QUANTUM MECHANICS THEORETICAL MINIMUM – NOTES

This book is half heuristic and half rigorous. My book notes fill in many of the gaps. For consistency I have preserved the book's equation numbers but adding rigor often required me to present topics in a slightly different order. Thus many of the book's numbered equations are out of order in these notes. I have also added my own numbered equations, usually by inserting an extra decimal place like (1.03). For my own readability I also attempted to be more uniform with symbology than the book and I added background details for a few topics that interested me such as Bell's Theorem.

In Chapters 1-3 I am especially careful at showing all steps to convince myself that the many "obvious" bra-ket manipulations are indeed legitimate.

Chapter 1. Systems and Experiments

Definition. When a **quantum system** is measured there is a discrete set of possible outcomes. The number of possible outcomes is called the **degrees of freedom** of the system. A **quantum state** is vector that represents the status of a system in terms of the possible outcomes. Thus, a quantum state is also called a **state vector.** We say that measurement of a state **prepares the system.**

In quantum experiments, performing the same measurement repeatedly gives the same answer until the system is prepared differently, such as measuring spin along a different axis

Notation. Denote a **vector** in $\mathbb{R}^n$ as $\vec{\boldsymbol{v}}$ and a **unit vector** as $\hat{\boldsymbol{v}}$.

Definition. **A spin state is a superposition (i.e., linear combination) of the outcomes**, such as a linear combination of the up- and down-spin vectors.

Notation. In a spin system we use the symbol σ **to represent measurement of spin** along an axis. The possible outcomes of a measurement are +1 ("up vector") and -1 ("down vector"). Thus, a spin system has 2 degrees of freedom.

A Quantum Experiment. Prepare a spin state of +1 in some direction $\hat{m}$. Then measure spin in some other direction $\hat{n}$. In a classical experiment the result would be **cos**θ where θ is the angle between $\hat{m}$ and $\hat{n}$. In a quantum experiment the result, strangely, is always ±1. However, repeat the quantum experiment many times, preparing the system in direction $\hat{m}$ and measuring in direction $\hat{n}$. The **average** obtained will be the classical result:

$$\langle\sigma\rangle = \hat{n}\cdot\hat{m} = \cos\theta \qquad (1.01)$$

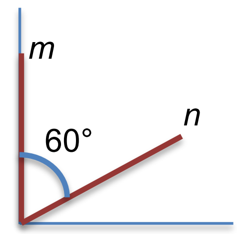
Definition. Denote the **probability that a measurement yields state 1** by **P(1)** and **state -1** by **P(-1).** Since no other outcome is possible, $\mathrm{P}(1)+\mathrm{P}(-1)=1$.

Theorem. $\langle\sigma\rangle \overset{(1.01)}{=} \cos\theta \;\Leftrightarrow\; \mathrm{P}(1)=\cos^2\frac{\theta}{2}$ (1.02)

Proof. $\langle\sigma\rangle=\cos\theta \;\Leftrightarrow\; \cos\theta=\mathrm{P}(1)\cdot 1+\left[1-\mathrm{P}(1)\right]\cdot(-1)=2\mathrm{P}(1)-1$

$$\Leftrightarrow\; \mathrm{P}(1)=\frac{1+\cos\theta}{2}=\cos^2\frac{\theta}{2} \qquad \blacksquare$$

Example. Let $\hat{m}$ and $\hat{n}$ be as in the figure. Then



$\langle\sigma\rangle \overset{(1.01)}{=} \cos(\theta)=\cos(60^\circ)=\frac{1}{2}$ and

$\mathrm{P}(1) \overset{(1.02)}{=} \cos^2\left(\frac{\theta}{2}\right)=\cos^2(30^\circ)=\left(\frac{\sqrt{3}}{2}\right)^2=\frac{3}{4}$. Then P(-1) = ¼.

Check: $\langle\sigma\rangle=1\left(\frac{3}{4}\right)+(-1)\left(\frac{1}{4}\right)=\frac{1}{2}$. ✔

Example. $\mathrm{P}(A\cup B)\neq \mathrm{P}(B\cup A)$

Spin is prepared as +1 in the *z*-direction.
Event A: Measure spin in the *z*-direction.
Event B: Measure spin in the *x*-direction.

$A\cup B$: $\mathrm{P}(A)=1$ so $\mathrm{P}(A\cup B)=1$.

$B\cup A$: $\mathrm{P}(B)=\frac{1}{2}$ and then $\mathrm{P}(A\,|{}^{\sim}B)=\frac{1}{4}$, resulting in $\mathrm{P}(B\cup A)=\frac{3}{4}$.

Remark. The foundations of set theory are different in the quantum domain!

Example. We measure +1 in *z*-direction (Event A). Then we measure +1 in the *y*-direction (Event B). **This experiment cannot be confirmed**. If we re-measure A we get a random number. This is an example of the **Heisenberg Uncertainty Principle**: we cannot with certainty know the outcomes of both events A and B.

We now introduce some mathematical frameworks to use for quantum states.

Definition. A **vector space** is a collection of elements called vectors where addition and scalar multiplication are defined and satisfy certain rules such as commutativity and associativity. A **Hilbert space** is a complete inner-product space. It is a vector space with an inner product operator, and thus also has a norm. If we represent vectors as points, "**complete**" means that every converging sequence has a limit point that belongs to the space.

A **basis** (also called an **algebraic basis**) for a vector space is a linearly independent set of vectors that spans the space in the sense that every vector can be expressed uniquely as a finite linear combination of basis vectors.

If the vector space is infinite-dimensional, in addition to an algebraic basis there is the concept of a **complete basis** (or **Schauder basis**) in which every vector can be uniquely expressed as a countably-infinite linear combination of basis vectors. Clearly every algebraic basis is a complete basis but not necessarily vice-versa. In a finite-dimensional vector space we drop the adjectives and simply use "basis" since every basis is algebraic.

In an inner product space if the basis vectors are mutually orthogonal (i.e., inner product equals 0), the basis is called an **orthogonal basis.** If in addition each basis vector is a unit vector, it is called an **orthonormal basis.**

Definition. $|\boldsymbol{A}\rangle$ is called a **ket**. It represents a state. In a finite-dimensional vector space a ket is represented by a column vector. $\langle\boldsymbol{A}|$ is the **bra** that corresponds to the ket $|A\rangle$. It is a row vector, the complex conjugate of the transpose of $|A\rangle$:

$$\langle A| = |A\rangle^{T^*} \tag{1.03}$$

We denote the **correspondence between a bra and a ket** by

$$|\boldsymbol{A}\rangle \Leftrightarrow \langle\boldsymbol{A}|. \tag{1.04}$$

Notation. For a given dimension n the set of kets (i.e., states) forms a complex vector space, specifically a Hilbert space. Let $\{\,|\,i\rangle\,\}_{i=1}^{n}$ be a basis for the vector space where

$$|\,i\rangle = \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix}. \tag{1.05}$$

Thus,

$$\langle i\,| = \begin{pmatrix} 0 & \cdots & 1_i & \cdots & 0 \end{pmatrix}, \tag{1.06}$$

Then a vector can be written as

$$\left|A\right\rangle = \sum_{i=1}^{n} \alpha_i \left|i\right\rangle = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{1.3}$$

where α_i are complex numbers called the **components of *A*.** So

$$\left\langle A\right| \overset{(1.03)}{=} \begin{pmatrix} \alpha_1^* & \cdots & \alpha_n^* \end{pmatrix} = \sum_i \left\langle i\right| \alpha_i^* \tag{1.07}$$

For example,

$$\text{if } \left|A\right\rangle = \begin{pmatrix} z \\ w \end{pmatrix} \text{ then } \left\langle A\right| = \left(z^* \;\; w^*\right) \tag{1.08}$$

Caution. $\left\langle A\right|$ consists of complex conjugates yet we don't label it $\left\langle A^*\right|$.

Definition. Let

$$\left|B\right\rangle = \sum_{j=1}^{n} \beta_j \left|j\right\rangle \text{ and } \left|C\right\rangle = \sum_{k=1}^{n} \gamma_k \left|k\right\rangle . \tag{1.09}$$

Vector space **addition** is defined by

$$\left|A\right\rangle + \left|B\right\rangle = \sum_i \left(\alpha_i + \beta_i\right) \left| i\right\rangle \tag{1.10}$$

and thus

$$\left\langle A\right| + \left\langle B\right| \overset{(1.07)}{=} \sum_i \left\langle i\right| \left(\alpha_i^* + \beta_i^*\right) . \tag{1.11}$$

Scalar multiplication by a complex number α is defined by

$$\left|\boldsymbol{\alpha A}\right\rangle \equiv \alpha \left|A\right\rangle = \left|A\right\rangle \alpha \overset{(1.3)}{=} \alpha \sum_{i=1}^{n} \alpha_i \left|i\right\rangle . \tag{1.12}$$

Thus

$$\left\langle \boldsymbol{\alpha A}\right| \equiv \left\langle A\right| \alpha^* = \alpha^* \left\langle A\right| = \alpha^* \sum_i \left\langle i\right| \alpha_i^* : \tag{1.13}$$

$$\left\langle \boldsymbol{\alpha A}\right| \overset{(1.03)}{=} \left|\alpha A\right\rangle^{T^*} \overset{(1.12)}{=} \left(\alpha\left|A\right\rangle\right)^{T^*} = \alpha^* \left|A\right\rangle^{T^*} \overset{(1.03)}{=} \alpha^* \left\langle A\right| \overset{(1.07)}{=} \alpha^* \sum_i \left\langle i\right| \alpha_i^* .$$

$$\left\langle A\right| \alpha^* = \alpha^* \left\langle A\right|$$

is nothing more than (scalar) (vector) = (vector) (scalar). ✔

The **Inner product** of two vectors is defined by

$$\left\langle A \,|\, B\right\rangle \equiv \left\{ \left\langle A\right| \right\} \left\{ \left|B\right\rangle \right\} . \tag{1.14}$$

Thus the Inner product of two *basis vectors* is

$$\left\langle i\middle|j\right\rangle = \delta_{ij} = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases} : \tag{1.15}$$

$$\langle i|j\rangle \overset{(1.14)}{=} \big(\langle i|\big)\big(|j\rangle\big) \overset{(1.05,\ 10.6)}{=} \begin{pmatrix} 0 & \cdots & 1_i & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1_j \\ \vdots \\ 0 \end{pmatrix} = \delta_{ij}$$

Hence

$$\langle A|B\rangle = \sum_i \alpha_i^* \beta_i : \tag{1.2}$$

$$\langle A|B\rangle \overset{(1.14)}{=} \big(\langle A|\big)\big(|B\rangle\big) \overset{(1.07,\ 1.09)}{=} \sum_i \sum_j \alpha_i^* \beta_j \langle i|j\rangle \overset{(1.15)}{=} \sum_i \sum_j \alpha_i^* \beta_j \delta_{ij} = \sum_i \alpha_i^* \beta_i$$

Definition. $\langle \boldsymbol{A}|\boldsymbol{\alpha}|\boldsymbol{B}\rangle \equiv \langle \boldsymbol{A}|\boldsymbol{\alpha B}\rangle$. (1.16)

Theorem. $\langle A|\alpha|B\rangle = \alpha\langle A|B\rangle = \langle \alpha^* A|B\rangle = \langle A|\alpha B\rangle$ (1.17)

Proof. $\langle A|\alpha|B\rangle \overset{(1.16)}{=} \langle A|\alpha B\rangle \overset{(1.14)}{=} \big(\langle A|\big)\big(|\alpha B\rangle\big) \overset{(1.12)}{=} \alpha\big(\langle A|\big)\big(|B\rangle\big) \overset{(1.14)}{=} \alpha\langle A|B\rangle$,

$$\alpha\big(\langle A|\big)\big(|B\rangle\big) \overset{(1.13)}{=} \big(\langle \alpha^* A|\big)\big(|B\rangle\big) \overset{(1.14)}{=} \langle \alpha^* A|B\rangle. \quad \blacksquare$$

Theorem. (Bra-Ket Correspondence Rules)

(1) $|A\rangle + |B\rangle \;\Leftrightarrow\; \langle A| + \langle B|$ (1.18)

(2) $\alpha|A\rangle \;\Leftrightarrow\; \langle A|\alpha^*$ (1.19)

(3) $|\alpha A\rangle \Leftrightarrow \langle \alpha A|$ (1.20)

Proof:

(1) Let

$$|D_i\rangle = (\alpha_i + \beta_i)|i\rangle \tag{a}$$

Claim $\langle D_i| = \langle i|(\alpha_i^* + \beta_i^*)$: (b)

$$\langle D_i| \overset{(1.03)}{=} |D_i\rangle^{T^*} \overset{(a)}{=} \big[(\alpha_i + \beta_i)|i\rangle\big]^{T^*} = (\alpha_i + \beta_i)^* |i\rangle^{T^*} \overset{(1.03)}{=} (\alpha_i^* + \beta_i^*)\langle i| \quad ✔$$

Thus

$$|A\rangle + |B\rangle \overset{(1.10)}{=} \sum_i (\alpha_i + \beta_i)|i\rangle \overset{(a)}{=} \sum_i |D_i\rangle \overset{(1.04)}{\Leftrightarrow} \sum_i \langle D_i|$$

$$\overset{(b)}{=} \sum_i \langle i|(\alpha_i^* + \beta_i^*) = \sum_i \langle i|\alpha_i^* + \sum_i \langle i|\beta_i^* \overset{(1.07)}{=} \langle A| + \langle B| \quad ✔$$

(2) Let

$$\left|A_i\right\rangle = \alpha\,\alpha_i\left|i\right\rangle \overset{(1.12)}{=} \left|\alpha\alpha_i\, i\right\rangle \cdot \tag{c}$$

Then

$$\left\langle A_i\right| \overset{(1.13)}{=} \left\langle i\right|\alpha^*\,\alpha_i^* \cdot \tag{d}$$

Thus

$$\alpha\left|A\right\rangle \overset{(1.3)}{=} \alpha\sum_i \alpha_i\left|\, i\right\rangle = \sum_i \alpha\,\alpha_i\left|\, i\right\rangle \overset{(c)}{=} \sum_i\left|\, A_i\right\rangle$$

$$\overset{(1.04,\ 1.18)}{\Leftrightarrow} \sum_i \left\langle A_i\right| \overset{(d)}{=} \sum_i \left\langle i\right|\alpha * \alpha_i^* \overset{(1.13)}{=} \alpha^* \sum_i \left\langle i\right|\alpha_i^* \overset{(1.07)}{=} \left\langle A\right|\alpha^* \quad ✔$$

(3) Let $\left|C\right\rangle = \left|\alpha A\right\rangle$. Then

$$\left|\alpha A\right\rangle = \left|C\right\rangle \overset{(1.04)}{\Leftrightarrow} \left\langle C\right| \overset{(1.03)}{=} \left|C\right\rangle^{T^*} = \left|\alpha A\right\rangle^{T^*} \overset{(1.12)}{=} \left(\alpha\left|A\right\rangle\right)^{T^*} = \alpha^*\left|A\right\rangle^{T^*} \overset{(1.03)}{=} \alpha^*\left\langle A\right| \overset{(1.13)}{=} \left\langle \alpha A\right| . \ \blacksquare$$

Theorem. (Inner Product Rules)

$$\left\langle C\right|\left\{\left|A\right\rangle + \left|B\right\rangle\right\} = \left\langle C\,|\,A\right\rangle + \left\langle C\,|\,B\right\rangle \tag{1.21}$$

$$\left\{\left\langle A\right| + \left\langle B\right|\right\}\left|C\right\rangle = \left\langle A|C\right\rangle + \left\langle B|C\right\rangle \tag{1.21b}$$

$$\left\langle B\,|\,A\right\rangle = \left\langle A\,|\,B\right\rangle^* \tag{1.22}$$

$$\left|A\right\rangle = \left|B\right\rangle \Rightarrow \left\langle C|A\right\rangle = \left\langle C|B\right\rangle \text{ and } \left\langle A|C\right\rangle = \left\langle B|C\right\rangle \tag{1.23}$$

Proof:

1.21: The first rule is a consequence of matrix multiplication as follows.

$$\left\langle C\right|\left\{\left|A\right\rangle + \left|B\right\rangle\right\} \overset{(1.07,\ 1.10)}{=} \left(\sum_j \left\langle j\right|\gamma_j^*\right)\left(\sum_i (\alpha_i + \beta_i)\left|i\right\rangle\right)$$

$$= \left[\sum_j \gamma_j^* \begin{pmatrix} 0 & \cdots & 1_j & \cdots & 0 \end{pmatrix}\right]\left(\sum_i (\alpha_i + \beta_i)\begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix}\right)$$

$$= \sum_j \sum_i \gamma_j^* (\alpha_i + \beta_i) \begin{pmatrix} 0 & \cdots & 1_j & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix}$$

$$= \sum_j \gamma_i^* (\alpha_i + \beta_i) = \sum_i \gamma_i^* \alpha_i + \sum_i \gamma_i^* \beta_i \overset{(1.2)}{=} \langle C \mid A \rangle + \langle C \mid B \rangle \quad ✔$$

1.22: $\langle B | A \rangle \overset{(1.2)}{=} \sum_i \beta_i^* \alpha_i = \left(\sum_i \alpha_i^* \beta_i \right)^* \overset{(1.2)}{=} \langle A | B \rangle^*$ ✔

1.23: $|A\rangle = |B\rangle$

$$\Rightarrow \quad \langle C | A \rangle \overset{(1.14)}{=} (\langle C|)(|A\rangle) = (\langle C|)(|B\rangle) \overset{(1.14)}{=} \langle C | B \rangle \quad ✔$$

From (1.04), $\langle A| = \langle B|$

$$\Rightarrow \quad \langle A | C \rangle = (\langle A|)(|C\rangle) = (\langle B|)(|C\rangle) = \langle B | C \rangle \quad ✔$$

1.21b: Let $\langle D| = \langle A| + \langle B|$. Then

$$\{ \langle A| + \langle B| \} |C\rangle = \langle D | C \rangle \overset{(1.20)}{=} \langle C | D \rangle^* \overset{(1.18)}{=} \left(\langle C| \{ |A\rangle + |B\rangle \} \right)^*$$

$$\overset{(1.21)}{=} \left(\langle C | A \rangle + \langle C | B \rangle \right)^* = \langle C | A \rangle^* + \langle C | B \rangle^*$$

$$\overset{(1.22)}{=} \langle A | C \rangle + \langle B | C \rangle \quad ✔ \qquad ■$$

Theorem.

$$\langle j \,|\, A \rangle = \alpha_j . \tag{1.5}$$

Proof.

$$\langle j | A \rangle \overset{(1.3)}{=} \langle j | \sum_i \alpha_i | i \rangle \overset{(1.21)}{=} \sum_{i=1}^{n} \langle j | \alpha_i | i \rangle \overset{(1.17)}{=} \sum_{i=1}^{n} \alpha_i \langle j \,|\, i \rangle \tag{1.4}$$

$$\overset{(1.15)}{=} \sum_{i=1}^{n} \alpha_i \delta_{ij} = \alpha_j \qquad ■$$

Corollary.

$$\langle A | j \rangle = \alpha_j^* \tag{1.5b}$$

Proof. $\langle A | j \rangle \overset{(1.22)}{=} \langle j | A \rangle^* \overset{(1.5)}{=} \alpha_j^*$ ■

Definition. The **magnitude** of a vector $\left|A\right\rangle$ is

$$\left\|A\right\| = \sqrt{\left\langle A \middle| A \right\rangle} \tag{1.24}$$

The vector is **normalized** if $\left\langle A \middle| A \right\rangle = 1$; i.e., $\left|A\right\rangle$ is a **unit vector**. Vectors $\left|A\right\rangle$ and $\left|B\right\rangle$ are **orthogonal** if $\left\langle A \middle| B \right\rangle = 0$.

Convention. **The spin states** constitute a 2-dimensional Hilbert space. For $\alpha \in \mathbb{C}$ we consider $\left|A\right\rangle$ and $\left|\alpha A\right\rangle$ to be the same (or equivalent) spin state, and we always use a unit vector (such as $\dfrac{\left|A\right\rangle}{\left\|A\right\|}$) to represent it. Identification of equivalent states can be made rigorous by defining the state space to be the **vector space of equivalence classes**, but that complication is side-stepped in this book by using unit vectors to represent states.

Definition. $z = e^{i\phi}$ is called a **phase factor**, and ϕ is called a **phase angle.** Note that zz^* = 1 where $z^* = e^{-i\phi}$ is the complex conjugate of z. Thus, if $\hat{A}$ is a unit vector then so is $e^{i\phi}\hat{A}$. Since z is a scalar, $\hat{A}$ and $e^{i\phi}\hat{A}$ are considered to be the same spin state, and either can be used to represent it.

Chapter 2. Quantum States

x	y	z
ℓ = left	i = in	u = up
r = right	o = out	d = down

Consider the spin states $|\ell\rangle$, $|r\rangle$, $|i\rangle$, $|o\rangle$, $|u\rangle$, and $|d\rangle$ as shown in the table. Since they are unit vectors:

$$\langle\ell|\ell\rangle = \langle r|r\rangle = \langle i|i\rangle = \langle o|o\rangle = \langle u|u\rangle = \langle d|d\rangle = 1 \qquad (2.01)$$

The following pairs of spin states are mutually exclusive, hence orthogonal:

$$\langle u|d\rangle = 0 = \langle d|u\rangle,\ \ \langle r|\ell\rangle = 0 = \langle \ell|r\rangle,\ \ \langle i|o\rangle = 0 = \langle o|i\rangle. \qquad (2.3)$$

These inner products are zero because if we measure, for example, along the *z*-axis and obtain "up" and then we measure again, we cannot obtain "down".

Notation. From (2.01) and (2.3), $\{|u\rangle, |d\rangle\}$ is an orthonormal **basis** for the spin states. A generic spin state $|A\rangle$ is a unit vector that can be expressed in terms of the basis as

$$|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle \text{ where } |\alpha_u|^2 + |\alpha_d|^2 = \alpha_u{}^*\alpha_u + \alpha_d{}^*\alpha_d = 1. \qquad (2.4)$$

The components α_u and α_d are called **probability amplitudes**. That is, although α_u and α_d could be negative or even complex numbers and thus not probabilities, $|\alpha_u|^2 \le 1$, $|\alpha_d|^2 \le 1$, and $|\alpha_u|^2 + |\alpha_d|^2 = 1$. So $\alpha_u^*\alpha_u = |\alpha_u|^2$ and $\alpha_d^*\alpha_d = |\alpha_d|^2$ are probabilities. (See definition (2.2), below.)

$$\alpha_u = \langle u|A\rangle \text{ and } \alpha_d = \langle d|A\rangle: \qquad (2.1)$$

$$\langle u|A\rangle \overset{(1.14)}{=} \langle u|\{\alpha_u|u\rangle + \alpha_d|d\rangle\} \overset{(1.21)}{=} \langle u|\alpha_u|u\rangle + \langle u|\alpha_d|d\rangle$$

$$\overset{(1.17)}{=} \alpha_u\langle u|u\rangle + \alpha_d\langle u|d\rangle \overset{(2.01,\ 2.3)}{=} \alpha_u.$$

$\langle d|A\rangle = \alpha_d$ is proved similarly. ✔

Definition. If spin is measured in the up direction, denote $\mathbf{P}_u$ as **the probability that the result is +1.** If spin is measured in the down direction, denote $\mathbf{P}_d$ as the probability that the result is +1. An equivalent definition for P_d is the probability of -1 when measuring in the up direction. We similarly define $\boldsymbol{P}_\ell$, $\boldsymbol{P}_r$, $\boldsymbol{P}_i$, and $\mathbf{P}_o$.

$$\begin{cases} \mathbf{P}_u \equiv \alpha_u^* \, \alpha_u \overset{(2.1)}{=} \langle A|u\rangle\langle u|A\rangle \\ \mathbf{P}_d \equiv \alpha_d^* \, \alpha_d \overset{(2.1)}{=} \langle A|d\rangle\langle d|A\rangle \end{cases} \tag{2.2}$$

Definition. When the initial spin measurement is in the $|A\rangle$ direction, we say that **spin is prepared in state $|A\rangle$**.

Convention. For the remainder of these notes, except as otherwise specified, we will suppose that spin has been prepared in some state $|A\rangle$ and then measured in the *z*-direction. In particular, this is the condition for equations 2.5, 2.6, and 2.10 – 2.13 to hold.

A further convention is that $\langle A|B\rangle$ represents a spin that is prepared in the $|B\rangle$ direction and then measured in the $|A\rangle$ direction. This convention is embedded in definition (2.2) and also below in equations (2.8) and (2.9).

If spin is prepared in state $|r\rangle$, then $\mathrm{P}_u = \mathrm{P}_d = \frac{1}{2}$. We define $|r\rangle$ by equation (2.5), the simplest such expression for $|r\rangle$ that satisfies this.

Theorem. Define $|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle$. (2.5)

Then $\mathrm{P}_u = \mathrm{P}_d = \frac{1}{2}$.

Proof. Let $\alpha_u = \frac{1}{\sqrt{2}} = \alpha_d$. Then $|r\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle$ and, from 2.4 (with $|A\rangle = |r\rangle$),

$$\langle u|r\rangle \overset{(2.1)}{=} \alpha_u = \frac{1}{\sqrt{2}} \text{ and } \langle r|u\rangle \overset{(2.1)}{=} \alpha_u^* = \frac{1}{\sqrt{2}}.$$

So

$$\mathrm{P}_u \overset{(2.2)}{=} \alpha_u^* \, \alpha_u = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = \frac{1}{2} \text{ and } \mathrm{P}_d = \langle r|d\rangle\langle d|r\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = \frac{1}{2}. \quad ■$$

Theorem. $|\ell\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle$ (2.6)

Proof. We write $|\ell\rangle$ in terms of the basis vectors: $|\ell\rangle = \alpha|u\rangle + \beta|d\rangle$. So,

$$0 \overset{(2.3)}{=} \langle r|\ell\rangle = \left\{\langle u|\frac{1}{\sqrt{2}} + \langle d|\frac{1}{\sqrt{2}}\right\} \Big| \left\{ \alpha|u\rangle + \beta|d\rangle\right\} \overset{(2.01,\, 2.3)}{=} \frac{1}{\sqrt{2}}(\alpha+\beta) \quad \Rightarrow \quad \beta = -\alpha\,.$$

(2.6) is the simplest expression satisfying $\beta = -\alpha$ and

$$\mathrm{P}_u \overset{(2.2)}{=} \alpha_u^* \alpha_u \overset{(2.1)}{=} \langle \ell|u\rangle\langle u|\ell\rangle = \frac{1}{2}\,. \quad \blacksquare$$

If spin is prepared in either state $|o\rangle$ or $|i\rangle$, then $\mathrm{P}_u = \mathrm{P}_d = \tfrac{1}{2}$:

$$\begin{cases} \dfrac{1}{2} = \mathrm{P}_u = \langle o|u\rangle\langle u|o\rangle & \dfrac{1}{2} = \mathrm{P}_u = \langle i|u\rangle\langle u|i\rangle \\ \dfrac{1}{2} = \mathrm{P}_d = \langle o|d\rangle\langle d|o\rangle & \dfrac{1}{2} = \mathrm{P}_d = \langle i|d\rangle\langle d|i\rangle \end{cases} \tag{2.8}$$

If spin is prepared in either state $|o\rangle$ or $|i\rangle$ and then measured in state $|r\rangle$ or $|\ell\rangle$, then $\mathrm{P}_r = \mathrm{P}_\ell = \tfrac{1}{2}$:

$$\begin{cases} \dfrac{1}{2} = \mathrm{P}_r = \langle o|r\rangle\langle r|o\rangle & \dfrac{1}{2} = \mathrm{P}_r = \langle i|r\rangle\langle r|i\rangle \\ \dfrac{1}{2} = \mathrm{P}_\ell = \langle o|\ell\rangle\langle \ell|o\rangle & \dfrac{1}{2} = \mathrm{P}_\ell = \langle i|\ell\rangle\langle \ell|i\rangle \end{cases} \tag{2.9}$$

Exercise 2.3. $|i\rangle = \dfrac{1}{\sqrt{2}}|u\rangle + \dfrac{i}{\sqrt{2}}|d\rangle$ and $|o\rangle = \dfrac{1}{\sqrt{2}}|u\rangle - \dfrac{i}{\sqrt{2}}|d\rangle$ (2.10)

Notation. Since we have agreed upon a convention to *prepare in any direction* but to always *measure in the z-direction*, we choose our basis to be

$$|u\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.11}$$

and

$$|d\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.12}$$

Then

$$|r\rangle \overset{(2.5)}{=} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},\quad |\ell\rangle \overset{(2.6)}{=} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix},\quad |i\rangle \overset{(2.10)}{=} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix},\ \text{and}\quad |o\rangle \overset{(2.10)}{=} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} \end{bmatrix}. \tag{2.13}$$

Chapter 3. Principles of Quantum Mechanics

Definition. An **operator *M*** is a black box: $|A\rangle$ goes in, $|B\rangle$ comes out:

$$M|A\rangle = |B\rangle$$

Notation: In the discussion that follows we denote $\{|i\rangle\}$ to be an orthonormal basis for the Hilbert space of states. $M|j\rangle$ is a ket. Call it $|C\rangle$. So, $\langle k|M|j\rangle = \langle k|C\rangle \in \mathbb{C}$ is just a complex number. Call it m_{kj}. Then

$$\langle k|M|j\rangle = m_{kj} \tag{3.01}$$

and we can write *M* as a matrix of complex numbers:

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} = \left(m_{ij}\right) \tag{3.2}$$

Definition. The **Hermitian conjugate of *M*** is $\boldsymbol{M}^\dagger \equiv \left(M^T\right)^*$, the complex conjugate of the transpose of *M*.

Definition. $\langle \boldsymbol{A}|\boldsymbol{LM}|\boldsymbol{B}\rangle \equiv \left(\langle A|L\right)\left(M|B\rangle\right)$. Because $M = IM = MI$ we also get that $\langle \boldsymbol{A}|\boldsymbol{M}|\boldsymbol{B}\rangle = \left(\langle A|\right)\left(M|B\rangle\right) = \left(\langle A|M\right)\left(|B\rangle\right)$.

Theorem: Bra-Ket Rules for Operators

Correspondence:

$$M|A\rangle \Leftrightarrow \langle A|M^\dagger \tag{3.02}$$

Operators are linear:

$$M\,\alpha|A\rangle = \alpha\,M|A\rangle \quad \text{and} \quad \mathrm{M}\left\{|A_1\rangle + |A_2\rangle\right\} = M|A_1\rangle + M|A_2\rangle \tag{3.03}$$

Inner product rules:

$$\begin{aligned}\langle A|LM|B\rangle &= \left(\langle A|L\right)\left(M|B\rangle\right) = \left(\langle A|LM\right)|B\rangle = \langle A|\left(LM|B\rangle\right) \\ &= \sum_i\sum_j\sum_k \alpha_j^*\beta_i \ell_{jk} m_{ki}\end{aligned} \tag{3.04}$$

$$\langle A|M|B\rangle = \langle A|\left(M|B\rangle\right) = \left(\langle A|M\right)|B\rangle = \sum_i\sum_j \alpha_j^*\beta_i m_{ji} \tag{3.05}$$

$$L|A\rangle = M|B\rangle \Rightarrow \langle C|L|A\rangle = \langle C|M|B\rangle \tag{3.06}$$

$$\langle A|L = \langle B|M \Rightarrow \langle A|L|C\rangle = \langle B|M|C\rangle \tag{3.06b}$$

$$M|A\rangle = \alpha|B\rangle \Rightarrow \langle C|M|A\rangle = \langle C|\alpha|B\rangle \tag{3.07}$$

$$\langle A|L = \langle B|\alpha \Rightarrow \langle A|L|C\rangle = \langle B|\alpha|C\rangle \tag{3.07b}$$

Proof: Let $|A\rangle = \sum_i \alpha_i |i\rangle$, $|B\rangle = \sum_i \beta_i |i\rangle$, $|C\rangle = M|A\rangle$, and $L = (\ell_{jk})$

3.02:

$$M|A\rangle = |C\rangle \overset{(1.04)}{\Leftrightarrow} \langle C| \overset{(1.03)}{=} |C\rangle^{\mathrm{T}^*} = (M|A\rangle)^{\mathrm{T}^*} = \left[\begin{pmatrix} & & \vdots & & \\ m_{i1} & \cdots & m_{ij} & \cdots & m_{in} \\ & & \vdots & & \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_n \end{pmatrix}\right]^{\mathrm{T}^*}$$

$$= \begin{pmatrix} \vdots \\ \sum_j m_{ij}\alpha_j \\ \vdots \end{pmatrix}^{\mathrm{T}^*} = \begin{pmatrix} \cdots & \sum_j m_{ij}^* \alpha_j^* & \cdots \end{pmatrix}.$$

$$\langle A|M^\dagger = \begin{pmatrix} \alpha_1^* & \cdots & \alpha_i^* & \cdots & \alpha_n^* \end{pmatrix}\begin{pmatrix} & m_{j1}^* & \\ & \vdots & \\ \cdots & m_{ji}^* & \cdots \\ & \vdots & \\ & m_{jn}^* & \end{pmatrix} = \begin{pmatrix} \cdots & \sum_i \alpha_i^* m_{ji}^* & \cdots \end{pmatrix}$$

$$= \begin{pmatrix} \cdots & \sum_j \alpha_j^* m_{ij}^* & \cdots \end{pmatrix} \Leftrightarrow M|A\rangle. \ ✔$$

3.03: This is true since matrix multiplication is linear.

3.04: By definition, $\langle A|LM|B\rangle = (\langle A|L)(M|B\rangle)$. So,

$$\langle A|LM|B\rangle = \langle A|(LM)I|B\rangle = (\langle A|LM)(I|B\rangle) = (\langle A|LM)|B\rangle$$

and

$$\langle A|LM|B\rangle = \langle A|I(LM)|B\rangle = (\langle A|I)(LM|B\rangle) = \langle A|(LM|B\rangle).$$

Moreover,

$$\langle A|L = \begin{pmatrix} \alpha_1^* & \cdots & \alpha_n^* \end{pmatrix}\begin{pmatrix} \ell_{11} & & \ell_{1n} \\ \vdots & \cdots & \vdots \\ \ell_{n1} & & \ell_{nn} \end{pmatrix} = \begin{pmatrix} \sum_j \alpha_j^* \ell_{j1} & \cdots & \sum_j \alpha_j^* \ell_{jn} \end{pmatrix},$$

$$M|B\rangle = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ & \vdots & \\ m_{n1} & \cdots & m_{nn} \end{pmatrix}\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \sum_i m_{1i}\beta_i \\ \vdots \\ \sum_i m_{ni}\beta_i \end{pmatrix}, \text{ and}$$

$$\langle A|LM|B\rangle = \left(\langle A|L\right)\left(M|B\rangle\right) = \sum_{i,j}\alpha_j^*\beta_i\ell_{j1}m_{1i} + \cdots + \sum_{i,j}\alpha_j^*\beta_i\ell_{jn}m_{ni}.$$

$$= \sum_{i,j,k}\alpha_j^*\beta_i\ell_{jk}m_{ki} \qquad ✔$$

3.05: This is the special case of (3.04) where $L = I = (\delta_{jk})$:

$$\langle A|M|B\rangle \overset{(3.04)}{=} \sum_i\sum_j\sum_k \alpha_j{}^*\beta_i\delta_{jk}m_{ki} = \sum_i\sum_j \alpha_j{}^*\beta_i m_{ji} \qquad ✔$$

3.06: $\langle C|L|A\rangle \overset{(3.05)}{=} \langle C|\left(L|A\rangle\right) = \langle C|\left(M|B\rangle\right) \overset{(3.05)}{=} \langle C|M|B\rangle$ ✔

3.06b, 3.07, and 3.07b are proven similarly ✔ ■

Alternate Ways to Represent $M|A\rangle = |B\rangle$

$$M|A\rangle = \begin{pmatrix} & \vdots & \\ m_{i1} & \cdots & m_{in} \\ & \vdots & \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \vdots \\ \sum_j m_{ij}\alpha_j \\ \vdots \end{pmatrix} \text{ and } |B\rangle = \begin{pmatrix} \vdots \\ \beta_i \\ \vdots \end{pmatrix}.$$

Thus $M|A\rangle = |B\rangle$ can be written

$$\sum_j m_{ij}\alpha_j = \beta_i \qquad (3.3)$$

or

$$\begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \qquad (3.4)$$

The next theorem is motivated by Bra-Ket Rule (3.02).

Theorem. If $|B\rangle = M|A\rangle$ then $\langle B| = \langle A|M^\dagger$ (3.08)

Proof: Since $\langle A| \overset{1.04}{\Leftrightarrow} |A\rangle$ and $\langle A| \overset{(1.03)}{=} |A\rangle^{T^*}$ we see that the bra $\langle A|$ that corresponds to a ket $|A\rangle$ is unique. Both $\langle B|$ and $\langle A|M^\dagger$ correspond to $M|A\rangle$:

$$\langle B| \overset{1.04}{\Leftrightarrow} |B\rangle = M|A\rangle \text{ and } \langle A|M^\dagger \overset{3.02}{\Leftrightarrow} M|A\rangle$$

Thus $\langle B| = \langle A|M^\dagger$. ■

Eigenvectors and Eigenvalues

Definition. $|\lambda\rangle \neq 0$ is an **eigenvector of *M*** if $M|\lambda\rangle = \lambda|\lambda\rangle$. λ is called the **eigenvalue of M corresponding to $|\lambda\rangle$.** Because of the corollary to the next theorem, multiples of an eigenvector are considered to be the same eigenvector. Eigenvectors that are linearly independent are called **distinct eigenvectors.** In quantum mechanics, distinct eigenvectors that have the same eigenvalue are called **degenerate eigenvectors.**

Theorem. If $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are eigenvectors of *M* having the same eigenvalue λ, then any linear combination is also is an eigenvector of *M* with eigenvector λ.

Proof. Let $|\lambda_3\rangle = \alpha|\lambda_1\rangle + \beta|\lambda_2\rangle$. Then

$$M|\lambda_3\rangle \overset{(3.03)}{=} \alpha M|\lambda_1\rangle + \beta M|\lambda_2\rangle = \lambda\alpha|\lambda_1\rangle + \lambda\beta|\lambda_2\rangle = \lambda|\lambda_3\rangle. \quad ■$$

Corollary. If $|\lambda\rangle$ is an eigenvector of *M* with eigenvalue λ then $|\alpha\lambda\rangle$ is also an eigenvector of *M* with eigenvalue λ for any scalar α.

Proof. Set $\lambda = \lambda_1$ and $\beta = 0$. Then

$$M|\alpha\lambda\rangle \overset{(1.12)}{=} M\alpha|\lambda\rangle \overset{(3.03)}{=} \alpha M|\lambda\rangle \overset{(\text{Eigenvector})}{=} \lambda\alpha|\lambda\rangle \overset{(1.12)}{=} \lambda|\alpha\lambda\rangle \quad ■$$

Definition. A **Hermitian Operator** is an operator such that $M = M^\dagger$, or $m_{ji} = m_{ij}^*$.

Rule. Operators are Hermitian:

$$M|A\rangle \Leftrightarrow \langle A|M \qquad (3.8)$$

Thus

$$M|\lambda\rangle \Leftrightarrow \langle\lambda|M \qquad (3.9)$$

The following theorem is needed in Examples 1 and 2 in Chapter 8. It is not trivial, but the book simply presents it as a fact.

Theorem. M is Hermitian iff $\langle\Psi|M|\Phi\rangle = \langle\Phi|M|\Psi\rangle^* \quad \forall\Psi,\Phi:$ (3.09)

Proof. Suppose M is Hermitian. Let $\{|i\rangle\}_{i=1}^n$ be an orthonormal basis for the Hilbert space of states, $|\Psi\rangle = \sum_{i=1}^n \psi_i|i\rangle$, and $|\Phi\rangle = \sum_{j=1}^n \phi_j|j\rangle$.

$$\langle\Psi|M|\Phi\rangle \overset{(3.05)}{=} \sum_{i,j} \psi_j^* \phi_i m_{ji}, \tag{a}$$

$$\langle\Phi|M|\Psi\rangle \overset{(3.05)}{=} \sum_{i,j} \phi_j^* \psi_i m_{ji} = \sum_{i,j} \phi_i^* \psi_j m_{ij}, \tag{b}$$

$$\langle\Phi|M|\Psi\rangle^* \overset{(b)}{=} \sum_{i,j} \psi_j^* \phi_i m_{ij}^* \overset{(\text{Hermitian})}{=} \sum_{i,j} \psi_j^* \phi_i m_{ji} \overset{(a)}{=} \langle\Psi|M|\Phi\rangle. \qquad ✔$$

Conversely, suppose M is not Hermitian. Then $M \neq M^\dagger$. We wish to find $|\Psi\rangle$ and $|\Phi\rangle$ such that $\langle\Phi|M|\Psi\rangle^* \neq \langle\Psi|M|\Phi\rangle$. Since $M \neq M^\dagger$, $m_{ji} \neq m_{ij}^*$ for some i and j. Set $|\Phi\rangle = |i\rangle$ and $|\Psi\rangle = |j\rangle$.

Denote $M = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix}$, $|i\rangle = \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix}$ and $|j\rangle = \begin{pmatrix} 0 \\ \vdots \\ 1_j \\ \vdots \\ 0 \end{pmatrix}$. Then,

$$M|i\rangle = \begin{pmatrix} m_{11} & \cdots & m_{1i} & \cdots & m_{1n} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \cdots & m_{ni} & \cdots & m_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} m_{1i} \\ \vdots \\ m_{ni} \end{pmatrix},$$

$$\langle i|M = \begin{pmatrix} 0 & \cdots & 1_i & \cdots & 0 \end{pmatrix} \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{i1} & \cdots & m_{in} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix} = \begin{pmatrix} m_{i1} & \cdots & m_{in} \end{pmatrix},$$

$$\langle\Psi|M|\Phi\rangle = \langle j|M|i\rangle = \langle j|\left(M|i\rangle\right) = \begin{pmatrix} 0 & \cdots & 1_j & \cdots & 0 \end{pmatrix} \begin{pmatrix} m_{1i} \\ \vdots \\ m_{ni} \end{pmatrix} = m_{ji},$$

$$\langle\Phi|M|\Psi\rangle = \langle i|M|j\rangle = \left(\langle i|M\right)|j\rangle = \begin{pmatrix} m_{i1} & \cdots & m_{in} \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix} = m_{ij},$$

and so

$$\langle\Phi|M|\Psi\rangle^* = m_{ij}^* \neq m_{ji} = \langle\Psi|M|\Phi\rangle \qquad \blacksquare$$

Definition. An **Observable** is something that can be measured (like spin in a certain direction). The result of a measurement is a real number. (Note. This may be worth challenging/investigating. Perhaps we only observe the real part of a measurement. For example, the current between the plates of a capacitor is imaginary. We don't observe it but it can be measured.) Observables are represented by Hermitian operators (see Principle 1, next page).

Theorem. The eigenvalues of a Hermitian matrix are real.

Proof. Let L be a Hermitian matrix and λ an eigenvalue. Then

$$(1)\ L|\lambda\rangle = \lambda|\lambda\rangle.$$

The corresponding bra equation is

$$(2)\ \langle\lambda|L \overset{(3.9,\,1.19)}{=} \langle\lambda|\lambda^*.$$

Therefore

$$(1')\ \langle\lambda|L|\lambda\rangle \overset{(3.07)}{=} \langle\lambda|\lambda|\lambda\rangle \overset{(1.17)}{=} \lambda\langle\lambda|\lambda\rangle$$

and

$$(2')\ \langle\lambda|L|\lambda\rangle \overset{(3.07b)}{=} \langle\lambda|\lambda^*|\lambda\rangle \overset{(1.17)}{=} \lambda^*\langle\lambda|\lambda\rangle.$$

Hence $\lambda = \lambda^*$, and so $\lambda \in \mathbb{R}$. $\blacksquare$

Lemma 1. If $\lambda_1 \neq \lambda_2$ are eigenvalues of a Hermitian operator L, then the corresponding eigenvectors are orthogonal.

Proof. $(1)\ \langle\lambda_1|L \overset{(2)}{=} \langle\lambda_1|\lambda_1^* \overset{(\text{Theorem})}{=} \langle\lambda_1|\lambda_1.$

Also,

$$(2)\ L|\lambda_2\rangle \overset{(1)}{=} \lambda_2|\lambda_2\rangle.$$

Therefore

$$(1')\ \langle\lambda_1|L|\lambda_2\rangle \overset{(3.07b)}{=} \langle\lambda_1|\lambda_1|\lambda_2\rangle \overset{(1.17)}{=} \lambda_1\langle\lambda_1|\lambda_2\rangle$$

and

$$(2')\ \langle\lambda_1|L|\lambda_2\rangle \overset{(3.07)}{=} \langle\lambda_1|\lambda_2|\lambda_2\rangle \overset{(1.17)}{=} \lambda_2\langle\lambda_1|\lambda_2\rangle.$$

Thus

$$\lambda_1\langle\lambda_1|\lambda_2\rangle = \lambda_2\langle\lambda_1|\lambda_2\rangle \;\Rightarrow\; \langle\lambda_1|\lambda_2\rangle = 0 \text{ since } \lambda_1 \neq \lambda_2. \quad ■$$

Lemma 2. If $|\lambda_1\rangle$ and $|\lambda_2\rangle$are distinct eigenvectors of a Hermitian operator L having a common eigenvalue λ, then a pair of orthonormal eigenvectors can be found that have eigenvalue λ.

Proof. Since $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are linearly independent, the Gram-Schmidt Orthogonalization Process yields a pair of orthonormal vectors that are linear combinations of $|\lambda_1\rangle$ and $|\lambda_2\rangle$. By the theorem on p. 14, both vectors have eigenvalue λ. ■

Fundamental Theorem (Exercise 3.1). The eigenvectors of a Hermitian operator on a complex vector space V can be chosen so as to form an orthonormal basis.

Principle 1. Observables are represented by Hermitian linear operators.

Principle 2. The result of a measurement is an eigenvalue of the corresponding Hermitian operator and the system state is the corresponding eigenvector.

Definition. Two states are **distinguishable** if there is a measurement that can clearly determine in which state the system was prepared.

There is no measurement that can unambiguously determine whether a system was prepared in state $|\ell\rangle$ vs $|u\rangle$. However, measuring in the *z*-direction will clearly distinguish whether the system was prepared in state $|u\rangle$ vs $|d\rangle$.

Principle 3. Distinguishable states are represented by orthogonal vectors.

The next principle formalizes equations 1.01 and 1.02. It also generalizes 2.2.

Definitions. Let $\{|i\rangle\}$ be a basis for our system and suppose the system is prepared in state $|A\rangle = \sum_i \alpha_i |i\rangle$. $|A\rangle$ is called the **state vector of the system**, or just the **system state** for short**.** Let L be an observable and λ a possible outcome. We define $\mathbf{P}_L(\boldsymbol{\lambda})$ to be **the probability that the outcome of the measurement is** $\boldsymbol{\lambda}$. When L is understood, we shorten this to just $\mathbf{P}(\boldsymbol{\lambda})$.

Principle 4. Let $|A\rangle$ be a normalized state-vector of a system and L an observable with eigenvalues $\{\lambda_i\}$ and corresponding eigenvectors $\{|\lambda_i\rangle\}$. If L is measured, then

$$\mathrm{P}(\lambda_i) = \langle A|\lambda_i\rangle\langle\lambda_i|A\rangle \tag{3.11}$$

Principle 4 says that if a system is prepared in state $|A\rangle$ and then is measured, the probability the system ends up in state $|\lambda_i\rangle$ is the square of the magnitude of the inner product of state $|\lambda_i\rangle$ with the initial state $|A\rangle$.

Theorem. Let σ_x, σ_y, and σ_z be the Hermitian operators that represent measurements along the *x*-, *y*-, and *z*-axes, respectively. Then

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \text{and} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.20}$$

Proof. From (2.11) and (2.12), $|u\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|d\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. By Principle 2, they are eigenvectors of σ_z with respective eigenvalues +1 and -1. Letting

$\sigma_z = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and solving $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

yields the result for σ_z . By Principle 2 the eigenvectors of σ_x are $|r\rangle \overset{(2.13)}{=} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

and $|\ell\rangle \overset{(2.13)}{=} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ with respective eigenvalues +1 and -1. We find σ_x by solving

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = -\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$ and similarly for σ_y

using $|i\rangle$ and $|o\rangle$. ■

Definition. σ_x, σ_y, and σ_z are called the **Pauli matrices**. They are called **components of the 3-vector operator** $\boldsymbol{\sigma} \equiv \begin{bmatrix} \boldsymbol{\sigma_x} \\ \boldsymbol{\sigma_y} \\ \boldsymbol{\sigma_z} \end{bmatrix}$. The 3-vector operator σ is a vector whose elements are matrices.

Common Misconception. Acting on a state with an operator *L* does not give the same answer as measuring the observable *L*. The former gives probability of an outcome, the latter gives a definite outcome.

Definition. Let $\hat{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}$ be a direction in 3-space and let σ_n be the spin state for the direction $\hat{n}$. We define the **spin operator in the direction** $\boldsymbol{\hat{n}}$ as

$$\sigma_n \equiv \sigma \bullet \vec{n} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z \tag{3.22}$$

Theorem.

$$\sigma_n = \begin{bmatrix} n_z & n_x - n_y i \\ n_x + n_y i & -n_z \end{bmatrix} \tag{3.23}$$

Proof. $\sigma_n \overset{(3.22)}{=} \sigma \bullet \hat{n} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z$

$$\overset{(3.20)}{=} n_x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + n_y \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + n_z \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} n_z & n_x - n_y i \\ n_x + n_y i & -n_z \end{bmatrix} \quad \blacksquare$$

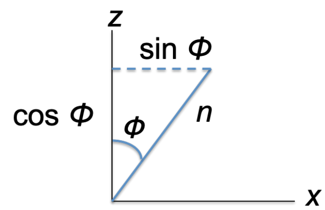
Exercise 3.4. Represent $\hat{n}$ in spherical coordinates: $\hat{n}=\begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}=\begin{bmatrix} \sin\phi\cos\theta \\ \sin\phi\sin\theta \\ \cos\phi \end{bmatrix}$.

Then the eigenvalues of σ_n are λ = ±1 and the eigenvectors are

$$\left|\lambda_1\right\rangle=\begin{bmatrix} \cos\dfrac{\phi}{2} \\ \sin\dfrac{\phi}{2}\,e^{i\theta} \end{bmatrix} \text{ and } \left|\lambda_2\right\rangle=\begin{bmatrix} \sin\dfrac{\phi}{2} \\ -\cos\dfrac{\phi}{2}\,e^{i\theta} \end{bmatrix}.$$

Definition. The **expected value of an operator *M*** is written $\langle \boldsymbol{M} \rangle$.

Example. If direction $\hat{n}$ lies in the *xz*-plane, find the eigenvectors and eigenvalues of the spin operator σ_n and use them to find the expected value $\langle \sigma_n \rangle$.



$\hat{n}=\begin{bmatrix} \sin\phi \\ 0 \\ \cos\phi \end{bmatrix}$. By (3.23), $\sigma_n=\begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix}$. Solving $\det(\sigma_n-\lambda I)=0$ yields eigenvalues $\lambda_1=1$ and $\lambda_2=-1$ for σ_n. The corresponding eigenvectors are

$$\left|\lambda_1\right\rangle=\begin{bmatrix} \cos\dfrac{\phi}{2} \\ \sin\dfrac{\phi}{2} \end{bmatrix} \text{ and } \left|\lambda_2\right\rangle=\begin{bmatrix} -\sin\dfrac{\phi}{2} \\ \cos\dfrac{\phi}{2} \end{bmatrix}, \text{ and so}$$

$$\mathrm{P}(\sigma_n=+1)=\mathrm{P}(\lambda_1)\overset{(3.11)}{=}\left|\left\langle u|\lambda_1\right\rangle\right|^2=\left|\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix} \cos\dfrac{\phi}{2} \\ \sin\dfrac{\phi}{2} \end{bmatrix}\right|^2=\cos^2\frac{\phi}{2} \text{ and } \langle\sigma_n\rangle\overset{(1.02)}{=}\cos\phi.$$

Theorem (**Spin Polarization Principle**). Given a spin state $\left|A\right\rangle=\alpha_u\left|u\right\rangle+\alpha_d\left|d\right\rangle$ there exists a direction $\hat{n}$ such that $(\vec{\sigma}\bullet\vec{n})\left|A\right\rangle=\left|A\right\rangle$.

Proof. WLOG $\alpha_u \in \mathbb{R}$:

Say $\alpha_u = r e^{i\omega}$. Set $|B\rangle = e^{-i\omega}|A\rangle = \beta_u|u\rangle + \beta_d|d\rangle$. Then $|B\rangle$ is equivalent to $|A\rangle$ and $\beta_u = e^{-i\omega}\alpha_u = r \in \mathbb{R}$. Moreover $\langle B|B\rangle = e^{i\omega}e^{-i\omega}\langle A|A\rangle = 1$, so

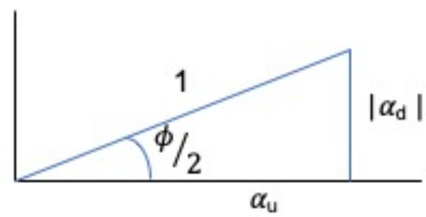
$$\beta_u{}^2 + |\beta_d|^2 = |\beta_u|^2 + |\beta_d|^2 = 1. \quad ✔ \tag{a}$$

So,

$$\alpha_u = r \in \mathbb{R}, \tag{b}$$

and

$$\alpha_d = s e^{i\theta} \in \mathbb{C}. \tag{c}$$



Since $\alpha_u{}^2 + |\alpha_d|^2 \overset{(a)}{=} 1$, there is an angle ϕ (see figure) such that

$$\alpha_u = \cos\frac{\phi}{2} \tag{d}$$

and

$$|\alpha_d| = \sin\frac{\phi}{2}. \tag{e}$$

Since $|\alpha_d| \overset{(c)}{=} s \overset{(c)}{=} e^{-i\theta}\alpha_d$ then

$$\alpha_d = |\alpha_d| e^{i\theta} \overset{(e)}{=} \sin\frac{\phi}{2} e^{i\theta}. \tag{f}$$

Define direction $\hat{n}$ in terms of θ and ϕ using spherical coordinates:

$$\hat{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \equiv \begin{bmatrix} \sin\phi\cos\theta \\ \sin\phi\sin\theta \\ \cos\phi \end{bmatrix}.$$

From Exercise 3.4 (p. 19), $|A\rangle = \begin{bmatrix} \alpha_u \\ \alpha_d \end{bmatrix} \overset{(d,\,f)}{=} \begin{bmatrix} \cos\frac{\phi}{2} \\ \sin\frac{\phi}{2} e^{i\theta} \end{bmatrix}$ is an eigenvector of σ_n having an eigenvalue of +1. That is, $\sigma_n|A\rangle = |A\rangle$ which finishes the proof since by definition (3.22) we have that $\sigma_n = \sigma \bullet \vec{n}$. ■

Corollary 1. For any spin state $|A\rangle$ there is a direction for which the spin is +1, and $|A\rangle$ is an eigenvector of $\vec{\sigma}_n$ with eigenvalue +1.

Corollary 2. Given a spin state $|A\rangle$, the expectation $\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = 1$.

Proof. From the theorem, $\vec{\sigma} \bullet \vec{n}|A\rangle = |A\rangle$ for some direction $\hat{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}$.

Spin is prepared in direction $\hat{n}$. For σ_x it is measured in the *x*-direction. So,

$$\langle\sigma_x\rangle \overset{(1.01)}{=} \vec{n} \bullet \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = n_x .$$

Similarly,

$$\langle\sigma_y\rangle = \vec{n} \bullet \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = n_y, \text{ and } \langle\sigma_z\rangle = \hat{n} \bullet \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = n_z .$$

Therefore

$$\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = n_x^2 + n_y^2 + n_z^2 = 1. \quad \blacksquare$$

Chapter 4. Time and Change

Definition. $\left|\boldsymbol{\Psi}(\boldsymbol{t})\right\rangle$ denotes the **state of a closed system at time *t*.** Let ***U*** be **the time-development operator** for the system, meaning

$$\left|\boldsymbol{\Psi}(\boldsymbol{t})\right\rangle = \boldsymbol{U}(\boldsymbol{t})\left|\boldsymbol{\Psi}(\boldsymbol{0})\right\rangle \tag{4.1}$$

Equation 4.1 says that a *state* $\left|\Psi(t)\right\rangle$ is computed deterministically from the initial state $\left|\Psi(0)\right\rangle$. Nonetheless, subsequent *measurements* are probabilistic. In conventional mechanics, states and measurements are one and the same, but not in quantum mechanics.

Theorem. $U^\dagger U = I$ (4.5)

Proof. Fix t and shorten $U(t)$ to U. Let $\left\{\left| i\right\rangle\right\}$ be an orthonormal basis for the system. Define $\left| i(t)\right\rangle = \left| i(0)\right\rangle = \left| i\right\rangle \forall t$. Letting $\left|\Psi(t)\right\rangle = \left| i(t)\right\rangle = \left| i\right\rangle$ in 4.1 yields $\left|i\right\rangle = U\left|i\right\rangle$ or

$$\left\langle i\right| = \left\langle i\right|U^\dagger . \tag{a}$$

Letting $\left|\Psi(t)\right\rangle = \left|j\right\rangle$ in 4.1 yields

$$\left|j\right\rangle = U\left|j\right\rangle . \tag{b}$$

So,

$$\left\langle i\right|U^\dagger U\left|j\right\rangle = \left\langle i|j\right\rangle : \tag{c}$$

$$\left\langle i\right|U^\dagger U\left|j\right\rangle \overset{(3.04)}{=} \left\{\left\langle i\right|U^\dagger\right\}\left\{U\left|j\right\rangle\right\} \overset{(a,\,b)}{=} \left\{\left\langle i\right|\right\}\left\{\left|j\right\rangle\right\} \overset{(1.14)}{=} \left\langle i|j\right\rangle$$

Equation (c) shows that $U^\dagger U$ behaves like I on all basis pairs.

Let $\left\langle A\right| = \sum_{i=1}^{n}\left\langle i\right|\alpha_i^*$ and $\left|B\right\rangle = \sum_{j=1}^{n}\beta_j\left|j\right\rangle$ be any two vectors.

$$\left\langle A\right|U^\dagger U\left|B\right\rangle = \left\langle A|B\right\rangle : \tag{d}$$

$$\left\langle A\right|U^\dagger U\left|B\right\rangle = \left[\sum_{i=1}^{n}\left\langle i\right|\alpha_i^*\right]U^\dagger U\left\{\sum_{j=1}^{n}\beta_j\left|j\right\rangle\right\}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i^*\beta_j\left\langle i\right|U^\dagger U\left|j\right\rangle \overset{(c)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i^*\beta_j\left\langle i|j\right\rangle \overset{(1.15)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i^*\beta_j\delta_{ij}$$

$$= \sum_{i=1}^{n}\alpha_i^*\beta_i \overset{(1.2)}{=} \left\langle A|B\right\rangle$$

Equation (d) shows that $U^\dagger U$ behaves like I on all vector pairs. To show that $U^\dagger U = I$, we need to show that $U^\dagger U\left|A\right\rangle = \left|A\right\rangle$ and $\left\langle A\right|U^\dagger U = \left\langle A\right|$. Suppose

$$U^\dagger U\left|A\right\rangle = \left|B\right\rangle. \tag{e}$$

Then

$$\alpha_i = \left(0\cdots 1_i \cdots 0\right)\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = \left\langle i\middle|A\right\rangle \overset{(d)}{=} \left\langle i\right|U^\dagger U\left|A\right\rangle \overset{(e)}{=} \left\langle i\middle|B\right\rangle = \beta_i \;\; \forall i.$$

That is, $\left|B\right\rangle = \left|A\right\rangle$ or, from (e), $U^\dagger U\left|A\right\rangle = \left|A\right\rangle$. ✔

Similarly $\left\langle A\right|U^\dagger U = \left\langle A\right|$. ■

Theorem. $U^\dagger U = I$ iff $U^\dagger = U^{-1}$.

Proof. If $U^\dagger = U^{-1}$ then $U^\dagger U = U^{-1}U = I$. Conversely, suppose $U^\dagger U = I$. Then $1 = \text{Det } I = \text{Det } U^\dagger U = \text{Det } U^\dagger \text{ Det } U \;\; \Rightarrow \;\; \text{Det } U \neq 0$. Thus, U^{-1} exists and $U^\dagger U = I = U^{-1}U$. Multiplying both sides on the right by U^{-1} yields $U^\dagger = U^{-1}$. ■

Definition. An operator U that satisfies $\boldsymbol{U^\dagger U = I}$ is called **unitary**.

Note. *Hermitian* means $U^\dagger = U$. *Unitary* means .$U^\dagger = U^{-1}$.

Convention. Henceforth the term “basis” will refer to an orthonormal basis.

Recall from Principle 3 that $\left|\Psi\right\rangle$ and $\left|\Phi\right\rangle$ are distinguishable iff $\left\langle \Psi\middle|\Phi\right\rangle = 0$.

Theorem: **Conservation of Distinctions Law.** If basis states $\left|\Psi(0)\right\rangle$ and $\left|\Phi(0)\right\rangle$ are distinguishable then so are $\left|\Psi(t)\right\rangle$ and $\left|\Phi(t)\right\rangle$:

$$\left\langle \Psi(t)\middle|\Phi(t)\right\rangle = 0 \;\; \forall t \in \mathbb{R} \tag{4.2}$$

Proof. $\left\langle \Psi(t)\middle|\Phi(t)\right\rangle \overset{(4.1)}{=} \left\langle \Psi(0)\right|U^\dagger U\left|\Phi(0)\right\rangle \overset{(4.5)}{=} \left\langle \Psi(0)\middle|\Phi(0)\right\rangle$. ■

Principle 5. The evolution of state vectors with time is unitary.

Principle 5 ensures that if two states were orthogonal before time *t*, they remain orthogonal. For example if $U(t)\left|r\right\rangle \perp U(t)\left|\ell\right\rangle$ at time *t* then $U(t+)\left|r\right\rangle \perp U(t+)\left|\ell\right\rangle$

just after time t. U is unitary although perhaps not continuous since Alice could make a sudden change to her system at time t such as making a measurement or applying a magnetic field. U is *mostly* continuous but the book does not pursue this topic.

Definition. For $t \neq 0$ define

$$H_t = \frac{1}{t\,i}\left[\, I - U(t) \,\right] \qquad (4.01)$$

where I is the identity matrix and $U(t)$ is the time-development operator (4.1).

Since $t \neq 0$, equation (4.01) is equivalent to

$$U(t) = I - t\,i\,H_t = U(0) - t\,i\,H_t\,. \qquad 4.02)$$

Observe that $-t\,i\,H_t$ represents the operator for the change that occurs between time 0 and time t. If U is continuous at $t = 0$ then the change is small for small t.

We define the **Hamiltonian** as

$$\boldsymbol{H = \lim_{t \to 0} H_t} \text{ when this limit exists.} \qquad (4.03)$$

This limit might not exist. For example, if $U(t) \to I$ then from (4.01) we get $H_t \to \frac{0}{0}$ where the numerator of the RHS is the zero operator.

Considering H as a limit of matrices, if H exists then clearly so does $H^\dagger$, and

$$H^\dagger = \lim_{t \to 0} H_t^\dagger \qquad (4.04)$$

Moreover, $H^\dagger H$ must also exist, and since each element of the matrix $H^\dagger H$ is a (finite) complex number, then

$$t\,i\,H_t^\dagger H_t \to t\,i\,H^\dagger H \to 0 \;\text{ as }\; t \to 0\,, \qquad (4.05),$$

a fact we will need in the next theorem.

There are some subtle issues lurking here. For U to be continuous, what exactly does it mean to say that H_t represents "a small change"? H_t is an operator, which is a linear transformation and can also be represented as a matrix. $-i\,t\,H_t$ represents a small change if $-i\,t\,H_t\left|A\right\rangle \approx \left|0\right\rangle$ for all vectors $\left|A\right\rangle$.

Also, we should not confuse the notion of continuity of U with respect to t with the concept of continuity of an individual $U(t)$ in terms of the vectors $\left|\Psi\right\rangle$ that it operates on. For the latter concept, it can be shown that a linear operator such as $U(t)$ is continuous if and only if it is bounded, meaning $U(t)$ maps the unit vectors into a bounded sphere.

Theorem. If H exists, then H is a Hermitian operator and

$$\left.\frac{\partial\left|\Psi(t)\right\rangle}{\partial t}\right|_{t=0}=-iH\left|\Psi(0)\right\rangle \tag{4.06}$$

Proof.

$$U(t)\overset{(4.02)}{=}I-tiH_t$$

$$\Rightarrow \quad U^{\dagger}(t)=U^{\mathsf{T}^*}(t)=I+tiH_t^{\dagger}$$

$$\Rightarrow \quad I\overset{(4.5)}{=}U^{\dagger}(t)U(t)=\left(I+tiH_t^{\dagger}\right)\left(I-tiH_t\right)=I+ti\left(H_t^{\dagger}-H_t\right)+t^2H_t^{\dagger}H_t$$

$$\Rightarrow \quad H_t=H_t^{\dagger}-tiH_t^{\dagger}H_t \tag{4.07}$$

$$\Rightarrow \quad H\overset{(4.03)}{=}\lim_{t\to0}H_t\overset{(4.07)}{=}\lim_{t\to0}H_t^{\dagger}-i\lim_{t\to0}tH_t^{\dagger}H_t\overset{(4.05)}{=}\lim_{t\to0}H_t^{\dagger}\overset{(4.04)}{=}H^{\dagger}$$

$$\Rightarrow \quad H \text{ is Hermitian} \quad ✔$$

$$\left|\Psi(t)\right\rangle\overset{(4.1)}{=}U(t)\left|\Psi(0)\right\rangle\overset{(4.02)}{=}\left(I-tiH_t\right)\left|\Psi(0)\right\rangle=\left|\Psi(0)\right\rangle-tiH_t\left|\Psi(0)\right\rangle$$

$$\Rightarrow \quad \frac{\left|\Psi(t)\right\rangle-\left|\Psi(0)\right\rangle}{t}=-iH_t\left|\Psi(0)\right\rangle$$

$$\Rightarrow \left.\frac{\partial\left|\Psi(t)\right\rangle}{\partial t}\right|_{t=0}=\lim_{t\to0}\frac{\left|\Psi(t)\right\rangle-\left|\Psi(0)\right\rangle}{t}=-i\lim_{t\to0}H_t\left|\Psi(0)\right\rangle\overset{(4.03)}{=}-iH\left|\Psi(0)\right\rangle \quad \blacksquare$$

We could have performed this process starting at any point t_0. Thus we can recast (4.06) as

$$\left.\frac{\partial\left|\Psi(t)\right\rangle}{\partial t}\right|_{t=t_0}=-iH\left|\Psi(t_0)\right\rangle. \tag{4.9}$$

Note 1. Since H is Hermitian, it is an observable and by the Fundamental Theorem it has a complete set of eigenvectors and eigenvalues.

Note 2. The classical mechanics Hamiltonian has units of energy: $\text{kg-m}^2/\text{sec}^2$. LHS of 4.9 has units of / sec. Planck's constant is

$$\hbar=\frac{h}{2\pi}\approx 1.05\text{x}10^{-34}\ \text{kg-m}^2/\text{sec},$$

just what we need to change units. So, we re-write 4.9 as

$$\hbar\left|\dot{\Psi}(t)\right\rangle = \hbar\frac{\partial\left|\Psi(t)\right\rangle}{\partial t} = -iH\left|\Psi(t)\right\rangle \tag{4.10}$$

or, equivalently,

$$\hbar\left\langle\dot{\Psi}(t)\right| = \hbar\frac{\partial\left\langle\Psi(t)\right|}{\partial t} = i\left\langle\Psi(t)\right|H \tag{4.10b}$$

Note 3. While this proof does not generate a formula for H, for many systems (e.g., for a magnetic field) the formula for energy, H, is already known. It is also possible that the limit of H_t does not exist. The study of conditions for which the limit exists is beyond the scope of this book.

Definition. Equation 4.10 is called the **time-dependent Schrödinger equation.** H is the **quantum Hamiltonian**. Because H represents energy, the observable values of energy are just the eigenvalues of H.

Theorem. Let L be an observable prepared in state $\left|A\right\rangle$ and having eigenvectors $\left\{\left|\lambda_i\right\rangle\right\}$ and eigenvalues $\left\{\lambda_i\right\}$. The **expected value of the observable *L*** is

$$\left\langle L\right\rangle = \sum_i \lambda_i \mathrm{P}(\lambda_i) \tag{4.11}$$

Proof. The outcomes of L in state $\left|A\right\rangle$ are its eigenvalues $\{\lambda_i\}$ with probabilities of occurrence $\{\mathrm{P}(\lambda_i)\}$. Thus (4.11) is simply the standard expected-value formula. ■

Caution. Recall that in equation (3.11) we shortened $\mathrm{P}_L(\lambda_i)$ to $\mathrm{P}(\lambda_i)$. Now, $\left\langle L^2\right\rangle = \sum_i \lambda_i^2 \mathrm{P}_{L^2}(\lambda_i^2)$. Also, L^2 has outcome λ_i^2 iff L has outcome λ_i. That is, $\mathrm{P}_{L^2}(\lambda_i^2) = \mathrm{P}_L(\lambda_i)$. So, using the shorthand notation $\mathrm{P} = \mathrm{P}_L$,

$$\left\langle L^2\right\rangle = \sum_i \lambda_i^2 \mathrm{P}(\lambda_i) \tag{4.11b}$$

not $\left\langle L^2\right\rangle = \sum_i \lambda_i^2 \mathrm{P}(\lambda_i^2)$.

Theorem. Let L be an observable with eigenvectors $\left\{\left|\lambda_i\right\rangle\right\}$, $\left|A\right\rangle = \sum_i \alpha_i\left|\lambda_i\right\rangle$ a normalized state vector, θ a phase angle, and $\left|B\right\rangle = e^{i\theta}\left|A\right\rangle$. Then

$$\left\langle L\right\rangle = \left\langle A\middle|L\middle|A\right\rangle \tag{4.13}$$

$$\left\langle L\right\rangle = \sum_i \alpha_i^* \alpha_i \lambda_i \tag{4.14}$$

$$\mathrm{P}(\lambda_i) = \langle B|\lambda_i\rangle\langle\lambda_i|B\rangle \qquad \text{(a)}$$

$$\langle L\rangle = \langle B|L|B\rangle \qquad \text{(b)}$$

Proof.

$$\langle\lambda_i|A\rangle \overset{(1.5)}{=} \alpha_i \qquad \text{(i)}$$

$$\langle A|\lambda_i\rangle \overset{(1.22)}{=} \langle\lambda_i|A\rangle^* \overset{(i)}{=} \alpha_i^* \qquad \text{(ii)}$$

$$L|A\rangle = \sum_i \alpha_i L|\lambda_i\rangle = \sum_i \alpha_i \lambda_i|\lambda_i\rangle \qquad \text{(iii)}$$

By Principle 4 (i.e., 3.11),

$$\langle A|L|A\rangle \overset{(iii)}{=} \{\langle A|\}\left\{\sum_i \alpha_i \lambda_i|\lambda_i\rangle\right\} = \sum_i \langle A|\lambda_i\rangle \alpha_i \lambda_i \overset{(ii)}{=} \sum_i \alpha_i^* \alpha_i \lambda_i \quad ✔$$

$$\overset{(ii,\, i)}{=} \sum_i \langle A|\lambda_i\rangle\langle\lambda_i|A\rangle \lambda_i \overset{(3.11)}{=} \sum_i \mathrm{P}(\lambda_i)\lambda_i \overset{(4.11)}{=} \langle L\rangle \quad ✔$$

$|A\rangle$ and $|B\rangle$ represent the same state since $e^{-i\theta}$ is a scalar, and this fact is corroborated by (a) and (b):

$$\langle B|\lambda_i\rangle = \langle A|e^{-i\theta}|\lambda_i\rangle = e^{-i\theta}\langle A|\lambda_i\rangle \overset{(ii)}{=} e^{-i\theta}\alpha_i^*$$

$$\Rightarrow \quad \langle\lambda_i|B\rangle = e^{i\theta}\alpha_i$$

$$\Rightarrow \quad \mathrm{P}(\lambda_i) \overset{(3.11)}{=} \alpha_i^* \alpha_i = \langle B|\lambda_i\rangle\langle\lambda_i|B\rangle \quad ✔$$

$$|A\rangle = e^{-i\theta}|B\rangle \quad \text{and} \quad \langle A| \overset{(1.13)}{=} \langle B|e^{i\theta}$$

$$\Rightarrow \quad \langle L\rangle \overset{(4.13)}{=} \langle A|L|A\rangle = \left(\langle B|e^{i\theta}\right)L\left(e^{-i\theta}|B\rangle\right) = \langle B|L|B\rangle \quad ✔ \quad ■$$

Definition. Given two operators L and M, the **commutator** is $[\boldsymbol{L}, \boldsymbol{M}] = \boldsymbol{LM} - \boldsymbol{ML}$.

In the next theorem the observable L does not change over time, yet $\langle L\rangle$ is time-dependent. How can that be? It is because the state vector $\langle\Psi|$ changes over time, hence the outcome probabilities due to L also change. For example, let L be the observable of spin along the positive z-axis. $\langle L\rangle$ would be +1 when $\langle\Psi(t)| = |u\rangle$ but it would be ½ when $|\Psi(t)\rangle = |r\rangle$. So $\langle L\rangle = \langle\Psi(t)|L|\Psi(t)\rangle$ is time-dependent.

Theorem. Let L be an observable, $\left|\Psi(t)\right\rangle$ a normalized time-dependent state vector for L, and H a Hermitian operator that satisfies the time-dependent Schrödinger equation (4.10). Then

$$\frac{d}{dt}\left\langle L\right\rangle = \frac{i}{\hbar}\left\langle \left[H, L\right] \right\rangle \tag{4.17}$$

or, equivalently

$$\frac{d}{dt}\left\langle L\right\rangle = -\frac{i}{\hbar}\left\langle \left[L,H\right] \right\rangle \tag{4.18}$$

Proof. Let

$$\left|\Psi(t)\right\rangle = \sum_i \alpha_i(t)\left|\lambda_i\right\rangle. \text{ Thus, } \left\langle\Psi(t)\right| = \sum_i \left\langle\lambda_i\right|\alpha_i^*(t). \tag{a}$$

So

$$\left|\dot\Psi(t)\right\rangle = \sum_i \dot\alpha_i(t)\left|\lambda_i\right\rangle \quad \text{and} \quad \left\langle\dot\Psi(t)\right| = \sum_i \left\langle\lambda_i\right|\dot\alpha_i^*(t). \tag{b}$$

Also

$$\left\langle\Psi(t)\right|L\left|\Psi(t)\right\rangle \overset{(4.13,\ 4.14)}{=} \sum_i \alpha_i^*(t)\alpha_i(t)\lambda_i\,. \tag{c}$$

Claim $\left\langle\dot\Psi(t)\right|L\left|\Psi(t)\right\rangle = \sum_i \dot\alpha_i^*(t)\alpha_i(t)\lambda_i$ and $\left\langle\Psi(t)\right|L\left|\dot\Psi(t)\right\rangle = \sum_i \alpha_i^*(t)\dot\alpha_i(t)\lambda_i$: (d)

$$L\left|\Psi(t)\right\rangle \overset{(a)}{=} \sum_i \alpha_i(t) L\left|\lambda_i\right\rangle = \sum_i \alpha_i(t)\lambda_i\left|\lambda_i\right\rangle$$

$$\left\langle\dot\Psi(t)\right|L\left|\Psi(t)\right\rangle \overset{(a,b)}{=} \sum_j \left\langle\lambda_j\right|\dot\alpha_j^*(t)\ L\ \sum_i \alpha_i(t)\left|\lambda_i\right\rangle = \sum_i\sum_j \alpha_i(t)\dot\alpha_j^*(t)\left\langle\lambda_j\right|L\left|\lambda_i\right\rangle$$

$$\overset{\text{(eigenvector)}}{=} \sum_i\sum_j \alpha_i(t)\dot\alpha_j^*(t)\left\langle\lambda_j\right|\lambda_i\left|\lambda_i\right\rangle \overset{(1.17)}{=} \sum_i\sum_j \alpha_i(t)\dot\alpha_j^*(t)\,\lambda_i\left\langle\lambda_j\middle|\lambda_i\right\rangle$$

$$\overset{(1.15)}{=} \sum_i\sum_j \alpha_i(t)\dot\alpha_j^*(t)\,\lambda_i\,\delta_{ij} = \sum_i \dot\alpha_i^*(t)\alpha_i(t)\lambda_i$$

The 2nd equation is proven similarly.

Therefore

$$\frac{d}{dt}\left\langle L\right\rangle \overset{(4.13)}{=} \frac{d}{dt}\left\langle\Psi(t)\right|L\left|\Psi(t)\right\rangle \overset{(c)}{=} \frac{d}{dt}\sum_i \alpha_i^*(t)\alpha_i(t)\lambda_i$$

$$= \sum_i \dot\alpha_i^*(t)\alpha_i(t)\lambda_i + \sum_i \alpha_i^*(t)\dot\alpha_i(t)\lambda_i$$

$$\overset{(d)}{=} \left\langle\dot\Psi(t)\right|L\left|\Psi(t)\right\rangle + \left\langle\Psi(t)\right|L\left|\dot\Psi(t)\right\rangle$$

$$\overset{(4.10)}{=} \frac{i}{\hbar}\langle\Psi(t)|HL|\Psi(t)\rangle - \frac{i}{\hbar}\langle\Psi(t)|LH|\Psi(t)\rangle$$
$$= \frac{i}{\hbar}\langle\Psi(t)|\,[H,L]\,|\Psi(t)\rangle$$
$$\overset{(4.13,\,4.14)}{=} \frac{i}{\hbar}\langle\,[H,L]\,\rangle \qquad \blacksquare$$

Exercise 4.2. If M and L are Hermitian, then $i\,[M, L]$ is Hermitian but not necessarily $[M, L]$.

Proof. Given

$$M^\dagger = M \text{ and } L^\dagger = L\,, \tag{i}$$

we wish to show that $\{i[M,L]\}^\dagger = i[M,L]$.

Note that

$$(A-B)^\dagger = (A-B)^{T^*} = A^{T^*} - B^{T^*} = A^\dagger - B^\dagger \tag{ii}$$

and

$$(AB)^\dagger = (AB)^{T^*} = (B^T A^T)^* = B^\dagger A^\dagger\,. \tag{iii}$$

Let $M = (m_{ij})$. Then $iM = (im_{ij})$ and

$$(iM)^\dagger = (im_{ji})^* = (i^* m_{ji}{}^*) = (-im_{ji}{}^*) = -i(m_{ji}{}^*) = -iM^\dagger\,. \tag{iv}$$

So

$$\{\,i[M,L]\,\}^\dagger = \{\,i(ML-LM)\,\}^\dagger = \{(iM)L - L(iM)\}^\dagger \overset{(\text{ii, iii})}{=} L^\dagger(iM)^\dagger - (iM)^\dagger L^\dagger$$
$$\overset{(\text{iv})}{=} -iL^\dagger M^\dagger + iM^\dagger L^\dagger = i(M^\dagger L^\dagger - L^\dagger M^\dagger) = i[M^\dagger, L^\dagger] \overset{(\text{i})}{=} i[M,L] \qquad ✔$$

However,

$$[M,L]^\dagger = (ML-LM)^\dagger \overset{(\text{ii, iii})}{=} L^\dagger M^\dagger - M^\dagger L^\dagger \overset{(\text{i})}{=} LM - ML = [L,M] = -[M,L],$$

not equal to $[M,L]$ unless M and L commute; i.e., unless $[M,L]=0$. $\blacksquare$

Corollary. The time derivative of a Hermitian operator is Hermitian.

Proof. $\frac{d}{dt}\langle L\rangle \overset{(4.18)}{=} -\frac{i}{\hbar}\langle\,[H,L]\,\rangle \qquad \blacksquare$

Definition. An **observable *L* is conserved** if it does not change over time. This occurs if $P(\lambda)$ is constant over time for all eigenvalues λ. This is equivalent to all the moments being constant over time: i.e., the mean of λ, 2nd moment, 3rd moment, etc. This causes every polynomial in λ to be constant over time, which means that every continuous function of λ is constant over time.

The next theorem is an immediate consequence of equation (4.18).

Theorem. An observable L is conserved iff $[L,H] = 0$; i.e., L commutes with the Hamiltonian.

Corollary. The Hamiltonian H is conserved.

Definition. H is the **energy of a system**.

We have thus proved that energy is conserved.

Definition. A **magnetic field** is a 3-vector $\vec{B} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix}$ where B_x, B_y, and B_z are the magnetic field components. The **energy of a rotor** (spinning charged object) **in a magnetic field** is $H \propto \vec{\sigma} \bullet \vec{B} = \sigma_x B_x + \sigma_y B_y + \sigma_z B_z$.

Magnetic Field Example. $\vec{B} = \begin{bmatrix} 0 \\ 0 \\ B_z \end{bmatrix}$ lies along the z-axis. Find the expected value of the 3-vector operator σ over time.

Solution. $H \propto \sigma_z$. Define ω such that

$$H = \frac{\hbar\omega}{2}\sigma_z \tag{a}$$

We wish to find $\langle \dot{\sigma}(t) \rangle$; i.e., to find $\langle \dot{\sigma}_x(t) \rangle$, $\langle \dot{\sigma}_y(t) \rangle$, and $\langle \dot{\sigma}_z(t) \rangle$. Since σ_x, σ_y, and σ_z are Hermitian operators (i.e., observables),

$$\begin{aligned} \langle \dot{\sigma}_x \rangle &\overset{(4.18)}{=} -\frac{i}{\hbar}\langle [\sigma_x, H] \rangle, \\ \langle \dot{\sigma}_y \rangle &\overset{(4.18)}{=} -\frac{i}{\hbar}\langle [\sigma_y, H] \rangle, \end{aligned} \tag{4.24}$$

$$\left\langle \dot{\sigma}_z \right\rangle \overset{(4.18)}{=} -\frac{i}{\hbar}\left\langle \left[\sigma_z, H\right] \right\rangle.$$

Thus,

$$\begin{aligned}
\left\langle \dot{\sigma}_x \right\rangle &\overset{(a)}{=} -\frac{i\omega}{2}\left\langle \left[\sigma_x, \sigma_z\right] \right\rangle, \\
\left\langle \dot{\sigma}_y \right\rangle &\overset{(a)}{=} -\frac{i\omega}{2}\left\langle \left[\sigma_y, \sigma_z\right] \right\rangle, \\
\left\langle \dot{\sigma}_z \right\rangle &\overset{(a)}{=} -\frac{i\omega}{2}\left\langle \left[\sigma_z, \sigma_z\right] \right\rangle = 0.
\end{aligned} \tag{4.25}$$

Recall

$$\sigma_x \overset{(3.20)}{=} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \overset{(3.20)}{=} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \overset{(3.20)}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

So,

$$\left[\sigma_x, \sigma_z\right] = -2i\sigma_y : \tag{4.26}$$

$$\begin{aligned}
\left[\sigma_x, \sigma_z\right] &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} = -2i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
&= -2i\sigma_y
\end{aligned}$$

and

$$\left[\sigma_y, \sigma_z\right] = 2i\sigma_x : \tag{4.26b}$$

$$\begin{aligned}
\left[\sigma_y, \sigma_z\right] &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2i \\ 2i & 0 \end{pmatrix} = 2i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
&= 2i\sigma_x.
\end{aligned}$$

(Also, $\left[\sigma_x, \sigma_y\right] = 2i\sigma_z$ but we do not need this.)

Therefore

$$\left\langle \dot{\sigma}_x \right\rangle \overset{(4.25,\,4.26)}{=} -\frac{i\omega}{2}\left\langle -2i\sigma_y \right\rangle = -\omega\left\langle \sigma_y \right\rangle \quad (\text{i.e., } x \to -y)$$

$$\left\langle \dot{\sigma}_y \right\rangle \overset{(4.25,\,4.26b)}{=} -\frac{i\omega}{2}\left\langle 2i\sigma_x \right\rangle = \omega\left\langle \sigma_x \right\rangle \quad (\text{i.e., } y \to x) \tag{4.27}$$

$$\left\langle \dot{\sigma}_z \right\rangle = 0.$$

Thus, the 3-vector operator $\vec{\sigma}$ precesses clockwise like a gyroscope around the direction of the magnetic field.

Note. The actual values for $\left\langle \sigma_x \right\rangle$ and $\left\langle \sigma_y \right\rangle$ depend on the state vector of the system; i.e., how the system is prepared. They also depend on solving the differential equations (4.27). The values are calculated in Exercise 4.6.

Definition. Let L be an observable, $\left|\Psi(t)\right\rangle$ a normalized time-dependent state vector for L, and $\left\{\left|E_j\right\rangle\right\}$ the eigenvectors (i.e., the energy states) of the Hamiltonian H (as defined by Eq. 4.10). Then

$$H\left|E_j\right\rangle = E_j\left|E_j\right\rangle \tag{4.28}$$

is the **time-independent Schrödinger equations** and E_j and $\left|E_j\right\rangle$ are the **energy eigenvalues and eigenvectors**, respectively, of the system.

Example. Find the eigenvalues and eigenvectors of H in the prior example.

Solution. $H \overset{(a)}{=} \frac{\hbar\omega}{2}\sigma_z$. The eigenvalues for σ_z are $\lambda_1 = +1$ and $\lambda_2 = -1$. Since $H\left|\lambda_i\right\rangle \overset{(a)}{=} \left(\frac{\hbar\omega}{2}\right)\sigma_z\left|\lambda_i\right\rangle = \left(\frac{\hbar\omega}{2}\right)\lambda_i\left|\lambda_i\right\rangle$, we see that the eigenvalues and eigenvectors for H are $E_1 = \frac{\hbar\omega}{2}\lambda_1 = \frac{\hbar\omega}{2}$, $E_2 = \frac{\hbar\omega}{2}\lambda_2 = -\frac{\hbar\omega}{2}$, $\left|E_1\right\rangle = \left|\lambda_1\right\rangle$ and $\left|E_2\right\rangle = \left|\lambda_2\right\rangle$:

$$H\left|E_1\right\rangle = \frac{\hbar\omega}{2}\sigma_z\left|\lambda_1\right\rangle = \frac{\hbar\omega}{2}\left|\lambda_1\right\rangle = E_1\left|E_1\right\rangle \quad ✔$$

$$H\left|E_2\right\rangle = \frac{\hbar\omega}{2}\sigma_z\left|\lambda_2\right\rangle = -\frac{\hbar\omega}{2}\left|\lambda_2\right\rangle = E_2\left|E_2\right\rangle \quad ✔ \quad \blacksquare$$

Exercise 4.5 (Corollary to Example). For an arbitrary direction $\vec{n} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}$

$$E_1 = \frac{\hbar\omega}{2},\ \ E_2 = -\frac{\hbar\omega}{2},\ \ \left|E_1\right\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix} \text{ and } \left|E_2\right\rangle = \begin{pmatrix} \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2}e^{i\phi} \end{pmatrix}.$$

Solution. From (3.23), the spin operator in the $\hat{n}$-direction is

$$\sigma_n = \begin{pmatrix} n_z & n_x - n_y i \\ n_x + n_y i & -n_z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & \text{-}\cos\theta \end{pmatrix}.$$

From Exercise 3.4, the eigenvalues and eigenvectors of σ_n are

$$\lambda_1 = +1,\ \ \lambda_2 = -1,\ \ \left|\lambda_1\right\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix} \text{ and } \left|\lambda_2\right\rangle = \begin{pmatrix} \sin\theta \\ -\cos\frac{\theta}{2}e^{i\phi} \end{pmatrix}.$$

Hence $E_1 = \frac{\hbar\omega}{2}\lambda_1 = \frac{\hbar\omega}{2}$, $E_2 = \frac{\hbar\omega}{2}\lambda_2 = -\frac{\hbar\omega}{2}$, $\left|E_1\right\rangle = \left|\lambda_1\right\rangle$, and $\left|E_2\right\rangle = \left|\lambda_2\right\rangle$. ■

Usually we have some apparatus to prepare the system, so an initial state $\left|\Psi(0)\right\rangle$ is known.

Example. Suppose an initial state along with the energy eigenvalues $\{E_j\}$ and eigenvectors $\{\left|E_j\right\rangle\}$ are given. Solve the time-dependent Schrödinger equation for $\left|\Psi(t)\right\rangle$.

Solution. We wish to solve

$$\text{(i)}\ \ \hbar\frac{\partial\left|\Psi\right\rangle}{\partial t} = -iH\left|\Psi\right\rangle.$$

The state vector $\left|\Psi\right\rangle$ has a unique representation in terms of the orthonormal basis $\{\left|E_j\right\rangle\}$:

$$\text{(ii)}\ \ \left|\Psi\right\rangle = \sum_j \alpha_j\left|E_j\right\rangle$$

We denote the time evolution of the state vector in terms of the basis $\left[\left|E_j\right\rangle\right]$:

$$\left|\Psi(t)\right\rangle = \sum_j \alpha_j(t)\left|E_j\right\rangle \tag{4.29}$$

Therefore

$$-\frac{i}{\hbar}\sum_j E_j\,\alpha_j(t)\left|E_j\right\rangle \overset{(4.28)}{=} -\frac{i}{\hbar}H\sum_j \alpha_j(t)\left|E_j\right\rangle \overset{(4.29)}{=} -\frac{i}{\hbar}H\left|\Psi(t)\right\rangle \overset{(i)}{=} \frac{\partial\Psi}{\partial t} \overset{(4.29)}{=} \sum_j \dot{\alpha}_j(t)\left|E_j\right\rangle$$

$$\Rightarrow\ \sum_j\left\{\dot{\alpha}_j(t)+\frac{i}{\hbar}E_j\,\alpha_j(t)\right\}\left|E_j\right\rangle = 0 \quad \Rightarrow\ \forall\,\mathrm{j}\ \ \frac{d\alpha_j(t)}{dt} = -\frac{i}{\hbar}E_j\,\alpha_j(t)$$

since $\left\{\left|E_j\right\rangle\right\}$ are linearly independent. The solution to this standard DE is

$$\alpha_j(t) = \alpha_j(0)\,e^{-\frac{i}{\hbar}E_j t} \tag{4.30}$$

This equation holds generally for any system, not just spin states, provided *H* does not depend on time.

Letting $\left|A\right\rangle = \left|\Psi(0)\right\rangle$ and $\left|j\right\rangle = \left|E_j\right\rangle$ in equation 1.5 yields

$$\alpha_j(0) = \left\langle E_j \middle| \Psi(0)\right\rangle \tag{4.31}$$

Therefore

$$\left|\Psi(t)\right\rangle \overset{(4.29)}{=} \sum_j \alpha_j(t)\left|E_j\right\rangle \overset{(430)}{=} \sum_j \alpha_j(0)\,e^{-\frac{i}{\hbar}E_j t}\left|E_j\right\rangle$$

$$\left|\Psi(t)\right\rangle \overset{(431)}{=} \sum_j \left\langle E_j \middle| \Psi(0)\right\rangle e^{-\frac{i}{\hbar}E_j t}\left|E_j\right\rangle \tag{4.32}$$

or, emphasizing the basis nature of $\left\{\left|E_j\right\rangle\right\}$,

$$\left|\Psi(t)\right\rangle = \sum_j \left|E_j\right\rangle\,\left\langle E_j \middle| \Psi(0)\right\rangle\, e^{-\frac{i}{\hbar}E_j t} \tag{4.33}$$

since $\left\langle\, E_j \middle| \Psi(0)\,\right\rangle$ is a scalar. ■

Note. Definition 7.01 will define $\left|E_j\right\rangle\,\left\langle E_j \middle| \Psi(0)\right\rangle$ to be $\left|E_j\right\rangle\left\langle E_j\right|\ \Psi(0)\rangle$, an outer product (an observable) times a vector. So (4.33) describe $\left|\Psi(t)\right\rangle$ as the product of $e^{-\frac{i}{\hbar}E_j t}$ and the sum of outer product observables acting on energy eigenvectors.

SUMMARY: Cookbook Recipe – Computing Time-evolution of a Quantum State

1) Derive, measure, approximate, or guess an equation for H
2) Use an apparatus to prepare an initial state $|\Psi(0)\rangle$; in other words, to prepare a state vector $|\Psi(0)\rangle$
3) Find the energy eigenvalues E_j and eigenvectors $|E_j\rangle$ of H by solving the time-independent Schrodinger equation $H|E_j\rangle \overset{(4.28)}{=} E_j|E_j\rangle$
4) Calculate $\alpha_j(0) = \langle E_j|\Psi(0)\rangle$
5) Compute $|\Psi(0)\rangle = \sum_j \alpha_j(0)|E_j\rangle$
6) Expand $|\Psi(t)\rangle = \sum_j \alpha_j(t)|E_j\rangle$ in terms of $\{\alpha_j(t)\}$
7) Replace $\alpha_j(t)$ in (6) with $\alpha_j(0)e^{-\frac{i}{\hbar}E_j t}$:
$$|\Psi(t)\rangle = \sum_j \alpha_j(0)e^{-\frac{i}{\hbar}E_j t}|E_j\rangle \tag{4.34}$$
8) Specify a new observable L at time t, compute its eigenvalues $\{\lambda\}$ and eigenvectors $\{|\lambda\rangle\}$ and calculate the probabilities of the outcomes:
$$\mathrm{P}_L(\lambda) = |\langle\lambda|\Psi(t)\rangle|^2$$

Exercise 4.6. Use the Cookbook Recipe to complete the magnetic field example on page 32: Given Hamiltonian $H = \frac{\hbar\omega}{2}\sigma_z$, state vector $|\Psi(0)\rangle = |r\rangle$, and basis $\{|u\rangle, |d\rangle\}$ show that $\langle\sigma_y\rangle$ varies sinusoidally over time.

Definition. Before measurement, a system is in a superposition state $\sum_j \alpha_j|\lambda_j\rangle$.

After measurement, it jumps to state $|\lambda_j\rangle$ for some j (with probability $|\alpha_j|^2$). This is called the **collapse of the wave function.** More light will be shed on this topic in Chapter 6 where we treat the measurement *apparatus* as quantum mechanical.

Chapter 5. Uncertainty and Time Dependence

Having established a good analytic framework for a single observable we now investigate the notion of two or more observables, say *L* and *M*. We wish to establish conditions for when *L* and *M* can be simultaneously (or even sequentially) measured.

Case A. *L* and *M* are both associated with Alice's system
In this situation it may or may not be possible to measure, say, *M* after *L* has been measured. For example *L* = σ_z measures **spin** in the up direction and disturbs the system so that a measurement σ_x of spin in the x-direction is not valid. However, if *L* measures the **position of a particle** it is possible to simultaneously measure the x-, *y*-, and *z*-positions.

Case B. *L* is associated with Alice's system and *M* with Bob's
In this case the measurements are independent and can both be performed no matter what the observables are.

There can be a difference in the mathematical notation between the 2 cases. Many of the equations developed in this chapter will pertain to both cases. Those that are specific to Case A will be called out. Chapter 6 is devoted to Case B.

Definition. The notation $\left|\boldsymbol{AB}\right\rangle$ will be used in Chapter 6 to denote a state vector of a combined system governed by Case B. In this chapter $\left|\boldsymbol{A,B}\right\rangle$ will be used to denote a combined state vector and applies to either Case A or Case B.

Suppose the observables *L* and *M* are represented by matrices. In Case A we can add or multiply the matrices as desired. In Case B, however, the matrices cannot be combined in any way. In Case B, when *L* or *M* are operating on a combined state vector, *L* operates solely on Alice's part of the vector, *M* is restricted to Bob's part, and *L* and *M* cannot be inter-mingled.

Definition. Let *L* and *M* be Hermitian operators. Let $\{\lambda\}$ and $\{\mu\}$ be the respective eigenvalues of *L* and *M*, and $\{\left|\lambda\right\rangle\}$ and $\{\left|\mu\right\rangle\}$ the respective eigenvectors. If *L* and *M* can be simultaneously measured, the outcome is an **eigenvalue pair** $(\boldsymbol{\lambda}, \boldsymbol{\mu})$. The **corresponding eigenvector** is denoted $\left|\boldsymbol{\lambda, \mu}\right\rangle$.
Because *L* operates only on $\left|\lambda\right\rangle$ and *M* operates only on $\left|\mu\right\rangle$ we have that

$$L\left|\lambda, \mu\right\rangle = \lambda\left|\lambda, \mu\right\rangle \text{ and } M\left|\lambda, \mu\right\rangle = \mu\left|\lambda, \mu\right\rangle.$$

We say that $|\lambda, \mu\rangle$ **are simultaneous eigenvectors of the operators *L* and *M*.** The **product space** is the vector space generated using $\{|\lambda, \mu\rangle\}$ as a basis.

Theorem. There is a basis $\{|\lambda, \mu\rangle\}$ of simultaneous eigenvectors of *L* and *M* if and only if *L* and *M* commute.

Proof. Suppose $\{|\lambda, \mu\rangle\}$ is a basis of simultaneous eigenvectors. Then

$$LM|\lambda, \mu\rangle = L\mu|\lambda,\mu\rangle = \lambda\mu|\lambda, \mu\rangle = \mu\lambda|\lambda, \mu\rangle = M\lambda|\lambda,\mu\rangle = ML|\lambda, \mu\rangle \;\; \forall\, |\lambda,\mu\rangle$$

$\Rightarrow \quad LM|A, B\rangle = ML|A, B\rangle$ for all product vectors $|A, B\rangle$:

Let $|A\rangle = \sum_{\lambda} \alpha_\lambda |\lambda\rangle$ and $|B\rangle = \sum_{\mu} \beta_\mu |\mu\rangle$. Then

$$|A,B\rangle = \left|\sum_{\lambda} \alpha_\lambda |\lambda\rangle, \sum_{\mu} \beta_\mu |\mu\rangle\right\rangle \overset{\text{(Linear)}}{=} \sum_{\lambda,\mu} \alpha_\lambda \beta_\mu |\lambda,\mu\rangle \text{ and}$$

$$LM|A,B\rangle = LM\sum_{\lambda,\mu} \alpha_\lambda \beta_\mu |\lambda,\mu\rangle = \sum_{\lambda,\mu} \alpha_\lambda \beta_\mu LM|\lambda,\mu\rangle$$

$$= \sum_{\lambda,\mu} \alpha_\lambda \beta_\mu ML|\lambda,\mu\rangle = ML|A,B\rangle$$

$\Rightarrow \quad LM = ML$. ✔

The book claims but does not prove the converse. Suppose *L* and *M* commute. The goal is to produce a collection $\{|\lambda,\mu\rangle\}$ that can be used as a basis to generate the product space. Let $|\lambda\rangle$ and $|\mu\rangle$ be respective eigenvectors of *L* and *M*. Since $LM|\lambda, \mu\rangle = ML|\lambda, \mu\rangle$, then in particular $LM|\lambda,\mu\rangle$ and $ML|\lambda,\mu\rangle$ exist. Thus, measurement of $|\mu\rangle$ by *M* does not prevent measurement of $|\lambda\rangle$ by *L*, and vice-versa. Hence, $|\lambda, \mu\rangle$ satisfies the requirement to be a simultaneous eigenvector, and the basis collection $\{|\lambda,\mu\rangle\}$ is a well-defined set. ■

Simultaneity Principle. The condition for operators *L*, *M*, *N*, … to be simultaneously measurable is that they pairwise commute or, equivalently, that their commutators all equal the zero operator.

Definition. If two or more observables commute we say they are **simultaneously measurable.** If two observables do not commute, it is generally not possible to have unambiguous knowledge of both. A **set of observables is complete** if their set of simultaneous eigenvectors forms a (complete) basis.

Exercise 5.1. Any 2x2 Hermitian matrix L can be written as

$$L = a\sigma_x + b\sigma_y + c\sigma_z + d\,I \quad \text{where } a,b,c,d \in \mathbb{R}.$$

Theorem. Let

$$L = a_1\sigma_x + b_1\sigma_y + c_1\sigma_z + d_1 I$$

and

$$M = a_2\sigma_x + b_2\sigma_y + c_2\sigma_z + d_2 I$$

be observables of Alice's system.

Then L and M can be simultaneously measured iff $\begin{cases} a_1 b_2 = a_2 b_1 \\ b_1 c_2 = b_2 c_1 \\ c_1 a_2 = c_2 a_1 \end{cases}$ (I)

Proof. They can be simultaneously measured iff [M, L] = 0. Observe that $\sigma_x\sigma_y = -\sigma_y\sigma_x$, $\sigma_y\sigma_z = -\sigma_z\sigma_y$, and $\sigma_z\sigma_x = -\sigma_x\sigma_z$:

$$\sigma_x\sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$\sigma_y\sigma_x = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -\sigma_x\sigma_y \qquad ✔$$

Etc.

Thus,

$$\begin{aligned} LM = {} & \left(a_1 a_2 \sigma_x^{\ 2} + b_1 b_2 \sigma_y^{\ 2} + c_1 c_2 \sigma_z^{\ 2} + d_1 d_2 I^2\right) \\ & + \left(a_1 b_2 - a_2 b_1\right)\sigma_x\sigma_y + \left(b_1 c_2 - b_2 c_1\right)\sigma_y\sigma_z + \left(c_1 a_2 - c_2 a_1\right)\sigma_z\sigma_x \\ & + \left(a_1 d_2 + a_2 d_1\right)\sigma_x + \left(b_1 d_2 + b_2 d_1\right)\sigma_y + \left(c_1 d_2 + c_2 d_1\right)\sigma_z \end{aligned}$$

and

$$\begin{aligned} ML = {} & \left(a_1 a_2 \sigma_x^{\ 2} + b_1 b_2 \sigma_y^{\ 2} + c_1 c_2 \sigma_z^{\ 2} + d_1 d_2 I^2\right) \\ & - \left(a_1 b_2 - a_2 b_1\right)\sigma_x\sigma_y - \left(b_1 c_2 - b_2 c_1\right)\sigma_y\sigma_z - \left(c_1 a_2 - c_2 a_1\right)\sigma_z\sigma_x \\ & + \left(a_1 d_2 + a_2 d_1\right)\sigma_x + \left(b_1 d_2 + b_2 d_1\right)\sigma_y + \left(c_1 d_2 + c_2 d_1\right)\sigma_z. \end{aligned}$$

Hence

$$\begin{aligned} \left[M,L\right] &= LM - ML \\ &= 2\left[\left(a_1 b_2 - a_2 b_1\right)\sigma_x\sigma_y + \left(b_1 c_2 - b_2 c_1\right)\sigma_y\sigma_z + \left(c_1 a_2 - c_2 a_1\right)\sigma_z\sigma_x\right] \\ &= A\sigma_x\sigma_y + B\sigma_y\sigma_z + C\sigma_z\sigma_x \end{aligned}$$

where

$$A = a_1 b_2 - a_2 b_1,\ \ B = b_1 c_2 - b_2 c_1,\ \text{ and }\ C = c_1 a_2 - c_2 a_1.$$