}(t_1,t_0)}_{e^{\frac{t_1-t_0}{i\hbar}\hat{H}}}|\psi_0\rangle \\ &= \sum_{i=1}^{n}\langle\varphi|\hat{U}(t_f,t_1)|b_i\rangle\langle b_i|\hat{U}(t_1,t_i)|\psi_0\rangle.\end{aligned}$$

Note that $\langle\varphi|\hat{U}(t_f,t_1)|b_i\rangle$ is the probability amplitude for the evolution $|b_i\rangle \to |\varphi\rangle$ and $\langle b_i|\hat{U}(t_f,t_1)|\psi_0\rangle$ is the probability amplitude for the evolution $|\psi_0\rangle \to |b_i\rangle$. Feynman observed that in quantum mechanics, the law of conditional probabilities applies to *probability amplitudes*, instead of simple probabilities. This is significant, because probability amplitudes can have phase factors (so in particular, the could be negative!). This is why quantum interference can happen.

6 Density Matrices and Mixed States

6.1 Density Matrices and Mixed States

So far, we have assumed that a state $|\psi\rangle$ of a system can be known with certainty. Then our expectation values are $\overline{f} = \langle\psi|\hat{f}|\psi\rangle$. A more realistic situation is that we do not know the state of the system with certainty, but rather know that it is in one of the states $|b_1\rangle, \ldots, |b_n\rangle$ with respective probabilities $p_1, \ldots, p_n$. Without loss of generality we can assume that the $|b_i\rangle$ form an O.N. basis (otherwise we could extend the set by adding new states with corresponding probabilities equal to zero), and that $\sum_{i=1}^{n} p_i = 1$ (we know for certain the system is in one of the n possible states). In this case, the expectation values are computed through

$$\overline{f} = \sum_{i=1}^{n} p_i \langle b_i|\hat{f}|b_i\rangle.$$

Note that the uncertainty in f now has two sources:

a) our uncertainty about which state the system is in, and

b) the quantum-mechanical uncertainty.

Definition 19 *We say that a system is in a* Mixed State *if more than one p_i is nonzero. Otherwise the system is said to be in a* Pure State.

Definition 20

$$\hat{\rho} = \sum_i p_i |b_i\rangle\langle b_i|$$

is called the Density Matrix *for the system.*

Definition 21 *The trace of an operator $\hat{A}$ is*

$$Tr(\hat{A}) = \sum_n \langle c_n|\hat{A}|c_n\rangle$$

where $\{|c_n\rangle\}$ is any O.N. basis for the space of $\hat{A}$.

Note: $\hat{\rho}$ is self-adjoint, and $Tr(\hat{\rho}) = 1$ (proof follows from cyclicity of trace).

Now, the equation $\overline{f} = \sum_i p_i \langle b_i|\hat{f}|b_i\rangle$ becomes

$$\overline{f} = Tr(\hat{\rho}\hat{f}).$$

Suppose we choose a basis $\{|c_n\rangle\}$. Then $\overline{f} = Tr(\hat{\rho}\hat{f})$ becomes

$$\overline{f} = \sum_{n,m} \langle c_n|\hat{\rho}|c_m\rangle\langle c_m|\hat{f}|c_n\rangle.$$

We can define coefficient matrices for $\hat{\rho}$ and $\hat{f}$ in the $\{|c_n\rangle\}$ basis:

$$\hat{\rho}_{n,m} = \langle c_n|\hat{\rho}|c_m\rangle \quad , \quad \hat{f}_{n,m} = \langle c_n|\hat{f}|c_m\rangle.$$

Then $\overline{f} = \sum_{n,m} \hat{\rho}_{n,m}\hat{f}_{n,m}$.

Remarks:

We can choose, equivalently, improper or mixed O.N. bases in all of the above.

For a mixed state, $\hat{\rho}$ is non-diagonal in almost all bases. But, because $\hat{\rho}$ is self-adjoint, there always exists a basis in which it is diagonal (the eigenbasis). Only then does $\hat{\rho}$ have the form

$$\hat{\rho} = \sum_n p_n |b_n\rangle\langle b_n|.$$

In general bases we have

$$\hat{\rho} = \sum_{n,m} \hat{\rho}_{n,m} |c_n\rangle\langle c_m|.$$

We know that the time-evolution of Schrödinger states $|\psi(t)\rangle$ obeys the Schrödinger equation. This leads us to a result for the time-evolution of density matrices.

Theorem 26 *The time-evolution of the density matrix for a closed system follows the von-Neumann eqution:*

$$i\hbar \frac{d}{dt}\hat{\rho}(t) = -\left[\hat{\rho}(t), \hat{H}\right].$$

Proof: Write

$$\hat{\rho}(t_0) = \sum_n p_n |b_n\rangle\langle b_n|.$$

From the Schrödinger equation:

$$i\hbar \frac{d}{dt}|b_n(t)\rangle = \hat{H}(t)|b_n(t)\rangle$$

and thus

$$-i\hbar \frac{d}{dt}\langle b_n(t)| = \langle b_n(t)|\hat{H}(t) \quad (\text{since } \hat{H}^\dagger(t) = \hat{H}(t)).$$

At time t, each $|b_n(t_0)\rangle$ has evolved to $|b_n(t)\rangle$. So $\hat{\rho}(t_0)$ has evolved to

$$\hat{\rho}(t) = \sum_n p_n |b_n(t)\rangle\langle b_n(t)| \quad (*).$$

Apply $i\hbar\frac{d}{dt}$ to both sides of $(*)$:

$$\begin{aligned}
i\hbar \frac{d}{dt}\hat{\rho}(t) &= i\hbar \frac{d}{dt} \sum_n p_n |b_n(t)\rangle\langle b_n(t)| \\
&= \sum_n p_n \left(i\hbar \frac{d}{dt} \left(|b_n(t)\rangle\langle b_n(t)| \right) \right) \\
&= \sum_n p_n \left(i\hbar \left(\frac{d}{dt}|b_n(t)\rangle \right) \langle b_n(t)| + i\hbar |b_n(t)\rangle \left(\frac{d}{dt}\langle b_n(t)| \right) \right) \\
&= \sum_n p_n \left(i\hbar \left(\frac{d}{dt}|b_n(t)\rangle \right) \langle b_n(t)| - |b_n(t)\rangle \left(-i\hbar \frac{d}{dt}\langle b_n(t)| \right) \right) \\
&= \sum_n p_n \left(\hat{H}(t)|b_n(t)\rangle\langle b_n(t)| - |b_n(t)\rangle\langle b_n(t)|\hat{H}(t) \right) \\
&= \hat{H}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{H}(t) \\
&= -\left[\hat{\rho}(t), \hat{H}\right] \qquad \square
\end{aligned}$$

Note: We also have $\hat{\rho}(t) = U(t)\hat{\rho}\hat{U}(t)^\dagger$.

Review: changing bases for operators

How do we convert $\hat{\rho}$ from the $\{|c_n\rangle\}$ basis to a $\{|d_n\rangle\}$ basis? We use the old trick of inserting the identity resolved in the new basis:

$$\begin{aligned}\hat{\rho} &= \sum_{n,m,r,s} \hat{\rho}_{n,m}|d_r\rangle\langle d_r|c_n\rangle\langle c_m|d_s\rangle\langle d_s| \\ &= \sum_{r,s} \widetilde{\rho}_{r,s}|d_r\rangle\langle d_s|\end{aligned}$$

where $\widetilde{\rho}_{r,s} = \sum_{n,m} \hat{\rho}_{n,m}\langle d_r|c_n\rangle\langle c_m|d_s\rangle$.

6.2 Example: quantum system $\mathcal{S}$ in a heat bath

Suppose a quantum system $\mathcal{S}$ is in a heat bath at temperature T. According to statistical physics, the probability for a subsystem of a heat bath to be at energy E is proportional to

$$\sim e^{-\frac{E}{kT}}$$

where $k = 1.38 \cdot 10^{-23} J/Kelvin$ is the *Boltzman Constant*.

Assume the system $\mathcal{S}$ has nondegenerate energy eigenvalues $\{E_n\}$. Then

$$prob(E_n) = Ne^{-\frac{E_n}{kT}},$$

where $N = \frac{1}{\sum_m e^{-\frac{E_m}{kT}}}$. In other words, the probability is

$$p_n = prob(E_n) = \frac{e^{-\frac{E_n}{kT}}}{\sum_m e^{-\frac{E_m}{kT}}}.$$

Choose the energy eigenbasis of the system. Then the density matrix is

$$\hat{\rho} = \sum_n p_n|E_n\rangle\langle E_n|.$$

Now, since $\{E_m\}$ are the eigenvalues of the Hamitonian $\hat{H}$, we have that $\hat{H}$ is diagonal in the basis of its eigenvectors $|E_m\rangle$, and so

$$e^{-\frac{\hat{H}}{kT}} = \sum_m e^{-\frac{E_m}{kT}}|E_m\rangle\langle E_m|$$

Using this observation, and the definition of Trace, we get

$$\begin{aligned}Tr(e^{-\frac{\hat{H}}{kT}}) &= \sum_n \langle E_n|e^{-\frac{\hat{H}}{kt}}|E_n\rangle \\ &= \sum_n \langle E_n|\sum_m e^{-\frac{E_m}{kT}}|E_m\rangle\langle E_m|E_n\rangle \\ &= \sum_{n,m} \langle E_n|E_m\rangle e^{-\frac{E_m}{kT}}\langle E_m|E_n\rangle \\ &= \sum_n e^{-\frac{E_n}{kT}}.\end{aligned}$$

Therefore we can write the density matrix for the system in the Heat bath as

$$\hat{\rho} = \frac{1}{Tr(e^{-\frac{\hat{H}}{kT}})} e^{-\frac{\hat{H}}{kT}}.$$

6.3 Decoherence and Mixed States

Consider a quantum system S_c in a pure state, evolving in time. If S_c is completely isolated from it's environment, then the time-evolution is unitary, and the state of S_c remains pure. If S_c is allowed to interact with the environment at time t_1, then the interaction carries away some information about the state of S_c, and in this sense constitutes a measurement of S_c (whether or not this interaction was intended as a controlled measurement). Thus the state of S_c collapses as a result of the interaction. Since we don't learn the outcome of this measurement, we don't know which pure state S_c is in, and so S_c is left in a mixed state $\hat{\rho}^{S_c}$ at time t_1. Suppose we know that the interaction with the environment was, in particular, with a subsystem of the environment which we will label S_b. Imagine S_b is a photon which comes in from the environment, bounces off S_c, and returns to the environment, carrying with it some information about the state of S_c. Let us investigate how to compute the mixed state $\hat{\rho}^{S_c}$.

Assume S_c has N particles. Assume S_c is in a pure state $|\psi^{(S_c)}\rangle$ at time t_0. We have $\hat{x}_i^{(j)}, \hat{p}_i^{(j)}$, where $i \in \{1,2,3\}$, $j \in \{1,2,\ldots,N\}$. The system S_c has some Hamiltonian $\hat{H}^{(S_c)}\left(\hat{x}_i^{(j)}, \hat{p}_i^{(j)}\right)$. Choose a maximal set of $3N$ commuting observables (self-adjoint operators), say $\hat{Q}^{(1)}, \ldots, \hat{Q}^{(3N)}$. Find a joint eigenbasis of all the $\hat{Q}^{(j)}$ spanning $\mathcal{H}_{S_c}$:

$$\{|q_{n_1}^{(1)}, q_{n_2}^{(2)}, \ldots, q_{n_{3N}}^{(3N)}\rangle\},$$

where $q_{n_i}^{(j)}$ are the eigenvalues of $\hat{Q}^{(j)}$. The pure state of S_c can be expanded in this basis as

$$|\psi^{(S_c)}\rangle = \sum_{n_1,\ldots,n_{3N}} \psi^{(S_c)}_{n_1,\ldots,n_{3N}} |q_{n_1}^{(1)}, q_{n_2}^{(2)}, \ldots, q_{n_{3N}}^{(3N)}\rangle$$

and any observable $\hat{f}$ can be expanded as

$$\hat{f} = \sum_{\substack{n_1,\ldots,n_{3N},\\ \tilde{n}_1,\ldots,\tilde{n}_{3N}}} \hat{f}_{n_1,\ldots,n_{3N},\tilde{n}_1,\ldots,\tilde{n}_{3N}} |q_{\tilde{n}_1}^{(1)}, q_{\tilde{n}_2}^{(2)}, \ldots, q_{\tilde{n}_{3N}}^{(3N)}\rangle\langle q_{n_1}^{(1)}, q_{n_2}^{(2)}, \ldots, q_{n_{3N}}^{(3N)}|.$$

Expectation values are calculated as

$$\begin{aligned}\overline{f} &= \langle\psi^{(S_c)}|\hat{f}|\psi^{(S_c)}\rangle \\ &= \sum_{\substack{n_1,\ldots,n_{3N},\\ \tilde{n}_1,\ldots,\tilde{n}_{3N}}} \psi^{*(S_c)}_{\tilde{n}_1,\ldots,\tilde{n}_{3N}} \hat{f}_{n_1,\ldots,n_{3N},\tilde{n}_1,\ldots,\tilde{n}_{3N}} \psi^{(S_c)}_{n_1,\ldots,n_{3N}}.\end{aligned}$$

Now, consider the subsystem of the environment, S_b. Assume S_b has M particles. We have operators $\hat{x}_i^{(j)}, \hat{p}_i^{(j)}$, where $i \in \{1,2,3\}$, $j \in \{N+1, N+2, \ldots, N+M\}$, and a Hamiltonian $\hat{H}^{(S_b)}\left(\hat{x}_i^{(j)}, \hat{p}_i^{(j)}\right)$. Choose a maximal set of $3M$ observables, say $\hat{R}^{(1)}, \ldots, \hat{R}^{(3M)}$. Obtain a joint orthonormal eigenbasis of all $\hat{R}^{(i)}$:

$$\{|r_{m_1}^{(1)}, r_{m_2}^{(2)}, \ldots, r_{m_{3M}}^{(3M)}\rangle\}.$$

Assume S_b is in a pure state $|\varphi^{(S_b)}\rangle \in \mathcal{H}_{S_b}$, and expand as

$$|\varphi^{(S_b)}\rangle = \sum_{m_1,\ldots,m_{3M}} \varphi^{(S_b)}_{m_1,\ldots,m_{3M}} |r_{m_1}^{(1)}, r_{m_2}^{(2)}, \ldots, r_{m_{3M}}^{(3M)}\rangle.$$

Now, consider the total system $S_{tot} = S_c \cup S_b$. We have operators $\hat{x}_i^{(j)}, p_i^{(j)}$, where $i \in \{1,2,3\}, j \in \{1,\ldots,N+M\}$. Now all $j \in \{1,\ldots,N+M\}$ can occur in the Hamiltonian $\hat{H}^{(tot)}\left(\hat{x}_i^{(j)}, p_i^{(j)}\right)$. Choose a

maximal set of $3(N+M)$ commuting self-adjoint operators $\hat{Q}^{(1)},\ldots,\hat{Q}^{(3N)},\hat{R}^{(1)},\ldots,\hat{R}^{(3M)}$. Again, obtain a joint orthonormal eigenbasis

$$\{|q_{n_1}^{(1)},\ldots,q_{n_{3N}}^{(3N)},r_{m_1}^{(1)},\ldots,r_{m_{3M}}^{(3M)}\rangle\}.$$

A general pure state of S_{tot} is

$$|\psi^{(tot)}(t)\rangle=\sum_{n_1,\ldots,m_{3M}}\psi^{(tot)}_{n_1,\ldots,m_{3M}}(t)|q_{n_1}^{(1)},\ldots,r_{m_{3M}}^{(3M)}\rangle.$$

Now, given that at time t_0, S_c and S_b are respectively in pure states $|\psi^{(S_c)}\rangle$ and $|\varphi^{(S_b)}\rangle$, then the state of S_{tot} has the components

$$\psi^{(tot)}_{n_1,\ldots,m_{3M}}(t_0)=\psi^{(S_c)}_{n_1,\ldots,n_{3N}}\varphi^{(S_b)}_{m_1,\ldots,m_{3M}}.$$

Consider the expectation of an observable $\hat{f}$ of S_c at time t_0:

$$\begin{aligned}\overline{f}&=\langle\psi^{(tot)}(t_0)|\hat{f}|\psi^{(tot)}(t_0)\rangle\\&=\sum_{\substack{n_1,\ldots,m_{3M},\\ \tilde{n}_1,\ldots,\tilde{m}_{3M}}}\psi^{*(tot)}_{\tilde{n}_1,\ldots,\tilde{m}_{3M}}(t_0)\hat{f}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},n_1,\ldots,n_{3N}}\psi^{(tot)}_{n_1,\ldots,m_{3M}}(t_0)\\&=\left(\sum_{\substack{n_1,\ldots,n_{3N},\\ \tilde{n}_1,\ldots,\tilde{n}_{3N}}}\psi^{*(S_c)}_{\tilde{n}_1,\ldots,\tilde{n}_{3N}}\hat{f}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},n_1,\ldots,n_{3N}}\psi^{(S_c)}_{n_1,\ldots,n_{3N}}\right)\\&\qquad\cdot\underbrace{\left(\sum_{\substack{m_1,\ldots,m_{3M},\\ \tilde{m}_1,\ldots,\tilde{m}_{3M}}}\varphi^{*(S_b)}_{\tilde{m}_1,\ldots,\tilde{m}_{3M}}\varphi^{(S_b)}_{m_1,\ldots,m_{3M}}\right)}_{=1}\\&=\langle\psi^{(S_c)}|\hat{f}|\psi^{(S_c)}\rangle.\end{aligned}$$

The evolution according to the joint Hamiltonian $\hat{H}^{(tot)}$ will destroy the product form of the state $|\psi^{(tot)}(t)\rangle$ and we obtain, at time t_1, the general components $\psi^{(tot)}_{n_1,\ldots,m_{3M}}(t_1)$.

Definition 22 *If $\psi^{(tot)}$ is a product, it is called an* unentangled, *or* separable *state. Otherwise, it is called an* entangled *state.*

Consider an observable $\hat{f}$ of S_c. We have

$$\hat{f}_{\tilde{n}_1,\ldots,\tilde{m}_{3M},n_1,\ldots m_{3M}}=\underbrace{\hat{f}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},n_1,\ldots n_{3N}}}_{\text{action on }\mathcal{H}_c}\ \underbrace{\delta_{\tilde{m}_1,m_1}\delta_{\tilde{m}_2,m_2}\cdots\delta_{\tilde{m}_{3M},m_{3M}}}_{\text{action on }\mathcal{H}_b}.$$

At time t_1, the expectation value of the observable $\overline{f}$ of S_c is:

$$\overline{f} = \underbrace{\langle\psi^{(tot)}(t_1)|\hat{f}|\psi^{(tot)}(t_1)\rangle}_{\text{pure-state calculation in } S^{(tot)}}$$

$$= \sum_{\substack{n_1,\ldots,n_{3N},m_1,\ldots,m_{3M},\\ \tilde{n}_1,\ldots,\tilde{n}_{3N},\tilde{m}_1,\ldots\tilde{m}_{3M}}} \psi^{*(tot)}_{\tilde{n}_1,\ldots,\tilde{m}_{3M}}(t_1)\hat{f}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},n_1,\ldots,n_{3N}}(\delta_{\tilde{m}_1,m_1}\cdots\delta_{\tilde{m}_{3M},m_{3M}})\psi^{(tot)}_{n_1,\ldots,m_{3M}}(t_1)$$

$$= \underbrace{\sum_{\substack{n_1,\ldots,n_{3N},\\ \tilde{n}_1,\ldots,\tilde{n}_{3N}}} \hat{f}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},n_1,\ldots,n_{3N}}\hat{\rho}^{(c)}_{n_1,\ldots,n_{3N},\tilde{n}_1,\ldots,\tilde{n}_{3N}}(t_1)}_{\text{mixed-state calculation in } S^{(c)}}$$

$$\text{where} \qquad \hat{\rho}^{(c)}_{n_1,\ldots,n_{3N},\tilde{n}_1,\ldots,\tilde{n}_{3N}}(t_1) = \sum_{m_1,\ldots,m_{3M}} \psi^{*(tot)}_{\tilde{n}_1,\ldots,\tilde{n}_{3N},m_1,\ldots m_{3M}}(t_1)\psi^{(tot)}_{n_1,\ldots,n_{3N},m_1,\ldots m_{3M}}(t_1).$$

7 Compositions of Several Systems

Consider two quantum systems S_1 and S_2. Observables of S_1 contain only $\hat{x}$ and $\hat{p}$ from S_1 acting on $|\psi\rangle \in \mathcal{H}_1$. Observables of S_2 contain only $\hat{x}$ and $\hat{p}$ from S_2 acting on $|\psi\rangle \in \mathcal{H}_2$. Observables of the total system $S_{tot} = S_1 \cup S_2$ may contain all $\hat{x}, \hat{p}$ of S_1 and S_2, acting on $\mathcal{H}_{tot}$. For example

$$\hat{H}^{(tot)} = \frac{\left(\hat{p}^{(1)}\right)^2}{2m_1} + \frac{\left(\hat{p}^{(2)}\right)^2}{m_2} + V\left(\hat{x}^{(1)} - \hat{x}^{(2)}\right).$$

We call V the *interaction term*.

A natural question is "how do we construct $\mathcal{H}_{tot}$ from $\mathcal{H}_1$ and $\mathcal{H}_2$"? A partial answer is that we often have $\mathcal{H}_{tot} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where "$\otimes$" denotes the *tensor product*, defined as follows.

Definition 23 *As a set, define* $\mathcal{H}_1 \otimes \mathcal{H}_2 = span\{|\psi\rangle \otimes |\varphi\rangle : |\psi\rangle \in \mathcal{H}_1, |\varphi\rangle \in \mathcal{H}_2\}$ *with the additional laws:*

I)

$$(|\psi_1\rangle + |\psi_2\rangle) \otimes |\varphi\rangle = |\psi_1\rangle \otimes |\varphi\rangle + |\psi_2\rangle \otimes |\varphi\rangle$$
$$|\psi\rangle \otimes (|\varphi_1\rangle + |\varphi_2\rangle) = |\psi\rangle \otimes |\varphi_1\rangle + |\psi\rangle \otimes |\varphi_2\rangle$$

II)

$$(\lambda|\psi\rangle) \otimes |\varphi\rangle = \lambda(|\psi\rangle \otimes |\varphi\rangle) = |\psi\rangle \otimes (\lambda|\varphi\rangle)$$

and with conjugation map $\dagger$*:*

$$\dagger : |\psi\rangle \otimes |\varphi\rangle \longrightarrow \langle\psi| \otimes \langle\varphi| \in (\mathcal{H}_1 \otimes \mathcal{H}_2)^\dagger$$

(extended linearly for linear combinations).

It can be shown that $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a Hilbert space. In particular, if $\{|b_n\rangle\}_n$ and $\{|c_m\rangle\}_m$ are orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $\{|b_n\rangle \otimes |b_m\rangle\}_{n,m}$ is an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Assume $\hat{f}$ is polynomial in the $\hat{x}$ and $\hat{p}$ of S_1. Define $\hat{f}$ to act on $|\psi\rangle \otimes |\varphi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ as

$$\hat{f} : |\psi\rangle \otimes |\varphi\rangle \longrightarrow (\hat{f}|\psi\rangle) \otimes |\varphi\rangle.$$

We write $\hat{f}$ on the acting on the whole system as $\hat{f} \otimes 1$.

Note: All the CCRs hold on $\mathcal{H}_1 \otimes \mathcal{H}_2$, and $\otimes$ respects Hermiteicity.

By the Stone/Von-Neumann uniqueness theorem, $\mathcal{H}_{tot} \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$. Thus every $|\psi\rangle \in \mathcal{H}_{(tot)}$ can be written as

$$|\psi\rangle = \sum_{n,m} \psi_{n,m} |b_n\rangle \otimes |c_m\rangle.$$

Definition 24 *$|\psi\rangle \in \mathcal{H}_{tot}$ is called* unentangled *(or* separable*) if it can be written as*

$$|\psi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$$

for $|\varphi_1\rangle \in \mathcal{H}_1$, $|\varphi_2\rangle \in \mathcal{H}_2$. Otherwise, $|\psi\rangle$ is called entangled*.*

Assume $|\psi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$ (i.e. $|\psi\rangle$ is unentangled). Now let $\hat{f}$ be any observable on the first subsystem. Consider the expectation of this observable on the whole system; i.e. the expectation of $\hat{f} \otimes 1$.

$$\begin{aligned}
\overline{f} &= \langle\psi|(\hat{f} \otimes 1)|\psi\rangle \\
&= (\langle\varphi_1| \otimes \langle\varphi_2|)(\hat{f} \otimes 1)(|\varphi_1\rangle \otimes |\varphi_2\rangle) \\
&= \langle\varphi_1|\hat{f}|\varphi_1\rangle\langle\varphi_2|\varphi_2\rangle \\
&= \langle\varphi_1|\hat{f}|\varphi_1\rangle.
\end{aligned}$$

But $\langle\varphi_1|\hat{f}|\varphi_1\rangle$ is exactly the expectation we would predict for $\hat{f}$ acting on system S_1 alone. Since this is true for any observable $\hat{f}$, system S_1 is in the pure state $|\varphi_1\rangle$. We have shown the following:

Fact 1 *$|\psi\rangle \in \mathcal{H}_{tot}$ is unentangled iff the states of subsystems S_1 and S_2 considered on their own are pure states.*

If, however, $|\psi\rangle$ is entangled, then the state of system S_1 is a mixed state whose density operator is obtained by the partial trace (i.e. we trace over the basis of the second system; sometimes referred to as *tracing-out* the second system). For the remainder of this section, we show that the partial trace operation gives the right result.

Let $\{|b_n\rangle\}$ be an orthonormal basis of subsystem S_1, and let $\{|c_m\rangle\}$ be an orthonormal basis of subsystem S_2. Suppose $|\psi\rangle$ is a general (possibly entangled) pure state on the total system:

$$|\psi\rangle = \sum_{n,m} \psi_{n,m} |b_n\rangle \otimes |c_m\rangle.$$

The subsystem S_1 is in a mixed state with density operator $\hat{\rho}^{(1)}$ iff

$$\overline{f} = Tr(\hat{f}\hat{\rho}^{(1)}) \qquad (*)$$

for arbitrary Hermitean observables $\hat{f}$ of S_1.

We compute the left-hand side (LHS) and right-hand side (RHS) of $(*)$ separately, and make sure they are equal.

RHS:

By definition of the reduced density operator and partial trace, we have

$$\begin{aligned}
\hat{\rho}^{(1)} &= Tr_{(2)}(|\psi\rangle\langle\psi|) \\
&= Tr_{(2)}\left(\sum_{n,m,\tilde{n},\tilde{m}} \psi_{n,m}\psi^*_{\tilde{n},\tilde{m}}(|b_n\rangle \otimes |c_m\rangle)(\langle b_{\tilde{n}}| \otimes \langle c_{\tilde{m}}|)\right) \\
&= \sum_r \sum_{\substack{n,m,\\ \tilde{n},\tilde{m}}} \psi_{n,m}\psi^*_{\tilde{n},\tilde{m}}|b_n\rangle\langle b_{\tilde{n}}|\langle c_r|c_m\rangle\langle c_{\tilde{m}}|c_r\rangle \\
&= \sum_r \sum_{\substack{n,m,\\ \tilde{n},\tilde{m}}} \psi_{n,m}\psi^*_{\tilde{n},\tilde{m}}|b_n\rangle\langle b_{\tilde{n}}|\delta_{r,m}\delta_{\tilde{m},r} \\
&= \sum_{n,\tilde{n},m} \psi_{n,m}\psi^*_{\tilde{n},m}|b_n\rangle\langle b_{\tilde{n}}|.
\end{aligned}$$

Now, $(\hat{f}\hat{\rho}^{(1)})$ is an operator on system S_1 alone, so by definition of Trace we have

$$\begin{aligned}
Tr(\hat{f}\hat{\rho}^{(1)}) &= \sum_u \langle b_u| \left(\hat{f} \sum_{n,\tilde{n},m} \psi_{n,m}\psi^*_{\tilde{n},m}|b_n\rangle\langle b_{\tilde{n}}|\right)|b_u\rangle \\
&= \sum_{u,m,n,\tilde{n}} \psi_{n,m}\psi^*_{\tilde{n},m}\langle b_u|\hat{f}|b_n\rangle\langle b_{\tilde{n}}|b_u\rangle \\
&= \sum_{u,m,n,\tilde{n}} \psi_{n,m}\psi^*_{\tilde{n},m}\langle b_u|\hat{f}|b_n\rangle\delta_{\tilde{n},u} \\
&= \sum_{m,n,\tilde{n}} \psi_{n,m}\psi^*_{\tilde{n},m}\langle b_{\tilde{n}}|\hat{f}|b_n\rangle. \qquad (1)
\end{aligned}$$

LHS:
To compute the LHS, we consider $\hat{f}$ acting on the total system as $\hat{f} \otimes 1$.

$$\begin{aligned}
\overline{f} &= \langle\psi|(\hat{f} \otimes 1)|\psi\rangle \\
&= \sum_{n,m,\tilde{n},\tilde{m}} \psi^*_{\tilde{n},\tilde{m}}\psi_{n,m}(\langle b_{\tilde{n}}| \otimes \langle c_{\tilde{m}}|)(\hat{f} \otimes 1)(|b_n\rangle \otimes |c_m\rangle) \\
&= \sum_{n,m,\tilde{n},\tilde{m},r,s,u,v} \psi^*_{\tilde{n},\tilde{m}}\psi_{n,m}(\langle b_{\tilde{n}}| \otimes \langle c_{\tilde{m}}|)(|b_r\rangle \otimes |c_s\rangle)(\langle b_r| \otimes \langle c_s|)(\hat{f} \otimes 1)(|b_u\rangle \otimes |c_v\rangle)(\langle b_u| \otimes \langle c_v|)(|b_n\rangle \otimes |c_m\rangle) \\
&= \sum_{n,m,\tilde{n},\tilde{m},r,s,u,v} \psi^*_{\tilde{n},\tilde{m}}\psi_{n,m}(\langle b_{\tilde{n}}|b_r\rangle\langle b_r|\hat{f}|b_u\rangle\langle b_u|b_n\rangle) \otimes (\langle c_{\tilde{m}}|c_s\rangle\langle c_s|c_v\rangle\langle c_v|c_m\rangle) \\
&= \sum_{n,m,\tilde{n},\tilde{m},r,s,u,v} \psi^*_{\tilde{n},\tilde{m}}\psi_{n,m}(\delta_{\tilde{n},r}\langle b_r|\hat{f}|b_u\rangle\delta_{u,n}) \cdot (\delta_{\tilde{m},s}\delta_{s,v}\delta_{v,m}) \\
&= \sum_{m,n,\tilde{n}} \psi^*_{\tilde{n},m}\psi_{n,m}\langle b_{\tilde{n}}|\hat{f}|b_n\rangle. \qquad (2)
\end{aligned}$$

Since (1) = (2), we have LHS=RHS, and so we are done. We have thus shown the following.

Fact 2 *For a general pure state $|\psi\rangle$ of the joint system $S_1 \cup S_2$, the subsystem S_1 considered alone is in a (generally mixed) state whose density operator is*

$$\hat{\rho}^{(1)} = Tr_{(2)}(|\psi\rangle\langle\psi|).$$

8 Identical Particles

8.1 Bosons and Fermions

Consider a system S. Choose a maximum set of commuting observables $\hat{Q}^{(1)}, \ldots, \hat{Q}^{(N)}$ with joint orthonormal eigenbasis $\{|q_{n_1}^{(1)}, \ldots, q_{n_N}^{(N)}\rangle\}$. We will write $\hat{Q}$ and $\{|q_n\rangle\}$ for short. Every $|\psi\rangle \in \mathcal{H}$ can be written $|\psi\rangle = \sum_n \psi_n |q_n\rangle$.

Consider having two identical such systems. The total system is in a state $|\psi\rangle \in \mathcal{H}^{tot}$. We know from the previous section that $\mathcal{H}^{tot} \subseteq \mathcal{H} \otimes \mathcal{H}$. Suppose we measure $\hat{Q}$ for both systems and obtain measurement outcomes q_n, q_m. If we assume that the systems are truly indistinguishable, there is no way of knowing which system was measured in q_n and which was measured in q_m. All that we can know is that one system was measured in each of the two. In fact, it is not even meaningful to ask "which system was measured in q_n?", since the measurement was of a *joint observable* of the overall system. Now consider the state of the system immediately after the collapse. Any state of the form

$$|\psi\rangle = \alpha |q_n\rangle \otimes |q_m\rangle + \beta |q_m\rangle \otimes |q_n\rangle$$

would yield q_n, q_m on an immediate remeasurement. So the question is, which one of these states does nature select? Does nature confine itself to a fixed choice of α and β? Before answering the question, note that an overall phase $e^{i\varphi}$ never matters, since

$$\overline{f} = \langle \psi | e^{-i\psi} \hat{f} e^{i\psi} | \psi \rangle = \langle \psi | \hat{f} | \psi \rangle$$

and so $|\psi\rangle$ and $e^{i\psi}|\psi\rangle$ describe the same physical state. Now, to answer the question above, it turns out that nature always makes one of two choices for the relative phase:

(I) $\alpha = \beta$ ("Bosonic case")

(II) $\alpha = -\beta$ ("Fermionic case").

So, two identical "Bosons" collapse into

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|q_n\rangle \otimes |q_m\rangle + |q_m\rangle \otimes |q_n\rangle)$$

and two identical "Fermions" collapse into

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|q_n\rangle \otimes |q_m\rangle - |q_m\rangle \otimes |q_n\rangle).$$

Bosons are particles which have integer values of angular momentum, and *Fermions* are particles which have half-integer values of angular momentum.

Consider carefully the situation for Fermions. Suppose the outcome of the measurement was q_n, q_n. Then the two identical Fermions will collapse into the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|q_n\rangle \otimes |q_n\rangle - |q_n\rangle \otimes |q_n\rangle) = 0.$$

But 0 is not a physical state, and so we see that identical Fermions can never be in the same state.

Note: If space had fewer than 3 dimensions, then any relative phase would be possible. Such a space would give rise to *Anyons* (e.g. quantum Hall effect).

Remark: $\hat{H}$ treats identical subsystems identically. One can show that symmetry (or anti-symmetry) is preserved under time-evolution. This means that $\mathcal{H}_{tot} \subsetneq \mathcal{H} \otimes \mathcal{H}$ for bosons and fermions.

8.2 Special Behaviour of Bosonic and Fermionic Systems

Consider a system S. Suppose a maximum set of commuting observables is $\hat{H}, \hat{Q}$. Assume the eigenvalues of $\hat{H}$ form a discrete set $\{E_m\}$, and that $\hat{Q}$ has only two eigenvalues $\{q_1, q_2\}$. The joint orthonormal eigenbasis is $\{|E_n, q_i\rangle\}_{n\in\mathbf{N}, i\in\{1,2\}}$. Now, suppose we put the system into a heat bath. The density matrix of the system is

$$\hat{\rho} = \sum_{n\in\mathbf{N}, i\in\{1,2\}} \rho_n |E_n, q_i\rangle\langle E_n, q_i| \quad , \quad \rho_n = \frac{1}{N} e^{-\frac{E_n}{kT}}.$$

Suppose at time t_1 we measure $\hat{H}$ and $\hat{Q}$. The probability for finding state $|E_n, q_1\rangle$ is proportional to $e^{-\frac{E_n}{kT}}$, and the probability for finding state $|E_n, q_2\rangle$ is also proportional to $e^{-\frac{E_n}{kT}}$. In both cases, the probability is independent of i. If we find E_n, then the probability of finding q_1 is $\frac{1}{2}$.

Now, suppose we have two similar systems in the heat bath together, and at time t_1 measure $\hat{H}$ and $\hat{Q}$ of each system. Suppose we obtain energy E_n for both systems (i.e. the measurement result is E_n, E_n). Consider the probability for finding the q_i to be different, verses the probability for finding the q_i the same. There are three cases.

Case 1: *Distinguishable systems*
We are sure to find one of 4 states:

$$|E_n, q_1\rangle \otimes |E_n, q_1\rangle \quad , \quad |E_n, q_1\rangle \otimes |E_n, q_2\rangle$$
$$|E_n, q_2\rangle \otimes |E_n, q_1\rangle \quad , \quad |E_n, q_2\rangle \otimes |E_n, q_2\rangle$$

All four states have equal energy $2E_n$, and so are equally probable. So,

the probability to find the q_i equal is $\frac{1}{2}$, and
the probability to find the q_i unequal is $\frac{1}{2}$.

Case 2: *Identical Bosons*
We are sure to find one of 3 states:

$$|E_n, q_1\rangle \otimes |E_n, q_1\rangle \quad , \quad |E_n, q_2\rangle \otimes |E_n, q_2\rangle$$
$$\frac{1}{\sqrt{2}} \left(|E_n, q_1\rangle \otimes |E_n, q_2\rangle + |E_n, q_2\rangle \otimes |E_n, q_1\rangle\right)$$

All three states have equal energy, and so are equally probable. So,

the probability to find the q_i equal is $\frac{2}{3}$, and
the probability to find the q_i unequal is $\frac{1}{3}$.

The probability for Bosons to be in the same state is enhanced. This is what makes Bose-Einstien condensates (many Bosons in the same state) possible.

Case 3: *Identical Fermions*
We are sure to find the state:

$$\frac{1}{\sqrt{2}} \left(|E_n, q_1\rangle \otimes |E_n, q_2\rangle - |E_n, q_2\rangle \otimes |E_n, q_1\rangle\right)$$

So,

the probability to find the q_i equal is 0, and
the probability to find the q_i unequal is 1.

This is known as the *Pauli exclusion principle*.

9 Angular Momentum

9.1 Generators of Symmetries

We begin by examining how which the (linear) momentum operator $\hat{p}$ can be thought to "generate" translations of the position operator. It can be shown that

$$e^{\frac{ib\hat{p}}{\hbar}}\hat{x}e^{-\frac{ib\hat{p}}{\hbar}} = \hat{x} + b \qquad \forall b \in \mathbb{R}$$

by expanding the second exponential in terms of a power series, and then commuting $\hat{x}$ through $\hat{p}^n$.

Now, consider the action of $\hat{x}$ on the state $e^{\frac{ib\hat{p}}{\hbar}}|x\rangle$.

$$\begin{aligned}
\hat{x}\left(e^{\frac{ib\hat{p}}{\hbar}}|x\rangle\right) &= e^{\frac{ib\hat{p}}{\hbar}}e^{-\frac{ib\hat{p}}{\hbar}}\hat{x}\left(e^{\frac{ib\hat{p}}{\hbar}}|x\rangle\right) \\
&= e^{\frac{ib\hat{p}}{\hbar}}\left(e^{-\frac{ib\hat{p}}{\hbar}}\hat{x}e^{\frac{ib\hat{p}}{\hbar}}\right)|x\rangle \\
&= e^{\frac{ib\hat{p}}{\hbar}}(\hat{x}+b)|x\rangle \\
&= (\hat{x}+b)\left(e^{\frac{ib\hat{p}}{\hbar}}|x\rangle\right)
\end{aligned}$$

and so we see that $e^{\frac{ib\hat{p}}{\hbar}}|x\rangle$ is an eigenvector of $\hat{x}$, with eigenvalue $(\hat{x}+b)$. This eigenvector is $|x+b\rangle$. So we have found

$$e^{\frac{ib\hat{p}}{\hbar}}|x\rangle = |x+b\rangle.$$

In the position basis,

$$\begin{aligned}
& e^{\frac{ib(-i\hbar)}{\hbar}\frac{d}{dx}}\psi(x) = \psi(x+b) \\
\implies\ & e^{b\frac{d}{dx}}\psi(x) = \psi(x+b) \\
\implies\ & \psi(x+b) = 1 + \left(\frac{d}{dx}\psi(x)\right)b + \frac{1}{2}\left(\frac{d^2}{dx^2}\psi(x)\right)b^2 + \ldots \qquad \text{(Taylor series).}
\end{aligned}$$

Consider $\hat{x}^2$. We have

$$\begin{aligned}
e^{\frac{ib\hat{p}}{\hbar}}\hat{x}\hat{x}e^{-\frac{ib\hat{p}}{\hbar}} &= e^{\frac{ib\hat{p}}{\hbar}}\hat{x}e^{-\frac{ib\hat{p}}{\hbar}}e^{\frac{ib\hat{p}}{\hbar}}\hat{x}e^{-\frac{ib\hat{p}}{\hbar}} \\
&= (\hat{x}+b)^2,
\end{aligned}$$

and by induction, $e^{\frac{ib\hat{p}}{\hbar}}\hat{x}^n e^{-\frac{ib\hat{p}}{\hbar}} = (\hat{x}+b)^n$.

Since $[\hat{p},\hat{p}] = 0$, we clearly have $e^{\frac{ib\hat{p}}{\hbar}}\hat{p}e^{-\frac{ib\hat{p}}{\hbar}} = \hat{p}$. Putting these observations together, for any Hamiltonian which is polynomial in $\hat{x},\hat{p}$ we have $e^{\frac{ib\hat{p}}{\hbar}}\hat{H}(\hat{x},\hat{p})e^{-\frac{ib\hat{p}}{\hbar}} = \hat{H}(\hat{x}+b,\hat{p})$.

For a Hamiltonian $\hat{H}$ to be *translation invariant* means

$$\hat{H}(\hat{x},\hat{p}) = \hat{H}(\hat{x}+b,\hat{p}) \qquad \forall b \in \mathbb{R}.$$

We have

$$\hat{H}(\hat{x}+b,\hat{p}) = \hat{H}(\hat{x},\hat{p}) = e^{\frac{ib\hat{p}}{\hbar}}e^{-\frac{ib\hat{p}}{\hbar}}\hat{H}(\hat{x},\hat{p}).$$

We also know from above that

$$\hat{H}(\hat{x}+b,\hat{p}) = e^{\frac{ib\hat{p}}{\hbar}}\hat{H}(\hat{x},\hat{p})e^{-\frac{ib\hat{p}}{\hbar}}.$$

Putting these together we get

$$e^{\frac{ib\hat{p}}{\hbar}}e^{-\frac{ib\hat{p}}{\hbar}}\hat{H}(\hat{x},\hat{p}) = e^{\frac{ib\hat{p}}{\hbar}}\hat{H}(\hat{x},\hat{p})e^{-\frac{ib\hat{p}}{\hbar}},$$

which implies that $[\hat{p}, \hat{H}] = 0$. By Heisenberg's equation, this implies that $\dot{\hat{p}} = 0$. So for a Hamiltonian $\hat{H}$ which is translation invariant, the generator $\hat{p}$ of the translation "symmetry" is conserved. This is an instance of a fact sometimes referred to as the "Noether Theorem"; namely, that every symmetry of the Hamiltonian implies a conserved quantity.

Now, keeping in mind this picture of operators behaving as generators of symmetries, we move onto angular momentum.

9.2 Orbital Angular Momentum

Definition 25 Orbital angular momentum *in three dimensions is defined as follows.*

$$\hat{L}_1 = \hat{x}_2\hat{p}_3 - \hat{x}_3\hat{p}_2 \quad , \quad \hat{L}_2 = \hat{x}_3\hat{p}_1 - \hat{x}_1\hat{p}_3 \quad , \quad \hat{L}_3 = \hat{x}_1\hat{p}_2 - \hat{x}_2\hat{p}_1.$$

The following notation is often used.

$$\hat{L}_i = \sum_{i,j,k=1}^{3} \varepsilon_{i,j,k}\hat{x}_j\hat{p}_k$$

where $\varepsilon_{1,2,3} = 1$*, and* ε *is totally antisymmetric (so, for example,* $\varepsilon_{2,1,3} = -1$*).*

In the same sense that $\hat{p}$ generates translations of $\hat{x}$, the orbital angular momentum $\hat{\mathbf{L}}$ generates rotations of $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$. Consider, for example, a rotation $\hat{\mathbf{L}} = (0, 0, \hat{L}_3)$ by angle θ about the z-axis:

$$\begin{aligned}
\hat{x}_1 &\longrightarrow \hat{x}_1\cos\theta + \hat{x}_2\sin\theta \\
\hat{x}_2 &\longrightarrow -\hat{x}_1\sin\theta + \hat{x}_2\cos\theta \\
\hat{x}_3 &\longrightarrow \hat{x}_3 \\
\mathbf{x} &\longrightarrow R\mathbf{x}.
\end{aligned}$$

One can show that

$$\begin{aligned}
e^{\frac{i\theta\hat{L}_3}{\hbar}}\hat{x}_1 e^{-\frac{i\theta\hat{L}_3}{\hbar}} &= \hat{x}_1\cos\theta + \hat{x}_2\sin\theta \\
e^{\frac{i\theta\hat{L}_3}{\hbar}}\hat{x}_2 e^{-\frac{i\theta\hat{L}_3}{\hbar}} &= -\hat{x}_1\sin\theta + \hat{x}_2\cos\theta \\
e^{\frac{i\theta\hat{L}_3}{\hbar}}\hat{x}_3 e^{-\frac{i\theta\hat{L}_3}{\hbar}} &= \hat{x}_3
\end{aligned}$$

and therefore

$$e^{\frac{i\theta\hat{L}_3}{\hbar}}|\mathbf{x}\rangle = |R\mathbf{x}\rangle$$

where R is the rotation matrix.

Similarly,

$$e^{\frac{i\theta\hat{L}_3}{\hbar}}|\mathbf{p}\rangle = |R\mathbf{p}\rangle.$$

Note that this gives the "Taylor expansion analogue":

$$e^{\alpha\left(\hat{x}_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right)}\psi(\mathbf{x}) = \psi(R\mathbf{x}).$$

For $\hat{H}$ to be *rotation invariant* means

$$H(\hat{\mathbf{x}}, \hat{\mathbf{p}}) = \hat{H}(R\hat{\mathbf{x}}, R\hat{\mathbf{p}}).$$

This means

$$e^{\frac{i\theta\hat{L}_3}{\hbar}}\hat{H}(\hat{x},\hat{p})e^{-\frac{i\theta\hat{L}_3}{\hbar}} = \hat{H}(\hat{x},\hat{p})$$

which gives $[\hat{L}_3, \hat{H}] = 0$, and so $\dot{\hat{L}}_3 = 0$. So for a Hamiltonian which is rotation invariant, the generator $\hat{L}_3$ of the rotation symmetry is conserved.

For example, consider the Hamiltonian for the Hydrogen atom

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} + \frac{\alpha}{\sqrt{\hat{\mathbf{x}}^2}}.$$

This Hamiltonian is rotation invariant, because $\mathbf{x}^2$, $\hat{\mathbf{p}}^2$ are scalars for rotations.

9.3 Angular Momentum and Identical Particles

Recall that when we studied identical particles we said that bosons are particles having integer values for angular momentum, and fermions are particles having half-integer values for angular momentum. Suppose we have a boson, and assume that $\hat{L}_3|\psi\rangle = \hbar m|\psi\rangle$, where $m \in \mathbb{Z}$. Then

$$2\pi \text{ rotated } |\psi\rangle \text{ is } e^{\frac{2\pi i\hat{L}_3}{\hbar}}|\psi\rangle = e^{\frac{2\pi i\hbar m}{\hbar}}|\psi\rangle = |\psi\rangle.$$

On the other hand, suppose we have a fermion, and assume that $\hat{L}_3|\psi\rangle = \hbar\frac{m}{2}|\psi\rangle$, where $m \in \mathbb{Z}$. Then

$$2\pi \text{ rotated } |\psi\rangle \text{ is } e^{\frac{2\pi i\hat{L}_3}{\hbar}}|\psi\rangle = e^{\frac{2\pi i\hbar m}{2\hbar}}|\psi\rangle = e^{\pi i m}|\psi\rangle = -|\psi\rangle.$$

This is the underlying reason why "swapping" two bosons leaves the state unchanged, but "swapping" two fermions introduces a (-1) factor.

9.4 Orbital Angular Momentum v.s. Spin

We proceed by finding joint eigenvectors of $\hat{\mathbf{L}}^2$ and L_3:

$$\hat{L}_3|\lambda,\mu,\tau\rangle = \mu\hbar|\lambda,\mu,\tau\rangle \qquad (I)$$
$$\hat{\mathbf{L}}^2|\lambda,\mu,\tau\rangle = \lambda\hbar^2|\lambda,\mu,\tau\rangle \qquad (II)$$

where τ are the eigenvalues with respect to the rest of a maximal set of commuting observables. For simplicity, we will omit writing τ.

One approach to solving (I) and (II) is to convert to a position representation: $\psi_{\lambda,\mu}(\mathbf{x}) = \langle\mathbf{x}|\lambda,\mu\rangle$. From Definition 25, and the rule for converting to the position basis, (I) and (II) become

$$\left(-i\hbar x_1\frac{\partial}{\partial x_2} + i\hbar x_2\frac{\partial}{\partial x_1}\right)\psi_{\lambda,\mu}(\mathbf{x}) = \mu\hbar\psi_{\lambda,\mu}(\mathbf{x}) \qquad (I')$$
$$\sum_{i=1}^{3}(\cdots)^2\,\psi_{\lambda,\mu}(\mathbf{x}) = \lambda\hbar^2\psi_{\lambda,\mu}(\mathbf{x}) \qquad (II')$$

These two partial differential equations can be solved, giving the result of "spherical harmonics":

$$\psi_{\lambda,\mu}(\mathbf{x}) = Y_\mu^\lambda(\mathbf{x})$$

with eigenvalues

$$\lambda = j(j+1) \quad \text{where } j \in \{0,1,2,\ldots\}$$
$$\mu = \{-j, -j+1, \ldots, 0, 1, 2, \ldots, j\}$$

Notice that we only get integer solutions. This means we have only found solutions corresponding to bosons. But we know that fermions exist, so our solution must have missed something. The calculation didn't miss any solutions... what went wrong was that we used the definition for *orbital angular momentum* when we set up $(I'), (II')$. That is, we assumed

$$\hat{L}_i = \sum_{j,k=1}^{3} \varepsilon_{i,j,k} \hat{x}_j \hat{p}_k.$$

It turns out that if we only make the weaker assumption that

$$[\hat{L}_i, \hat{L}_j] = i\hbar \sum_{k=1}^{3} \varepsilon_{i,j,k} \hat{L}_k \qquad (*)$$

the resulting $\hat{L}_i$ can be shown to generate of rotations of the vector $(\hat{L}_1, \hat{L}_2, \hat{L}_3)$. For example, the rotation generated by $\hat{L}_3$ is:

$$\begin{aligned}
e^{\frac{i\theta \hat{L}_3}{\hbar}} \hat{L}_1 e^{-\frac{i\theta \hat{L}_3}{\hbar}} &= \hat{L}_1 \cos\theta + \hat{L}_2 \sin\theta \\
e^{\frac{i\theta \hat{L}_3}{\hbar}} \hat{L}_2 e^{-\frac{i\theta \hat{L}_3}{\hbar}} &= -\hat{L}_1 \sin\theta + \hat{L}_2 \cos\theta \\
e^{\frac{i\theta \hat{L}_3}{\hbar}} \hat{L}_3 e^{-\frac{i\theta \hat{L}_3}{\hbar}} &= \hat{L}_3
\end{aligned}$$

So, to be general, we should study (2) directly.

Note: A direct consequence of $(*)$ is

$$\left[\hat{\mathbf{L}}^2, \hat{L}_i\right] = 0.$$

Remark: (2) is a special case of a *Lie Algebra*.

Definition 26

$$\begin{aligned}
\hat{L}_+ &= \hat{L}_1 + i\hat{L}_2 \\
\hat{L}_- &= \hat{L}_1 - i\hat{L}_2
\end{aligned}$$

$L_\pm$ has the following easily verifiable properties.

$$\begin{aligned}
&[\hat{L}_+, \hat{L}_-] = 2\hbar \hat{L}_3 \\
&[\hat{L}_3, \hat{L}_\pm] = \pm\hbar \hat{L}_\pm \\
&[\hat{\mathbf{L}}, \hat{L}_\pm] = 0 \qquad (I) \\
&\hat{\mathbf{L}}^2 = \hat{L}_- \hat{L}_+ + \hat{L}_3 + \hbar \hat{L}_3 \qquad (II)
\end{aligned}$$

Now assume $|\lambda, \mu\rangle$ is a joint eigenvector of $\mathbf{L}^2$ and $\hat{L}_3$.

Proposition 1 *$\hat{L}_\pm |\lambda, \mu\rangle$ is proportional to $|\lambda, \mu \pm 1\rangle$.*

Check:

$$\begin{aligned}
\hat{\mathbf{L}}^2 (\hat{L}_\pm |\lambda, \mu\rangle) &= \hat{L}_\pm (\hat{\mathbf{L}}^2 |\lambda, \mu\rangle) \qquad \text{by } (I) \\
&= \lambda\hbar^2 (\hat{L}_\pm |\lambda, \mu\rangle)
\end{aligned}$$

$$\hat{L}_3(\hat{L}_\pm|\lambda,\mu\rangle) = (\hat{L}_\pm\hat{L}_3 \pm \hbar\hat{L}_\pm)|\lambda,\mu\rangle = \hbar(\mu\pm 1)(\hat{L}_\pm|\lambda,\mu\rangle) \quad \square$$

Now, we calculate $\|\hat{L}_\pm|\lambda,\mu\rangle\|$ (we set $\hbar = 1$ for simplicity).

$$\begin{aligned}\|\hat{L}_\pm|\lambda,\mu\rangle\|^2 &= \langle\lambda,\mu|\hat{L}_\mp\hat{L}_\pm|\lambda,\mu\rangle \\ &= \langle\lambda,\mu|(\hat{\mathbf{L}}^2 - \hat{L}_3^2 \mp \hat{L}_3|\lambda,\mu\rangle \\ &= \lambda - \mu^2 \mp \mu \quad \text{(because } \langle\lambda,\mu|\lambda,\mu\rangle = 1)\end{aligned}$$

In general, we have

$$\|(\hat{L}_\pm)^n|\lambda,\mu\rangle\|^2 = \prod_{\nu=0}^{n-1}(\lambda - (\mu\pm\nu)^2 \mp (\mu\pm\nu)).$$

Clearly, this may become negative for sufficiently large choices of n. Since any vector's norm is zero or positive, the assumption that $|\lambda,\mu\rangle$ is a joint eigenvector of $\hat{\mathbf{L}}^2$ and $\hat{L}_3$ can be true only if λ and μ are such that the product on the RHS never becomes negative. So λ and μ must be such that for some n, a factor in the product equals zero. In this way, the application of high powers on $\hat{L}_\pm$ on $|\lambda,\mu\rangle$ merely leads to the zero vector in the Hilbert space for sufficiently large n.

This means that the possible eigenvalues λ,μ are such that there exist nonnegative integers ν,ν' so that

$$\lambda - (\mu+\nu)^2 - (\mu+\nu) = 0 \quad (1)$$
$$\lambda - (\mu-\nu')^2 + (\mu-\nu') = 0 \quad (2)$$

Subtracting the two equations yields:

$$-2\mu\nu - 2\mu\nu' - \nu^2 + \nu'^2 - 2\mu - \nu + \nu' = 0$$

i.e.

$$2(\nu+\nu'+1)\mu + (\nu+\nu'+1)(\nu-\nu') = 0.$$

Clearly, for this equation to hold, the eigenvalues of μ must be of the form

$$\mu = \frac{\nu'-\nu}{2}$$

and substituting into (2), the eigenvalues λ must be of the form

$$\lambda = \frac{\nu'+\nu}{2}\left(\frac{\nu'+\nu}{2}+1\right).$$

We allowed ν and ν' to be any two nonnegative integers. Define $j = \frac{\nu'+\nu}{2}$. We can now read-off that joint eigenvectors $|\lambda,\mu\rangle$ of $\hat{\mathbf{L}}^2$ and $\hat{L}_3$ have these (and only these) eigenvalues:

$$\lambda = j(j+1) \qquad \text{where } j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots\}$$

and

$$\mu \in \{-j, \ldots, -1, 0, 1, \ldots, j\}.$$

Note: Since j is almost the root of λ, j is essentially the length of the angular momentum vector. The condition for μ says that the component of the angular momentum vector in the z-direction cannot be longer than the total length of the angular momentum vector.

Note: It would seem we have only shown that all eigenvalues other than the above mentioned ones are excluded. In fact, we have shown their (mathematical) existence by explicitly constructing their eigenvectors from one vector $|\lambda,\mu\rangle$ by the repeated application of $\hat{L}_\pm$.

Note: An often used notation for the state $|\lambda,\mu\rangle$ is $|j,\mu\rangle$.

9.5 Spin

We studied the commutation relation

$$[\hat{L}_i, \hat{L}_j] = i\hbar \sum_{k=1}^{3} \varepsilon_{i,j,k} \hat{L}_k \qquad (*)$$

and found that for any $\hat{\mathbf{L}}$ satisfying this the eigenvalues corresponding to

$$\hat{\mathbf{L}}^2 |\lambda, \mu\rangle = \hbar^2 \lambda |\lambda, \mu\rangle$$
$$\hat{L}_3 |\lambda, \mu\rangle = \hbar \mu |\lambda, \mu\rangle$$

are

$$\lambda = j(j+1) \quad \text{where} \quad j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots\}$$
$$\mu = \{-j, -j+1, \ldots, 0, 1, \ldots, j\}.$$

We saw that an example of $\hat{\mathbf{L}}$ which obeys $(*)$ is the *orbital angular momentum*:

$$\hat{L}_i^{(orb)} = \sum_{j,k=1}^{3} \varepsilon_{i,j,k} \hat{x}_j \hat{p}_k.$$

We saw that the possible eigenvalues for orbital angular momentum are only those corresponding to $j \in \{0, 1, 2, \ldots\}$. But we found that for operators obeying $(*)$ the mathematics admits the possibility of those having half-integer eigenvalues. We cannot interpret these as orbital angular momentum vectors. But the presence of Fermions in the universe indicates that half-integer angular momentum values *do occur in nature.* Recall we saw that for Fermions, a 2π-rotation introduced a relative phase factor of (-1). We know that all our usual spatial degrees of freedom are unaffected by such a 2π-rotation. So it must be that an angular momentum operator having half-integer eigenvalues acts on something *other* than the Hilbert space that we have dealt with so far. This means we have a new degree of freedom for these operators. We call this the *spin* degree of freedom. We introduce a new, finite-dimensional, Hilbert space for the spin, which we attach to our familiar Hilbert space using a tensor product.

Physical Assumption 6 *Quantum mechanical systems have a physically meaningful* spin *degree of freedom which is independent of the position/momentum degrees of freedom. The corresponding spin operators $\hat{L}^{(spin)}$ act on a (new) finite dimensional Hilbert space. A description of a system including it's position/momentum* and *spin degrees of freedom is given by a unit vector in $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_s$, where $\mathcal{H}_0$ is the Hilbert space on which the $\hat{x}$ and $\hat{p}$ act, and $\mathcal{H}_s$ is the new finite-dimensional Hilbert space on which the spin operators act.*

Note: The *total angular momentum* operator for a system is the sum of all its angular momentum operators; i.e. the sum of all the orbital angular momentum operators of its constituent particles and all the spin operators of its constituent particles.

The following is also known from physical experiment.

Physical Assumption 7 *Each species of particle has a fixed eigenvalue λ for its spin.*

Some examples:

e^{-1}, quarks, neutrinos: $\lambda = \frac{1}{2}(1 + \frac{1}{2})$ (i.e. $j = \frac{1}{2}$) *"spinor"*

Higg's particles: $\lambda = 0$ (i.e. $j = 0$) *"scalar"*

Photons, gluons, w, z: $\lambda = 1(1+1)$ (i.e. $j = 1$) *"vector"*

Since the spin operators act nontrivially only on these new degrees of freedom, they commute with all the operators which act on the old Hilbert space (such as the $\hat{x}$, $\hat{p}$, orbital angular momenta, etc). This means we can add, for example, $\hat{\mathbf{L}}^{(spin)^2}$, $\hat{L}_3^{(spin)}$ to a maximal set of commuting observables. So a new maximal set of observables for a single particle could be

$$\hat{Q}^{(1)}, \hat{Q}^{(2)}, \hat{Q}^{(3)}, \hat{\mathbf{L}}^{(spin)^2}, \hat{L}_3^{(spin)}$$

where the $\hat{Q}^{(i)}$ are commuting observables which only contain the $\hat{x}$ and $\hat{p}$ (i.e. which do not contain any of the spin operators $\hat{S}_i$). The operators $\hat{x} \otimes 1, \hat{p} \otimes 1$ and $1 \otimes \hat{S}_i$ act on the full Hilbert space $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_s$. In general, the Hamiltonian may contain any polynomial (or other functions) of these operators.

Note that, in practise, it is often a good approximation to assume that the Hamiltonian is of the form $\hat{H} = \hat{H}_0 + \hat{H}_s$, where $\hat{H}_0$ contains only the $\hat{x}$ and $\hat{p}$ operators, and where $\hat{H}_s$ contains only the spin operators $\hat{S}_i$. (This is completely analogous to the case of two noninteracting subsystems of a larger system, the total Hamiltonian being the sum of the subsystems' Hamiltonians without any additional interaction Hamiltonian.) We say in this case that the spin degree of freedom is *decoupled* from the usual motion degrees of freedom of the particle.

9.6 Working With Spin-$\frac{1}{2}$ Particles

For a spin-$\frac{1}{2}$ particle, we choose as a maximal set of commuting observables

$$\hat{Q}^{(1)}, \hat{Q}^{(2)}, \hat{Q}^{(3)}, \hat{\mathbf{L}}^{(spin)^2}, \hat{L}_3^{(spin)}$$

with joint eigenbasis

$$\{|q_{n_1}, q_{n_2}, q_{n_3}, \lambda, \mu\rangle\}$$

where $\lambda = \frac{1}{2}(1 + \frac{1}{2})$.

Sometimes these basis vectors are denoted

$$\{|q_{n_1}, q_{n_2}, q_{n_3}\rangle \otimes |\lambda, \mu\rangle\}.$$

Since λ is fixed, $|\lambda, \mu\rangle$ is often shortened to $|\mu\rangle$ (i.e. $|\frac{1}{2}\rangle, |-\frac{1}{2}\rangle$ or also as $|\uparrow\rangle, |\downarrow\rangle$).

Let $\hat{\mathbf{S}} = \hat{\mathbf{L}}^{(spin)}$. Consider the matrix elements of $\hat{S}_i$ in the $\{|\frac{1}{2}\rangle, |-\frac{1}{2}\rangle\}$-basis. We have

$$\begin{aligned}\hat{S}_3|\tfrac{1}{2}\rangle &= \tfrac{1}{2}\hbar|\tfrac{1}{2}\rangle \\ \hat{S}_3|-\tfrac{1}{2}\rangle &= -\tfrac{1}{2}\hbar|-\tfrac{1}{2}\rangle.\end{aligned}$$

So

$$\langle\mu|\hat{S}_3|\mu'\rangle = \begin{pmatrix} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{pmatrix}_{\mu,\mu'}.$$

Also, using the operators $\hat{L}_\pm$ we saw in Section 9.4 we can derive

$$\begin{aligned}\langle\mu|\hat{S}_1|\mu'\rangle &= \begin{pmatrix} 0 & \frac{1}{2}\hbar \\ \frac{1}{2}\hbar & 0 \end{pmatrix}_{\mu,\mu'} \\ \langle\mu|\hat{S}_2|\mu'\rangle &= \begin{pmatrix} 0 & -\frac{i}{2}\hbar \\ \frac{i}{2}\hbar & 0 \end{pmatrix}_{\mu,\mu'}.\end{aligned}$$

These are sometimes more compactly represented as

$$\hat{S}_i = \frac{\hbar}{2}\sigma_i$$

where σ_i are the *Pauli Matrices*:

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

Suppose we choose $\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{S}_3$ as a maximal set of commuting observables, with joint orthonormal basis $\{|\hat{x}_1, \hat{x}_2, \hat{x}_3, \mu\rangle\}$.

We have a resolution of the identity:

$$\mathbf{1} = \sum_{\mu \in \{\frac{1}{2}, -\frac{1}{2}\}} \int_{\mathbb{R}^3} |\mathbf{x}, \mu\rangle\langle \mathbf{x}, \mu| d^3\mathbf{x}$$

and then

$$|\psi\rangle = \sum_{\mu \in \{\frac{1}{2}, -\frac{1}{2}\}} \int_{\mathbb{R}^3} \psi_\mu(\mathbf{x})|\mathbf{x}, \mu\rangle d^3\mathbf{x} \qquad \text{where } \psi_\mu(\mathbf{x}) = \langle \mathbf{x}, \mu|\psi\rangle.$$

Scalar products are

$$\langle\varphi|\psi\rangle = \sum_{\mu \in \{\frac{1}{2}, -\frac{1}{2}\}} \int_{\mathbb{R}^3} \varphi^*_\mu(\mathbf{x})\psi_\mu(\mathbf{x}) d^3\mathbf{x}.$$

Matrix elements of operators are

$$\langle \mathbf{x}, \mu|\hat{S}_3|\mathbf{x}', \mu'\rangle = \delta^3(\mathbf{x} - \mathbf{x}')\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mu,\mu'}$$
$$\langle \mathbf{x}, \mu|\hat{x}_i|\mathbf{x}', \mu'\rangle = x_i\delta^3(\mathbf{x} - \mathbf{x}')\delta_{\mu,\mu'} = x_i\delta^3(\mathbf{x} - \mathbf{x}')\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{\mu,\mu'}$$
$$\langle \mathbf{x}, \mu|\hat{p}_j|\mathbf{x}', \mu'\rangle = i\hbar\frac{\partial}{\partial x_j}\delta^3(\mathbf{x} - \mathbf{x}')\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{\mu,\mu'}.$$

Wavefunctions now become 2-vectors

$$\begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix}$$

and the scalar product in terms of these is

$$\langle\varphi|\psi\rangle = \int_{\mathbb{R}^3} \left(\varphi^*_{\frac{1}{2}}(\mathbf{x}), \varphi^*_{-\frac{1}{2}}(\mathbf{x})\right)\begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix} d^3x.$$

The action of operators $\hat{Q}|\psi\rangle = |\varphi\rangle$ in the $\hat{\mathbf{x}}$, $\hat{S}_3$ basis becomes

$$\hat{Q} = \hat{x}_j : \qquad \langle \mathbf{x}, \mu | \hat{x}_j | \psi \rangle = x_j \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix}$$

$$\hat{Q} = \hat{p}_j : \qquad \langle \mathbf{x}, \mu | \hat{p}_j | \psi \rangle = -i\hbar \frac{\partial}{\partial x_j} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix}$$

$$\hat{Q} = \hat{S}_1 : \qquad \langle \mathbf{x}, \mu | \hat{S}_1 | \psi \rangle = \frac{\hbar}{2} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix} \right].$$

9.7 Example: e^- in a B-Field

$$\hat{H} = \frac{\hat{p}^2}{2m} + \alpha \hat{\mathbf{L}}^{(orb)} \cdot \mathbf{B} + \beta \hat{\mathbf{S}} \cdot \mathbf{B}.$$

Suppose, for example, the B-field is constant in the x-direction. Then Schrödinger's equation in the $\hat{\mathbf{x}}$, $\hat{S}_3$-basis is:

$$i\hbar \frac{d}{dt} \begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix} = \Bigg(-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$+ \left(-i\hbar x_2 \frac{\partial}{\partial x_3} + i\hbar x_3 \frac{\partial}{\partial x_2} \right) b \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta b \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Bigg) \begin{pmatrix} \psi_{\frac{1}{2}}(\mathbf{x}) \\ \psi_{-\frac{1}{2}}(\mathbf{x}) \end{pmatrix}.$$

10 Time-Independent Perturbation Theory

As we have just seen in the example above, Hamiltonians can very quickly become extremely complicated, even for relatively simple physical systems. The can give rise to systems of many coupled differential equations, which could be very difficult to solve, in general. *Perturbation theory* is a systematic procedure for obtaining approximate solutions by building on the known exact solutions to a simplified case.

Suppose the Hamiltonian $\hat{H}$ can be split into the sum of an easy-to-diagonalise term $\hat{H}_0$, and the remaining term is $\varepsilon \hat{W}$, where ε is some small number:

$$\hat{H} = \hat{H}_0 + \varepsilon \hat{W}$$
and $\qquad \hat{H}_0 |E_n^{(0)}\rangle = E_n^{(0)} |E_n^{(0)}\rangle \quad , \quad$ where $E_n^{(0)}, |E_n^{(0)}\rangle$ are assumed easy to calculate.

The goal is to find $\{E_n, |E_n\rangle\}$ which solve $\hat{H}|E_n\rangle = E_n|E_n\rangle$. (1)

Assume $spec(\hat{H}_0)$ is nondegenerate and discrete (the more general case is just more complicated). Treat $\hat{H}(\varepsilon)$ as a function of ε. Then the eigenvalues $E_n(\varepsilon)$ and the eigenvectors $|E_n(\varepsilon)\rangle$ are functions of ε as well. Taylor expand E_n and $|E_n\rangle$ in powers of ε, and use the fact that ε is small to ignore terms of ε-degree ≥ 3 in the expansion:

$$E_n = E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \mathcal{O}(\varepsilon^3) \qquad (2)$$
$$|E_n\rangle = |E_n^{(0)}\rangle + \varepsilon |E_n^{(1)}\rangle + \varepsilon^2 |E_n^{(2)}\rangle + \mathcal{O}(\varepsilon^3) \qquad (3)$$

(assume the $|E_n^{(i)}\rangle$ have been normalised). Substituting (2) and (3) into (1) we get:

$$\left(\hat{H}_0 + \varepsilon\hat{W}\right)\left(|E_n^{(0)}\rangle + \varepsilon|E_n^{(1)}\rangle + \varepsilon^2|E_n^{(2)}\rangle + \mathcal{O}(\varepsilon^3)\right)$$
$$= \left(E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \mathcal{O}(\varepsilon^3)\right)\left(|E_n^{(0)}\rangle + \varepsilon|E_n^{(1)}\rangle + \varepsilon^2|E_n^{(2)}\rangle + \mathcal{O}(\varepsilon^3)\right).$$

Now compare terms of like powers of ε in the above equation to get:

ε^0-terms: $\hat{H}_0|E_n^{(0)}\rangle = E_n^{(0)}|E_n^{(0)}\rangle$ (4)

ε^1-terms: $\hat{W}|E_n^{(0)}\rangle + \hat{H}_0|E_n^{(1)}\rangle = E_n^{(0)}|E_n^{(1)}\rangle + E_n^{(1)}|E_n^{(0)}\rangle$ (5)

ε^2-terms: $\hat{W}|E_n^{(1)}\rangle + \hat{H}_0|E_n^{(2)}\rangle = E_n^{(2)}|E_n^{(0)}\rangle + E_n^{(1)}|E_n^{(1)}\rangle + E_n^{(0)}|E_n^{(2)}\rangle$ (6).

Equation (4) was assumed to be easy to solve. Now, multiply equation (5) on the left by $\langle E_n^{(0)}|$:

$$\langle E_n^{(0)}|\hat{W}|E_n^{(0)}\rangle + \langle E_n^{(0)}|\hat{H}_0|E_n^{(1)}\rangle = E_n^{(0)}\langle E_n^{(0)}|E_n^{(1)}\rangle + E_n^{(1)}\langle E_n^{(0)}|E_n^{(0)}\rangle.$$

This gives the first-order correction to the eigenvalues:

$$E_n^{(1)} = \langle E_n^{(0)}|\hat{W}|E_n^{(0)}\rangle.$$

Since, by assumption, the eigenvectors $|E_n^{(0)}\rangle$ are explicitly known in some basis, the calculation of the first-order correction in practise boils down to some explicit integral or sum (depending on whether one uses a discrete or continuous basis).

To compute the first-order correction $|E_n^{(1)}\rangle$ to the eigenstate, multiply equation (5) on the left by $\langle E_{n'}^{(0)}|$:

$$\langle E_{n'}^{(0)}|\hat{W}|E_n^{(0)}\rangle + E_{n'}^{(0)}\langle E_{n'}^{(0)}|E_n^{(1)}\rangle = E_n^{(0)}\langle E_{n'}^{(0)}|E_n^{(1)}\rangle + E_n^{(1)}\langle E_{n'}^{(0)}|E_n^{(0)}\rangle.$$

Therefore:

$$\langle E_{n'}^{(0)}|E_n^{(1)}\rangle = \frac{\langle E_{n'}^{(0)}|\hat{W}|E_n^{(0)}\rangle}{E_n^{(0)} - E_{n'}^{(0)}}. \quad (7)$$

Note that substituting the expansion equation (3) into $\langle E_n|E_n\rangle = 1$ yields:

$$\langle E_n^{(0)}|E_n^{(1)}\rangle = 0 \quad (8).$$

Equation (8) immediately gives:

$$|E_n^{(1)}\rangle = \sum_{n'\neq n} |E_{n'}^{(0)}\rangle\langle E_{n'}^{(0)}|E_n^{(1)}\rangle$$

which together with equation (7) gives the first order correction to the eigenstate:

$$|E_n^{(1)}\rangle = \sum_{n'\neq n} \frac{\langle E_{n'}^{(0)}|\hat{W}|E_n^{(0)}\rangle}{E_n^{(0)} - E_{n'}^{(0)}}|E_{n'}^{(0)}\rangle.$$

Finally, we calculate the second-order correction $E_n^{(2)}$ to the energy eigenvalues. Multiply equation (6) on the left by $\langle E_n^{(0)}|$:

$$\langle E_n^{(0)}|\hat{W}|E_n^{(1)}\rangle + \langle E_n^{(0)}|\hat{H}_0|E_n^{(2)}\rangle = E_n^{(2)}\langle E_n^{(0)}|E_n^{(0)}\rangle + E_n^{(1)}\langle E_n^{(0)}|E_n^{(1)}\rangle + E_n^{(0)}\langle E_n^{(0)}|E_n^{(2)}\rangle.$$

Therefore

$$\begin{aligned} E_n^{(2)} &= \langle E_n^{(0)}|\hat{W}|E_n^{(1)}\rangle \\ &= \sum_{n'\neq n} \frac{\langle E_n^{(0)}|\hat{W}|E_{n'}^{(0)}\rangle\langle E_{n'}^{(0)}|\hat{W}|E_n^{(0)}\rangle}{E_n^{(0)} - E_{n'}^{(0)}} \end{aligned}$$

and the second-order energy eigenvalue correction is thus

$$E_n^{(2)} = \sum_{n'\neq n} \frac{|\langle E_n^{(0)}|\hat{W}|E_{n'}^{(0)}\rangle|^2}{\left(E_n^{(0)} - E_{n'}^{(0)}\right)}.$$

11 Bell's Theorem

A profound question is the following: *why is quantum mechanics probabilistic?*. In the 1930's, Einstein, Podolsky and Rosen (EPR) put forth a hypothesis to answer this question. They believed that the probabilistic nature of quantum mechanics was due to incompleteness in the theory. The hypothesis was that the state of any system is completely specified by giving

$$\{|\psi\rangle, \lambda_1, \ldots, \lambda_N\}$$

where $\lambda_1, \ldots, \lambda_N$ are some "hidden variables" not accounted for in quantum mechanics. The hypothesis says that every outcome of any observable $\hat{Q}$ is exactly predictable when the hidden variables are taken into account. That is, the outcome q of measurement of $\hat{Q}$ is given by

$$q = f(|\psi\rangle, \lambda).$$

Since quantum mechanics doesn't account for the hidden variables, we have to assign a probability distribution to the possible values of λ:

$$w(\lambda_1, \ldots, \lambda_N) = \text{ probability the hidden variables take values } \lambda_1, \ldots, \lambda_N.$$

These probabilities must sum to 1:

$$\int_\lambda w(\lambda) d\lambda = 1.$$

Then along came John Bell. He proved that if EPR were right about the hidden variables, then there is an experimentally testable inequality for correlations. In fact, the experiments have been done extensively since Bell's time, and they have confirmed beyond any doubt that EPR were *wrong*, and there are no such hidden variables.

Consider two spin-$\frac{1}{2}$ particles in a maximally entangled state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle\right)$$

where $|\uparrow\rangle$ is an eigenvector to $\hat{S}_3^{(i)}$ with eigenvalue $+\frac{\hbar}{2}$ and $|\downarrow\rangle$ is an eigenvector to $\hat{S}_3^{(i)}$ with eigenvalue $-\frac{\hbar}{2}$. Now consider two observables:

$$\hat{Q}_{\mathbf{a}}^{(1)} = \frac{2}{\hbar}\hat{\mathbf{S}}^{(1)} \cdot \mathbf{a} \qquad \hat{Q}_{\mathbf{b}}^{(2)} = \frac{2}{\hbar}\hat{\mathbf{S}}^{(2)} \cdot \mathbf{b} \qquad \text{where } \mathbf{a}, \mathbf{b} \text{ are unit vectors.}$$

The following are easy consequences:

$$\begin{aligned} \overline{Q}_{\mathbf{a}}^{(1)} &= \langle\psi|\hat{Q}_{\mathbf{a}}^{(1)}|\psi\rangle = 0 \\ \overline{Q}_{\mathbf{b}}^{(2)} &= \langle\psi|\hat{Q}_{\mathbf{b}}^{(2)}|\psi\rangle = 0. \end{aligned}$$

An ensemble of n runs of measurements of these observables on (identical copies of) the given state would some sequence of pairs of 1s and -1's: $\{(q_1^{(1)}, q_1^{(2)}), (q_2^{(1)}, q_2^{(2)}), \ldots, (q_n^{(1)}, q_n^{(2)})\}$. Consider the expectation for the *correlation* between the outcomes of the two observables.

Definition 27

$$K(\boldsymbol{a}, \boldsymbol{b}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{\infty} q_n^{(1)} q_n^{(2)}.$$

The quantum mechanical prediction for K is

$$K(\mathbf{a}, \mathbf{b}) = \langle\psi|\hat{Q}_{\mathbf{a}}^{(1)} \hat{Q}_{\mathbf{b}}^{(2)}|\psi\rangle.$$

For example, suppose we choose

$$\mathbf{a} = (0, 0, 1)^T \quad , \quad \mathbf{b} = (\sin(\theta), 0, \cos(\theta))^T.$$

Then

$$K(\mathbf{a}, \mathbf{b}) = \frac{4}{\hbar^2} \langle\psi|\hat{S}_3^{(1)} \left(\sin(\theta)\hat{S}_1^{(2)} + \cos(\theta)\hat{S}_3^{(2)}\right)|\psi\rangle$$
$$= -\cos(\theta)$$

In general, we have

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

This is experimentally confirmed.

Remark: Choose for example $\mathbf{a} = \mathbf{b}$. Then $K(\mathbf{a}, \mathbf{b}) = -1$, and so spins are always found anti-aligned in this state if $\mathbf{a} = \mathbf{b}$ is chosen.

If EPR were correct, then all outcomes would be predetermined:

$$q^{(1)} = A(\mathbf{a}, \lambda) \quad , \quad q^{(2)} = B(\mathbf{b}, \lambda).$$

The λ have a distribution $w(\lambda)$ with $\int_\lambda w(\lambda) d\lambda = 1$. So EPR's claim demands

$$K(\mathbf{a}, \mathbf{b}) = \int_\lambda w(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) d\lambda.$$

We know already that $K(\mathbf{a}, \mathbf{a}) = -1$. So $A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda)$ (except possibly for aset of λ of measure 0). So this means

$$K(\mathbf{a}, \mathbf{b}) = -\int_\lambda w(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) d\lambda.$$

Now, we eliminate $w(\lambda)$ from the equation:

$$K(\mathbf{a}, \mathbf{b}) - K(\mathbf{a}, \mathbf{c}) = -\int_\lambda w(\lambda) A(\mathbf{a}, \lambda) \left(A(\mathbf{b}, \lambda) - A(\mathbf{c}, \lambda)\right) d\lambda$$
$$= -\int_\lambda w(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) \left(1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)\right) d\lambda.$$

Since $A(\mathbf{b}, \lambda)^2 = 1$, and $A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) \leq 1$, we get

$$|K(\mathbf{a}, \mathbf{b}) - K(\mathbf{a}, \mathbf{c})| \leq \int_\lambda w(\lambda) \left(1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)\right) d\lambda$$

which leads immediately to *Bell's Inequality*:

$$|K(\mathbf{a},\mathbf{b}) - K(\mathbf{a},\mathbf{c})| \leq 1 - K(\mathbf{b},\mathbf{c}).$$

Now, recall the experimentally confirmed quantum mechanical prediction that $K(\mathbf{a},\mathbf{b}) = -\mathbf{a}\cdot\mathbf{b}$, and choose $\theta = \frac{\pi}{4}$. Then, in Bell's inequality we would have

$$\left| -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \right| \leq 0$$

which is clearly false. So Bell's inequality is violated, proving that EPR must have been wrong. This results is called *Bell's Theorem*.

Remark: Bell's Theorem does not imply information transmission faster than the speed of light, since each subsystem of $|\psi\rangle$ is in a maximally mixed state whose density matrix is the identity, and the identity is the same in any basis.

12 The Adiabatic Theorem

12.1 Dynamical Phase for a Time-Independent Hamiltonian

First consider a system with a time-*in*dependent Hamiltonian $\hat{H}$. Suppose at time t_0 we have a system initially in an eigenstate $|E_n(t_0)\rangle = |E_n\rangle$ of $\hat{H}$, with energy eigenvalue E_n. Now, let $U(t)$ be the time-evolution operator for the system. Since $\hat{H}$ is time-independent, we have

$$\hat{U}(t) = e^{\frac{1}{i\hbar}(t-t_0)\hat{H}} \quad , \quad U(t_0) = \mathbf{1}.$$

Consider the eigenstate $|E_n(t_0)\rangle$ time-evolved for time t.

$$\begin{aligned}
|E_n(t)\rangle &= \hat{U}(t)|E_n(t_0)\rangle \\
&= e^{\frac{1}{i\hbar}(t-t_0)\hat{H}}|E_n(t_0)\rangle \\
&= \sum_{r=1}^{\infty} \frac{\left(\frac{1}{i\hbar}(t-t_0)\hat{H}\right)^r}{r!}|E_n(t_0)\rangle \\
&= \sum_{r=1}^{\infty} \frac{\left(\frac{1}{i\hbar}(t-t_0)\right)^r}{r!}\hat{H}^r|E_n(t_0)\rangle \\
&= \sum_{r=1}^{\infty} \frac{\left(\frac{1}{i\hbar}(t-t_0)\right)^r}{r!}E_n^r|E_n(t_0)\rangle \\
&= \sum_{r=1}^{\infty} \frac{\left(\frac{1}{i\hbar}(t-t_0)E_n\right)^r}{r!}|E_n(t_0)\rangle \\
&= e^{-\frac{i}{\hbar}(t-t_0)E_n}|E_n(t_0)\rangle
\end{aligned}$$

So we see that for any time t the system remains in the eigenstate $|E_n\rangle = |E_n(t_0)\rangle$ of $\hat{H}$, only picking-up a phase factor of $-\frac{1}{\hbar}(t-t_0)E_n$. As we know, such phase factors cancel out in the calculation of measurement expectations, and so are physically insignificant. A phase like this, which arises as a result of the time-evolution of the system, is called a *dynamical phase*.

12.2 The Adiabatic Theorem

Consider now the case of a time-*de*pendent Hamiltonian $\hat{H}(t)$, with eigenvectors $|E_n\rangle$ and eigenvalues $\{E_n\}$. We make the following assumptions about $\hat{H}(t)$.

Assumptions

1) $\hat{H}(t)$ is varying *continuously* over the time interval of interest, so $E_1(t), E_2(t), \ldots$ and $|E_1(t)\rangle, |E_2(t)\rangle, \ldots$ are *continuous* functions of t.

2) The spectrum of $\hat{H}(t)$ is *discrete* and *nondegenerate* throughout the time $0 \leq t \leq 1$ during which the Hamiltonian is changing.

Consider assumption (2). If there is a degeneracy at time t_c, then at that time there is an ambiguity in the ordering of the eigenvalues. It is to resolve such ambiguities that we must use the time-ordering symbol T in the time-evolution operator for a general time-dependent Hamiltonian. If, however, we assume that the spectrum remains nondegenerate for all $t_i \leq t \leq t_f$, we have a well-defined ordering of the eigenvalues and we can write the time-evolution operator as

$$\hat{U}(t) = e^{\frac{1}{i\hbar}\int_{t_0}^{t} \hat{H}(t')dt'} \quad , \quad U(t_0) = \mathbf{1}$$

without the need for the time-ordering symbol.

Note that assumption (2) may be relaxed, but it makes the analysis quite a bit more complicated, as we need something like the time-ordering symbol to allow us to "track" the eigenstates.

Now, suppose $\hat{H}$ is changing in compliance with above assumptions over the time interval $\Delta t = t_f - t_i$. Suppose the initial Hamiltonian is $\hat{H}(t_i) = \hat{H}_i$ and the final Hamiltonian is $\hat{H}(t_f) = \hat{H}_f$. Note that $\Delta(t)$ provides a measure of how fast the Hamiltonian varies between t_i and t_f. We introduce a new time variable s, which is offset to $s = 0$ at time t_i, and scaled by Δt:

$$s = \frac{t - t_i}{\Delta t}.$$

Note the time-correspondences

$s = 0 \Leftrightarrow t = t_i$, and

$s = 1 \Leftrightarrow t = t_f$.

Notice that the assumptions (1) and (2) above for the functions of t imply the equivalent conditions for the corresponding functions of s. There is one more technical assumption that must be made; that is, we assume that the first and second derivatives of $|E_j(s)\rangle\langle E_j(s)|$ with respect to s are well-defined and piecewise continuous throughout the time $0 \leq s \leq 1$. We express $\hat{H}(s)$ and $\hat{U}(s)$ as functions of the new time variable s. Our present goal is to determine the dependence of $U(s)$ on Δt. To emphasize this, we write $U_{\Delta t}(s)$. In the limiting case $\Delta t \to 0$ (i.e. a sudden change in the Hamiltonian) we have

$$\lim_{\Delta t \to 0} U_{\Delta t}(1) = \mathbf{1}.$$

So the state of the system remains unchanged between t_i and t_f.

We are interested in very gradual changes in the Hamiltonian, which are called *adiabatic* changes. These correspond to the limiting case $\Delta t \to \infty$. We consider this case in more detail. Suppose the eigenvalues of $\hat{H}$ are $E_1, E_2, \ldots$, and the eigenvectors are $|E_1\rangle, |E_2\rangle, \ldots$. These are, of course, all functions of t, and therefore also of s.

Theorem 27 ***(Adiabatic Theorem)***

$$\lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(s)|E_j(0)\rangle\langle E_j(0)| = |E_j(s)\rangle\langle E_j(s)|\lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(s) \qquad j=0,1,2,\dots$$

Now, to see consider what the theorem is saying, multiply on both sides by the j^{th} eigenstate of $\hat{H}_i$ (namely $|E_j(0)\rangle$).

$$\begin{aligned} & \lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(s)|E_j(0)\rangle\langle E_j(0)|E_j(0)\rangle = |E_j(s)\rangle\langle E_j(s)|\lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(t)|E_j(0)\rangle \\ \implies & \lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(s)|E_j(0)\rangle = \hat{P}_j(s)\lim_{\Delta t\to\infty}\hat{U}_{\Delta t}(s)|E_j(0)\rangle \qquad (*) \end{aligned}$$

The operator $\hat{P}_j(t) = |E_j(s)\rangle\langle E_j(s)|$ is the *projector* onto the j^{th} eigenspace of $\hat{H}(s)$. If we think in the Scrödinger picture of the eigenstates as time-dependent Schrödinger states, then $\hat{U}_{\Delta t}(s)|E_j(0)\rangle$ is the j^{th} eigenstate of $\hat{H}_i$ time-evolved to time s. So what equation $(*)$ is saying is that, in the limit, once we evolve the eigenstate forward to time s, the projector onto the j^{th} eigenspace of $\hat{H}(s)$ acts on the new state as the identity. This means that the new state *is* the j^{th} eigenstate of $\hat{H}(s)$. So the adiabatic theorem can be stated more informally as follows.

Theorem 28 ***(Adiabatic Theorem (restated))*** *For a slowly varying Hamiltonian, a system in the j^{th} eigenstate of the initial Hamiltonian follows to the j^{th} eigenstate of the new Hamiltonian.*

Loosely speaking, what is meant by "slowly varying Hamiltonian" is that the *characteristic time* for a noticeable change in the Hamiltonian is very large compared to the characteristic time required for a transition from one eigenstate to the next. The adiabatic theorem, when used in practice, is often called the *adiabatic approximation*.

12.3 Example

Consider an electron e^- in a magnetic field $\mathbf{B} = (B_x, B_y, B_z)$. Considering *only* the spin degree of freedom, the Hamiltonian is $\hat{H} = \beta\hat{\mathbf{S}}\cdot\mathbf{B}$, where $\mathbf{S} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$. Suppose the magnetic field has constant magnitude $|\mathbf{B}| = B_0$, and is precessing about the z-axis with a constant angular-momentum ω, sweeping out a cone with constant opening angle α.

The time-dependent magnetic field vector is

$$\mathbf{B}(t) = B_0\left(\sin(\alpha)\cos(\omega t), \sin(\alpha)\sin(\omega t), \cos(\alpha)\right).$$

So the Hamiltonian in the eigenbasis of $\hat{S}_z$ (spin operator in the z-direction) is

$$\hat{H}(t) = \beta\hat{\mathbf{S}}\cdot\mathbf{B} = \frac{\hbar\beta B_0}{2}\left(\sin(\alpha)\cos(\omega t)\sigma_x + \sin(\alpha)\sin(\omega t)\sigma_y + \cos(\alpha)\sigma_z\right),$$

where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices. As a matrix (in the $\hat{S}_z$-eigenbasis), this is

$$\hat{H}(t) = -\frac{\hbar\omega_1}{2}\begin{pmatrix}\cos\alpha & e^{-i\omega t}\sin\alpha \\ e^{i\omega t}\sin\alpha & -\cos\alpha\end{pmatrix}$$

where $\omega_1 = -\beta B_0$. Its normalised eigenvectors (still in the $\hat{S}_z$-eigenbasis) are

$$|\mu_\uparrow(t)\rangle = \begin{pmatrix}\cos\left(\frac{\alpha}{2}\right) \\ e^{i\omega t}\sin(\frac{\alpha}{2})\end{pmatrix} \quad , \quad |\mu_\downarrow(t)\rangle = \begin{pmatrix}\sin\left(\frac{\alpha}{2}\right) \\ -e^{i\omega t}\cos(\frac{\alpha}{2})\end{pmatrix}$$

with corresponding eigenvalues

$$\mu_\uparrow = -\frac{\hbar\omega_1}{2} \quad , \quad \mu_\downarrow = \frac{\hbar\omega_1}{2}$$

respectively. These eigenvectors represent *spin-up* and *spin-down* along the instantaneous direction of $\mathbf{B}(t)$. In the language we used in the discussion of the adiabatic theorem above, $|\mu_\uparrow(t)\rangle$ is the $j = 1^{st}$ eigenstate of $\hat{H}(t)$, and $|\mu_\downarrow(t)\rangle$ is the $j = 2^{nd}$ eigenstate of $\hat{H}(t)$. Suppose the electron is initially in the first (spin-up) eigenstate of $\hat{H}(t_0)$ (to simplify the expression, let $t_0 = 0$):

$$|\mu(0)\rangle = \begin{pmatrix} \cos\left(\frac{\alpha}{2}\right) \\ \sin\left(\frac{\alpha}{2}\right) \end{pmatrix}.$$

In general, there will be some positive probability that a transition occurs to the second eigenstate (spin-down) of $\hat{H}(t)$, but in the adiabatic limit, we should find that this probability goes to zero.

The exact solution to the Schrödinger equation can be found to be

$$|\mu(t)\rangle = \begin{pmatrix} \left[\cos\left(\frac{\lambda t}{2}\right) + i\frac{(\omega_1+\omega)}{\lambda}\sin\left(\frac{\lambda t}{2}\right)\right]\cos(\frac{\alpha}{2})e^{\frac{-i\omega t}{2}} \\ \left[\cos\left(\frac{\lambda t}{2}\right) + i\frac{(\omega_1-\omega)}{\lambda}\sin\left(\frac{\lambda t}{2}\right)\right]\sin(\frac{\alpha}{2})e^{\frac{i\omega t}{2}} \end{pmatrix},$$

where $\lambda = \sqrt{\omega^2 + \omega_1^2 + 2\omega\omega_1\cos\alpha}$. In terms of the eigenvectors $|\mu_\uparrow(t)\rangle, |\mu_\downarrow(t)\rangle$, this is

$$|\mu(t)\rangle = \left[\cos\left(\frac{\lambda t}{2}\right) + i\frac{(\omega_1 + \omega\cos\alpha)}{\lambda}\sin\left(\frac{\lambda t}{2}\right)\right]e^{\frac{-i\omega t}{2}}|\mu_\uparrow(t)\rangle + i\left[\frac{\omega}{\lambda}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]e^{\frac{-i\omega t}{2}}|\mu_\downarrow(t)\rangle.$$

So the probability of a transition to spin-down at any time t (along the current direction of $\mathbf{B}(t)$) is

$$|\langle\mu(t)|\mu_\downarrow(t)\rangle|^2 = \left[\frac{\omega}{\lambda}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]^2.$$

The adiabatic approximation is valid when the characteristic time for changes in the Hamiltonian is much greater than the characteristic time for a state-transition from one eigenstate to another. The characteristic time for changes in the Hamiltonian in the present example is $1/\omega$ (since the $\mathbf{B}$-field is changing at the rate ω). The characteristic time for a phase change from spin-up to spin-down is $1/\omega_1$. So the adiabatic approximation is valid when $\omega \ll \omega_1$. In this case $\lambda \approx \omega_1$, and so

$$|\langle\mu(t)|\mu_\downarrow(t)\rangle|^2 = \left[\frac{\omega}{\omega_1}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]^2 \to 0$$

and so we see that in the adiabatic limit the probability of a transition from spin-up to spin-down is 0. This is exactly what the adiabatic theorem tells us; the system will remain in the first (spin-up) eigenstate $|\mu_\uparrow(t)\rangle$ as long as the Hamiltonian is changing adiabatically.

13 The Berry Phase

13.1 Geometric Phase

Suppose we are standing at the North Pole holding a (ideal, frictionless) pendulum. Say we set the pendulum in motion at the north pole, so that it swings directly along the longitudinal line passing through Waterloo. If we want to carry the pendulum without disturbing its natural oscillation, we must do so adiabatically. Any quick or sudden movements could disturb the orientation of the plane in which the pendulum is swinging. Now suppose we (adiabatically) carry the Pendulum down the longitudinal line, keeping the pendulum swinging north-south. We walk through Waterloo, and keep heading south, until we reach the equator.

Then, we walk (still adiabatically) along the equator, still keeping the pendulum swinging north-south. We walk some distance along the equator (say 1000 km) and finally walk north again directly back to the north pole. During the *entire journey*, the motion has been adiabatic, and the pendulum has been swinging north-south. It is fairly easy to see that when we return to the north pole, the pendulum is swinging in a different plane than it was originally; now it swings in the direction of the longitudinal line along which we travelled back to the north pole. Let θ be the angle between the original plane of the pendulum's oscillation, and the plane of it's oscillation after the motion.

The movement of the pendulum around the path while keeping the plane of its oscillation oriented in the north/south direction is an example of *parallel transport*. Consider the area A bounded by the path along which the pendulum was carried. It is a fraction $\frac{\theta}{2\pi}$ of the northern hemisphere of the earth:

$$A = \left(\frac{4\pi r^2}{2}\right)\left(\frac{\theta}{2\pi}\right) = \theta r^2$$

where r is the earth's radius. So we can write the angle θ by which the plane of the pendulums motion has rotated in terms of the area bounded by the path of the parallel transport:

$$\theta = \frac{A}{r^2}$$

(notice that this expression for θ is independent of the shape of the path, and depends *only* on the area enclosed by the path). This angle θ is an example of a *geometric phase*. In this (classical) example, it arises because of the intrinsic curvature of the sphere (the geometric phase would not occur for parallel transport on the surface of a flat manifold).

13.2 Development of the Berry Phase

It is crucial to note that the adiabatic theorem leaves open the possibility that the eigenstate may accumulate an additional phase factor, on top of the dynamical phase we saw earlier. This is because "the j^{th} eigenstate" multiplied by some extra phase-factor is still "the j^{th} eigenstate". Let $|E_j(t)\rangle$ denote the j^{th} eigenstate of $\hat{H}(t)$, with eigenvalue $E_j(t)$. For a system initially in state $|E_j(t_0)\rangle$, the adiabatic theorem tells us that at time t the state is approximated by (and in the adiabatic limit is equal to)

$$|\psi(t)\rangle = e^{i\gamma_j(t)} e^{-\frac{i}{\hbar}\int_{t_0}^{t} E_j(t')dt'} |E_j(t)\rangle. \quad (1)$$

Here $-\frac{1}{\hbar}\int_{t_0}^{t} E_j(t')dt'$ is the dynamical phase. It is obtained in a manner analogous to what we did in Section 12.1 to find the dynamical phase $\frac{1}{\hbar}(t-t_0)E_n$ for time-independent $\hat{H}$. Where in that derivation we used $\hat{U}(t) = e^{\frac{1}{i\hbar}(t-t_0)\hat{H}}$, we now have $\hat{U}(t) = e^{\frac{1}{i\hbar}\int_{t_0}^{t}\hat{H}(t')dt'}$ (recall that our assumptions allow us to ignore the time-ordering symbol in the expression for $\hat{U}(t)$). The factor $\gamma_j(t)$ in equation (1) is the additional phase allowed by the adiabatic theorem. This $\gamma_j(t)$ is a *geometric phase* like the type we saw in Section 13.1, and is called the *Berry phase*. We wish to obtain an expression for $\gamma_j(t)$. To simplify the calculations, let $\theta_j(t)$ denote the dynamical phase:

$$\theta_j(t) := -\frac{1}{\hbar}\int_{t_0}^{t} E_j(t')dt' \quad (2).$$

To evaluate the Berry phase, plug equation (1) into the Schrödinger equation, and use the eigenvalue equation, obtaining

$$\begin{aligned} & i\hbar\frac{\partial}{\partial t}\left[e^{i\gamma_j(t)}e^{i\theta_j(t)}|E_j(t)\rangle\right] = \hat{H}(t)\left[e^{i\gamma_j(t)}e^{i\theta_j(t)}|E_j(t)\rangle\right] \\ \implies \; & i\hbar\left[i\frac{d\gamma_j}{dt}e^{i\gamma_j(t)}e^{i\theta_j(t)}|E_j(t)\rangle - \frac{i}{\hbar}E_j(t)e^{i\gamma_j(t)}e^{i\theta_j(t)}|E_j(t)\rangle + e^{i\gamma_j(t)}e^{i\theta_j(t)}\frac{\partial}{\partial t}|E_j(t)\rangle\right] \\ & \qquad = e^{i\gamma_j(t)}e^{i\theta_j(t)}E_j(t)|E_j(t)\rangle \end{aligned}$$

which simplifies to

$$\frac{\partial}{\partial t}|E_j(t)\rangle + i|E_j(t)\rangle\frac{\partial\gamma_j}{\partial t} = 0.$$

Multiply this equation by $\langle E_j(t)|$ to obtain

$$\frac{\partial\gamma_j}{\partial t} = i\langle E_j(t)|\frac{\partial}{\partial t}|E_j(t)\rangle,$$

and an expression for the Berry phase

$$\gamma_j(t) = i\int_{t_0}^{t}\langle E_j(t')|\frac{\partial}{\partial t'}|E_j(t')\rangle dt' \quad (3).$$

13.3 When is the Berry Phase Nonzero?

The Hamiltonian $\hat{H}$ is dependent on time through some parameters $\{Q_i(t)\}$ which we assume are changing (adiabatically) in time. First, suppose there is only one such parameter Q which changes from Q_0 at time t_0 to Q_f at some later time t_f. Then E_j and $|E_j\rangle$ are both dependent on time through Q, and we have

$$\frac{\partial}{\partial t}|E_j(t)\rangle = \frac{\partial}{\partial Q}|E_j(Q)\rangle\frac{dQ}{dt}$$

Plugging this expression into equation (3) we get

$$\begin{aligned}\gamma_j(t_f) &= i\int_{t_0}^{t_f}\langle E_j(Q)|\frac{\partial}{\partial Q}|E_j(Q)\rangle\frac{dQ}{dt}dt \\ &= i\int_{Q_0}^{Q_f}\langle E_j(Q)|\frac{\partial}{\partial Q}|E_j(Q)\rangle dQ\end{aligned}$$

In particular, if the Hamiltonian returns to its original form at time t_f, then $Q_0 = Q_f$ and we have

$$\gamma_j(t_f) = i\int_{Q_0}^{Q_0}\langle E_j(Q)|\frac{\partial}{\partial Q}|E_j(Q)\rangle dQ = 0$$

and so the Berry Phase is 0. So in the case that only a single parameter in the Hamiltonian is changing in time, the Berry phase equals 0.

Now suppose that there are N such adiabatically-changing parameters $Q_1(t), Q_2(t), \ldots, Q_n(t)$. Then $\hat{H}, E_j$, and $|E_j\rangle$ are dependent on t through $\mathbf{Q} = (Q_1(t), Q_2(t), \ldots, Q_n(t))$, and we have

$$\begin{aligned}\frac{\partial}{\partial t}|E_j(t)\rangle &= \frac{\partial}{\partial Q_1}|E_j(\mathbf{Q})\rangle\frac{dQ_1}{dt} + \frac{\partial}{\partial Q_2}|E_j(\mathbf{Q})\rangle\frac{dQ_2}{dt} + \cdots + \frac{\partial}{\partial Q_N}|E_j(\mathbf{Q})\rangle\frac{dQ_N}{dt} \\ &= (\nabla_Q|E_j(Q)\rangle)\cdot\frac{d\mathbf{Q}}{dt},\end{aligned}$$

where ∇_Q is the gradient with respect to these parameters. So now we have

$$\gamma_j(t_f) = i\int_{\mathbf{Q}_0}^{\mathbf{Q}_f}\langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle\cdot d\mathbf{Q}$$

for the Berry phase. If the parameters of the Hamiltonian return to their initial form after time t_f, then this becomes

$$\gamma_j(t_f) = i\oint\langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle\cdot d\mathbf{Q} \quad (4)$$

which is a *line integral* around a closed loop in this parameter space. From complex analysis, we know that such an integral is *not* in general equal to zero. Note that $\gamma_j(t_f)$ depends only on the path taken, and not on the speed at which the parameters changed along that path. This is in contrast with the dynamical phase $\theta_j(t_f) = -\frac{1}{\hbar}\int_{t_0}^{t_f}E_j(t')dt'$ which depends on the elapsed time $t_f - t_0$.

13.4 The Berry Phase is Real

Note also that we can show that the Berry phase is real. If it were purely imaginary, then $e^{i\gamma_j(t)}$ would be an exponential factor and would be lost in the normalisation of $|E_j(t)\rangle$.

Proof that $\theta_j(t) \in \mathbb{R}$:

$$\begin{aligned}
& \nabla_Q \left(\langle E_j(\mathbf{Q})|E_j(\mathbf{Q})\rangle\right) = 0 \\
\implies \quad & \left(\nabla_Q \langle E_j(\mathbf{Q})|\right)|E_j(\mathbf{Q})\rangle + \langle E_j(\mathbf{Q})|\left(\nabla_Q|E_j(\mathbf{Q})\rangle\right) \\
& \qquad = \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle^* + \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle = 0 \\
\implies \quad & \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle \text{ is imaginary} \\
\implies \quad & \gamma_j(t) = i\int_{\mathbf{Q}_i}^{\mathbf{Q}_f} \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle \cdot d\mathbf{Q} \text{ is real.} \qquad \square
\end{aligned}$$

13.5 Alternative Expression for $\gamma_j(t)$

Recall equation (4) expresses $\gamma_j(t)$ after a complete cycle as a line-integral around a closed loop C in the parameter space:

$$\gamma_j(t_f) = i\oint \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle \cdot d\mathbf{Q}.$$

The assumption we made in Section 12.2 about the piecewise continuity of the first and second derivatives of $|E_j(s)\rangle\langle E_j(s)|$ with respect to s will ensure that the curve C is piecewise-smooth. If we assume that the parameter space is three-dimensional (so only parameters Q_1, Q_2, Q_3 of $\hat{H}$ are changing in time) then C is the boundary of some orientable surface S in the parameter space. Recalling vector calculus we can apply Stoke's theorem, and write $\gamma_j(t_f)$ as a *surface integral*

$$\gamma_j(t_f) = i\int_S [\nabla_Q \times \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle d] \cdot \mathbf{a},$$

where $\nabla_Q \times \langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle$ is the *curl* of $\langle E_j(\mathbf{Q})|\nabla_Q|E_j(\mathbf{Q})\rangle$.

Analogous results can be stated for parameter spaces with higher dimensions, but we will not pursue these here.

13.6 Example

Recall the example in Section 12.3, illustrating the adiabatic evolution of a spin-$\frac{1}{2}$ particle (an electron) in an adiabatically rotating magnetic field. We assumed the electron was initially in the spin-up eigenstate at time $t_0 = 0$, with the corresponding eigenvalue $\mu_\uparrow = -\frac{\hbar\omega_1}{2}$. Recall equation (2) for the dynamical phase:

$$\theta_j(t) := -\frac{1}{\hbar}\int_{t_0}^{t} E_j(t')dt'.$$

For the present example, the state is the $j = 1^{st}$ eigenstate (the spin-up eigenstate), and the eigenvalue $E_j(t)$ is a constant value $-\frac{\hbar\omega_1}{2}$. So the dynamical phase is found to be

$$\begin{aligned}
\theta_\uparrow(t) &= -\frac{1}{\hbar}\int_0^t \left(-\frac{\hbar\omega_1}{2}\right) dt' \\
&= \frac{\omega_1 t}{2}
\end{aligned}$$

Now recall equation (3) for the Berry phase:

$$\gamma_j(t) = i\int_{t_0}^{t}\langle E_j(t')|\frac{\partial}{\partial t'}|E_j(t')\rangle dt'.$$

Recalling the development of the example in Section 12.3, we have the $j = 1^{st}$ eigenstate written in the $\hat{S}_3$-eigenbasis:

$$|\mu_\uparrow(t)\rangle = \begin{pmatrix}\cos\left(\frac{\alpha}{2}\right)\\ e^{i\omega t}\sin\left(\frac{\alpha}{2}\right)\end{pmatrix},$$

and so

$$\frac{\partial}{\partial t}|\mu_\uparrow(t)\rangle = \begin{pmatrix}0\\ i\omega e^{i\omega t}\sin\left(\frac{\alpha}{2}\right)\end{pmatrix}$$

and

$$\langle\mu_\uparrow(t)|\frac{\partial}{\partial t}|\mu_\uparrow(t)\rangle = i\omega\sin^2\left(\frac{\alpha}{2}\right).$$

A complete cycle of the magnetic field takes time $t = 2\pi/\omega$, and so equation (3) now becomes

$$\begin{aligned}\gamma_\uparrow(t) &= i\int_0^{\frac{2\pi}{\omega}} i\omega\sin^2\left(\frac{\alpha}{2}\right)\\ &= \pi(\cos\alpha - 1) \qquad \left(\text{using the identity } 2\sin^2\left(\frac{\alpha}{2}\right) = 1-\cos\alpha\right).\end{aligned}$$

Now, consider the more general case in which the magnetic field vector $\mathbf{B}$ traces out an arbitrary closed path C (while $|\mathbf{B}| = B_0$ remains constant), returning to its starting position at time t_f. Now the angular momentum ω is not assumed to be constant; nor is there a constant "opening-angle" α. The spin-up eigenstate along $\mathbf{B}(t)$ is now

$$|\mu_\uparrow(t)\rangle = \begin{pmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\varphi}\sin\left(\frac{\theta}{2}\right)\end{pmatrix},$$

where $\varphi(t)$ and $\theta(t)$ are functions of time. From vector calculus, we recall how to write the gradient in spherical coordinates:

$$\begin{aligned}\nabla|\mu_\uparrow\rangle &= \frac{\partial}{\partial r}|\mu_\uparrow\rangle\hat{r} + \frac{1}{r}\frac{\partial}{\partial\theta}|\mu_\uparrow\rangle\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}|\mu_\uparrow\rangle\hat{\varphi}\\ &= \frac{1}{r}\begin{pmatrix}-\frac{1}{2}\sin\left(\frac{\theta}{2}\right)\\ \frac{1}{2}e^{i\varphi}\cos\left(\frac{\theta}{2}\right)\end{pmatrix}\hat{\theta} + \frac{1}{r\sin\theta}\begin{pmatrix}0\\ ie^{i\varphi}\sin\left(\frac{\theta}{2}\right)\end{pmatrix}\hat{\varphi}.\end{aligned}$$

So we get

$$\begin{aligned}\langle\mu_\uparrow|\nabla|\mu_\uparrow\rangle &= \frac{1}{2r}\left[-\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\hat{\theta} + \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\hat{\theta} + 2i\frac{\sin^2\left(\frac{\theta}{2}\right)}{\sin\theta}\hat{\varphi}\right]\\ &= i\frac{\sin^2\left(\frac{\theta}{2}\right)}{r\sin\theta}\hat{\varphi}.\end{aligned}$$

The curl is

$$\begin{aligned}\nabla\times\langle\mu_\uparrow|\nabla|\mu_\uparrow\rangle &= \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\left(\frac{i\sin^2\left(\frac{\theta}{2}\right)}{r\sin\theta}\right)\right]\hat{r}\\ &= \frac{i}{2r^2}\hat{r}.\end{aligned}$$

Recalling Section $\langle \mu_\uparrow | \nabla | \mu_\uparrow \rangle$, we can write $\gamma_\uparrow(t_f)$ as a surface integral over the surface S bounded by the curve C swept-out by the the magnetic field vector $\mathbf{B}$:

$$\gamma_\uparrow(t_f) = -\frac{1}{2} \int \frac{1}{r^2} \hat{r} \cdot d\mathbf{a}.$$

By the geometry of this example $d\mathbf{a} = r^2 d\Omega \hat{r}$, where Ω is the angle subtended at the origin, and so our integral becomes

$$\gamma_\uparrow(t_f) = -\frac{1}{2} \int d\Omega = -\frac{1}{2}\Omega.$$

14 APPENDIX: The δ-Function and Fourier Transforms

Here we give a brief summary of some aspects of δ-functions and Fourier Transforms which are relevant to quantum mechanics.

14.1 Definition of $\delta(x)$

The "Dirac delta function" $\delta(x)$ is a continuous generalization of the Kronecker delta symbol:

$$\delta_{n,m} = \begin{cases} 0 \text{ if } & m \neq n \\ 1 \text{ if } & m = n \end{cases} .$$

Suppose $f(x)$ is a function which is well-defined at the point $x = x_0$. Then the dirac delta function $\delta(x)$ is defined by

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0).$$

Thus $\delta(x)$ is a function which satisfies

$$\delta(x - x_0) = \begin{cases} 0 \text{ if } & x \neq x_0 \\ \infty \text{ if } & x = x_0 \end{cases}$$

and

$$\int_{-\infty}^{\infty} \delta(x - x_0)dx = 1.$$

14.2 Properties of $\delta(x)$

The following properties can be proved rigorously by means of distribution theory. We just state the properties without proof.

$\delta(x) = \delta(-x)$

$\delta(ax) = \frac{1}{|a|}\delta(x) \quad a \neq 0$

$\delta[g(x)] = \sum_{n:g(x_n)=0, g'(x_n)\neq 0} \frac{1}{|g'(x_n)|}\delta(x - x_n)$

$x\delta(x) = 0$

$f(x)\delta(x - a) = f(a)\delta(x - a)$

$\int_{-\infty}^{\infty} \delta(x - y)\delta(y - a)dy = \delta(x - a)$

$\delta(x) = \frac{a}{2\pi} \int_{-\infty}^{\infty} e^{ikx}dx$

Also, the first derivative $\delta'(x)$ of $\delta(x)$ has the following properties:

$\int_{-\infty}^{\infty} \delta'(x)f(x)dx = -f'(0)$

$\delta'(x) = -\delta'(-x)$

$\int_{-\infty}^{\infty} \delta'(x - y)\delta(y - a)dy = \delta'(x - a)$

$x\delta'(x) = -\delta(x)$

$x^2\delta'(x) = 0$

$\delta'(x) = \frac{i}{2\pi} \int_{-\infty}^{\infty} ke^{ikx}dk$

14.3 The Fourier Transform

If the "Fourier Transform" of a function $f(x)$ exists, it is the function

$$F(u) \equiv \mathcal{F}[f] = \sqrt{\frac{\alpha}{2\pi}} \int_{-\infty}^{\infty} e^{-i\alpha u x} f(x) dx$$

where α is a fixed constant. In quantum mechanics, α is taken to be $\frac{1}{\hbar}$, and so the Fourier transform for us is

$$F(u) \equiv \mathcal{F}[f] = \sqrt{\frac{1}{2\pi\hbar}} \int_{-\infty}^{\infty} e^{\frac{-iux}{\hbar}} f(x) dx.$$

The inverse fourier transform is

$$f(x) \equiv \mathcal{F}^{-1}[F] = \sqrt{\frac{1}{2\pi\hbar}} \int_{-\infty}^{\infty} e^{\frac{iux}{\hbar}} F(u) du.$$

Generally, if $f(x_1, \ldots, x_n)$ is a function of n variables, its Fourier Transform is

$$F(u_1, \ldots, u_n) \equiv \mathcal{F}[f] = \left(\frac{1}{2\pi\hbar}\right)^{\frac{n}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\frac{-i(u_1 x_1 + \ldots + u_n x_n)}{\hbar}} f(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

and the inverse Fourier Transform is

$$f(x_1, \ldots, x_n) \equiv \mathcal{F}[F] = \left(\frac{1}{2\pi\hbar}\right)^{\frac{n}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\frac{i(u_1 x_1 + \ldots + u_n x_n)}{\hbar}} F(u_1, \ldots, u_n) du_1 \ldots du_n.$$

14.4 Properties of the Fourier Transform

An *integrable function* $f(x)$ is one satisfying

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty.$$

Any integrable function has a Fourier transform.

In the following, we assume the Fourier Transforms of f and g exist, and $F(u) = \mathcal{F}[f]$, $G(u) = \mathcal{F}[g]$.

14.4.1 *Important Limit Property*

$$lim_{|u| \to \infty} F(u) = 0$$

14.4.2 *Linearity*

$$\mathcal{F}[af(x) + bg(x)] = aF(u) + bG(u)$$

14.4.3 *Fourier Transform of Derivatives*

If $f(x)$ is differentiable, and its derivative is integrable, then

$$\mathcal{F}[f'(x)] = \left(\frac{iu}{\hbar}\right) F(u).$$

Suppose $f(x)$ is m-times continuously differentiable, and its derivatives are integrable, then

$$\mathcal{F}\left[\frac{d^m f}{dx^m}\right] = \left(\frac{iu}{\hbar}\right)^m F(u).$$

The property is essentially that derivatives of $f(x)$ get transformed into multiplication by $\frac{iu}{\hbar}$. The converse is essentially the same:

$$\mathcal{F}[x^m f(x)] = \left(\frac{i}{\hbar}\right)^m \frac{d^m}{du^m}[F(u)].$$

14.4.4 *Multiplication $\Longleftrightarrow$ Convolution*

The *convolution* of two functions $f(x)$, $g(x)$ is defined as

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt = \int_{-\infty}^{\infty} f(x-t)g(t)dt$$

The following properties essentially say that multiplication of functions Fourier Transforms to convolution of functions (and vice-versa).

$$\mathcal{F}[(f * g)(x)] = \sqrt{2\pi\hbar}F(u)G(u)$$

$$\mathcal{F}[f(x)g(x)] = \frac{1}{\sqrt{2\pi\hbar}}(F * G)(u)$$

14.4.5 *Scaling and Shifting Properties*

For a constant c, we have

$$\mathcal{F}[f(cx_1, \ldots, cx_n)] = \frac{1}{|c|^n} F(\frac{u_1}{c}, \ldots, \frac{u_n}{c})$$

and

$$\mathcal{F}^{-1}[F(cu_1, \ldots, cu_n)] = \frac{1}{|c|^n} f(\frac{x_1}{c}, \ldots, \frac{x_n}{c})$$

We give the shifting properties for the case of a function in one variable:

$$\mathcal{F}[f(x-a)] = e^{-iua}F(u)$$

and

$$\mathcal{F}[e^{-i\lambda x}f(x)] = F(u-\lambda)$$

14.4.6 *Fourier Transform of* $\delta(x)$

$$\mathcal{F}[\delta(x-a)] = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \delta(x-a)e^{-iux}dx = \frac{1}{2\pi\hbar}e^{iua}.$$