PY 451 Notes — February 1, 2018
McIntyre Chap. 2 — Operators and Measurement

Professor Kenneth Lane[1]

Much of the material of Chapter 2 was covered in my notes and lectures on Chapter 1: extended discussions of QM Postulates 1–5, the importance of hermitian operators, especially as observables, and their eigenvalues ("ev's") and eigenvectors ("evecs"), and projection operators. In postulates 2 and 3, it is absolutely essential that you remember to insert the word *hermitian*:

Postulate 2: A physical observable $\mathcal{A}$ is represented mathematically by an abstract **hermitian operator** $A = A^\dagger$ that acts on state vectors, kets: $|\psi_A\rangle = A|\psi\rangle$. The corresponding bra vector is $\langle\psi_A| = \langle\psi|A^\dagger \equiv \langle\psi|A$ for hermitian A.

Postulate 3: The only possible result of a measurement of observable $\mathcal{A}$ is one of the *real* eigenvalues a_i of the corresponding hermitian operator A.

Note: From now on, I will use the subscripts i, j, k, l to label the individual ev's a_i, a_j, etc. and evecs $|a_i\rangle, |a_j\rangle, \ldots$ of A. As is standard practice among physicists, I prefer to use the letter n to denote the total number of ev's and *linearly independent* evecs, thus:

$$A|a_i\rangle = a_i|a_i\rangle, \qquad i = 1, 2 \ldots n. \tag{1}$$

As I mentioned in class, this notation allows for the possibility that the eigenvalue a_i is *d-fold degenerate*, i.e., there are d *orthonormalized* evecs with ev a_i, denoted by $|a_i, 1\rangle, |a_i, 2\rangle, \ldots |a_i, d\rangle$:

$$\begin{aligned} A|a_i, j\rangle &= a_i|a_i, j\rangle, \\ \langle a_i, j|a_1, k\rangle &= \delta_{jk}, \qquad j, k = 1, 2 \ldots, d. \end{aligned} \tag{2}$$

The very meaning of the statement that the ev a_i is d-fold degenerate is that there are d *linearly independent* eigenvectors of A with the same eigenvalue. We'll see better how the different evecs get labeled when I discuss commuting observables and complete sets of commuting observables (CSCO's) later in these notes. Any such set of eigenvectors can be made orthogonal to one another by what is called the Gram-Schmidt process (which I'll describe in discussion section if you like or you can Google it).

Furthermore, evecs of A corresponding to different ev's are orthogonal. Here's the proof: For $A = A^\dagger$,

$$(a_i - a_j)\langle a_i|a_j\rangle = \left(\langle a_i|A^\dagger\right)|a_j\rangle - \langle a_i|\left(A|a_j\rangle\right) = \langle a_i|A|a_j\rangle - \langle a_i|A|a_j\rangle = 0. \tag{3}$$

Therefore, $\langle a_i|a_j\rangle = 0$ if $a_i \neq a_j$. QED

In the Chapter 1 notes, I defined the projection operator on the evec $|a_i\rangle$ of the hermitian operator A:

$$P_{a_i} = |a_i\rangle\langle a_i|. \tag{4}$$

If ev a_i is nondegenerate, we can also say that P_{a_i} projects state vectors onto a_i. For any state vector $|\psi\rangle$,

$$|\psi(a_i)\rangle \equiv P_{a_i}|\psi\rangle = c_{a_i}|a_i\rangle, \text{ where } c_{a_i} = \langle a_i|\psi\rangle \tag{5}$$

satisfies $A|\psi(a_i)\rangle = a_i|\psi(a_i)\rangle$. If a_i is d-fold degenerate, then the projection operator onto it is

$$P_{a_i} = \sum_{j=1}^{d} |a_i, j\rangle\langle a_i, j|. \tag{6}$$

Note that projection operators are hermitian and "idempotent":

$$\begin{aligned} P_{a_i}^\dagger &= (|a_i\rangle\langle a_i|)^\dagger = |a_i\rangle\langle a_i| = P_{a_i}, & (7)\\ P_{a_i}^2 &= |a_i\rangle\langle a_i|\,|a_i\rangle\langle a_i| \equiv |a_i\rangle\langle a_i|a_i\rangle\langle a_i| = P_{a_i}. & (8)\end{aligned}$$

It is straightforward to extend this to the case of degenerate ev's.

Finally, a *very important theorem*, that is obvious for 2-state systems and which I state without proof for a state vector space of *any* dimensionality: **The evecs of hermitian operator A corresponding to observable $\mathcal{A}$ form a complete orthonormal set in the state-vector space of a physical system.** That is, for *any* state vector $|\psi\rangle$,

$$|\psi\rangle = \sum_{i=1}^{n} \langle a_i|\psi\rangle\, |a_i\rangle \equiv \sum_{i=1}^{n} |a_i\rangle\langle a_i|\psi\rangle \Longrightarrow \sum_{i=1}^{n} P_{a_i} \equiv \sum_{i=1}^{n} |a_i\rangle\langle a_i| = \mathbf{1}, \tag{9}$$

where $\mathbf{1}$ is the *identity* or *unit* operator: $\mathbf{1}$ acting on "anything" equals "anything". A very useful trick in doing certain calculations or proving results is to *insert a complete set of states* between two operators or an operator and a vector. We'll see examples of this below. And, lastly, note the following important result for any hermitian operator:

$$A = \sum_{i}^{n} a_i |a_i\rangle\langle a_i| \equiv \sum_{i}^{n} a_i P_{a_i}. \tag{10}$$

Now a collection of remarks to supplement Chapter 2 in McIntyre:

1.) Using Eq. (10), we have for the operator S_z of a spin-$\frac{1}{2}$ system:

$$S_z = \frac{\hbar}{2}\big(|+\rangle\langle+| - |-\rangle\langle+|\big), \tag{11}$$

where, as usual (and unless stated otherwise) we are working in the eigenbasis of S_z, namely $|\pm\rangle = |\pm\rangle_z$. In this basis, the matrix representing S_z is

$$S_z \doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{12}$$

Now let's use our tools to calculate the representation of S_x in this basis (also, see the note (2) on page 5 below):

$$S_x = \sum_{a_i, a_j = \pm} |a_i\rangle\langle a_i|S_x|a_j\rangle\langle a_j|. \tag{13}$$

Using $|\pm\rangle = (1/\sqrt{2})\big(|+\rangle_x \pm |-\rangle_x\big)$, the evecs of S_x, we obtain the hermitian result (in the S_z basis!)

$$S_x = \frac{\hbar}{2}\big(|+\rangle\langle-| + |-\rangle\langle+|\big) \doteq \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{14}$$

The numbers $\langle a_i|S_x|a_j\rangle$ are called the *matrix elements* of S_x in the S_Z-basis $\{|a_1\rangle = |+\rangle, |a_2\rangle = |-\rangle\}$. To get used to this sort of calculation, you should repeat it to determine the matrix representing S_y. Finally, note that while S_z is diagonal in the $|\pm\rangle$ basis, S_x is not! This is a very important feature of the operators of angular momentum, discussed further when we discuss the "commutators" of operators.

2.) **Diagonalization of Operators** — using Dirac notation:
Suppose that $A|a_i\rangle = a_i|a_i\rangle$, $(i = 1, 2, \ldots, n)$. How do we find A's eigenbasis when we are in some other orthonormal basis? Suppose that we are working in the complete orthonormal basis $\{|\chi_1\rangle, |\chi_2\rangle, \ldots, |\chi_n\rangle\}$ with $\sum_{i=1}^n |\chi_i\rangle\langle\chi_i| = \mathbf{1}$ and $\langle\chi_i|\chi_j\rangle = \delta_{ij}$. In this basis

$$A \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \ldots \\ A_{21} & A_{22} & A_{23} & \ldots \\ A_{31} & A_{32} & A_{33} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{15}$$

where $A_{ij} = \langle\chi_i|A|\chi_j\rangle$ and

$$|a_i\rangle = \sum_{j=1}^n |\chi_j\rangle\langle\chi_j|a_i\rangle \equiv \sum_{j=1}^n c_{ji}|\chi_j\rangle. \tag{16}$$

Then, inserting a complete set of states twice:

$$a_i|a_i\rangle = A|a_i\rangle = \sum_{k,l=1}^n |\chi_k\rangle\langle\chi_k|A|\chi_l\rangle\langle\chi_l|a_i\rangle \equiv \sum_{k,l=1}^n |\chi_k\rangle A_{kl}c_{li} = a_i\sum_{k=1}^n |\chi_k\rangle c_{ki}. \tag{17}$$

Multiplying on the left by $\langle\chi_j|$ gives

$$\sum_{l=1}^n A_{jl}c_{li} = a_i c_{ji} \Longrightarrow \sum_{l=1}^n (A_{jl} - \delta_{jl}a_i)c_{li} = 0. \tag{18}$$

These homogeneous equations have nontrivial solutions for the eigenvectors $|a_i\rangle = \sum_{k=1}^n c_{ki}|\chi_k\rangle$ if and only if

$$\det(A - \lambda\mathbf{1}) = \det\begin{pmatrix} A_{11}-\lambda & A_{12} & A_{13} & \ldots \\ A_{21} & A_{22}-\lambda & A_{23} & \ldots \\ A_{31} & A_{32} & A_{33}-\lambda & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = 0, \tag{19}$$

i.e., iff $A - \lambda\mathbf{1}$ has *no* inverse; otherwise you could prove that all $c_{ki} = 0$.

The 2×2 case of, e.g., spin-$\frac{1}{2}$ is easy because you only have to solve a quadratic equation for its two solutions, the ev's a_1 and a_2. In the $n \times n$ case, once the n solutions $\lambda = (a_1, .a_2, \ldots, a_n)$ are found, you can obtain the eigenvectors $|a_i\rangle$ in the χ-basis by solving

$$\sum_{k=1}^n A_{jk}c_{ki} = a_i c_{ji}. \tag{20}$$

The example of diagonalizing S_y — i.e., finding its ev's and evecs in the S_z-basis — is worked out in the text.

3.) The spin-$\frac{1}{2}$ operator along an arbitrary axis, $\hat{\boldsymbol{n}} = \hat{\boldsymbol{x}} \sin\theta\cos\phi + \hat{\boldsymbol{y}} \sin\theta\sin\phi + \hat{\boldsymbol{z}}\cos\theta$, for polar angle θ and azimuthal angle ϕ, is given by

$$S_n = \boldsymbol{S}\cdot\hat{\boldsymbol{n}} = S_x \sin\theta\cos\phi + S_y \sin\theta\sin\phi + S_z\cos\theta \doteq \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}. \tag{21}$$

Unsurprisingly, it's ev's are $\pm\hbar/2$. The corresponding evecs are

$$|+\rangle_n \doteq \begin{pmatrix} \cos(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} \end{pmatrix} \text{ and } |-\rangle_n \doteq \begin{pmatrix} -\sin(\theta/2)\, e^{-i\phi/2} \\ \cos(\theta/2)\, e^{i\phi/2} \end{pmatrix}. \tag{22}$$

These differ from those given in the text by a factor of $e^{-i\phi/2}$ and $-e^{i\phi/2}$, respectively. The reason I have done this is due to the following theorem:

Theorem: Any $n\times n$ hermitian matrix $A = A^\dagger$ is diagonalized by a unitary $n\times n$ matrix U ($U^\dagger = U^{-1}$),

$$A_{\text{diag}} \equiv \begin{pmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = U^\dagger A U. \tag{23}$$

The columns of U are — up to an overall phase of U — the orthonormalized eigenvectors of U.

This theorem is almost self-evident. You can demonstrate it for S_n with the unitary matrix U_n of its eigenvectors in Eq. (22):

$$(\ |+\rangle_n\, ,\ \ |-\rangle_n\ ) \doteq \begin{pmatrix} \cos(\theta/2)\, e^{-i\phi/2} & -\sin(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} & \cos(\theta/2)\, e^{i\phi/2} \end{pmatrix} \equiv U_n. \tag{24}$$

When you calculate S_nU_n, S_n multiplies each column of U_n to give back the evec times its ev:

$$S_nU_n \doteq \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix} \cos(\theta/2)\, e^{-i\phi/2} & -\sin(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} & \cos(\theta/2)\, e^{i\phi/2} \end{pmatrix} \doteq \frac{\hbar}{2}(\ |+\rangle_n\, ,\ \ -|-\rangle_n\ ). \tag{25}$$

Finally, since the evecs $|\pm\rangle_n$ are orthonormalized,

$$U_n^\dagger S_n U_n = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \doteq S_z\ \ !!! \tag{26}$$

This establishes the theorem in this rather general 2×2 case.

Before leaving this nice little toy, S_n, play with it some more for various values of θ and ϕ, e.g. (1) both small but nonzero; (2) one equal zero and the other $\pi/2$; (3) both near $\pi/2$; etc. Look at the probabilities $\mathcal{P}_{x,y,z}$. Do you understand the results you are getting?

Three comments:

(1) As my calculation of diagonalizing S_n by U_n shows, it is becoming painful to keep writing $\doteq$ to distinguish between an operator or a ket/bra vector and its matrix representation. Therefore, except in cases where the distinction is important, I will stop using the $\doteq$ equality. The important thing to remember is that the elements of a column (row) vector representing a ket (bra) vector are

really its <u>amplitudes</u> in a particular complete orthonormal basis, and the square matrix representing an operator is composed of the operator's <u>matrix elements</u> (which are also just amplitudes!) in that basis.

(2) In my calculation of S_x in the S_z-basis, I inserted a complete sets of S_x-evecs and made use of the relations $\langle\pm|+\rangle_x = 1/\sqrt{2}$, $\langle\pm|-\rangle_x = \pm 1/\sqrt{2}$. As Qianyi and probably others of you realized in class, what I was doing was exactly equivalent to "rotating S_z into S_x". That is, since the matrix for S_x in the x-basis is the same as for S_z in the z-basis,

$$(S_x)_{z-\text{basis}} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}\begin{pmatrix} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{pmatrix}\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}. \tag{27}$$

I did just the inverse of the same thing in Eq. (26 To diagonalize S_n, I rotated it into S_z by first rotating S_n about the z-axis by $-\phi$, and then rotating about the y-axis by $-\theta$.

(3) Now read the analyses of Experiments 3 and 4 on your own. Pay attention to the use of projection operators and how Experiment 4 is explained by summing $P_{+x} + P_{-x} = |+\rangle_x\langle+|_x + |-\rangle_x\langle-|_x = \mathbf{1}$. The discussion of $\mathcal{P}_+$ and $\mathcal{P}_-$ on pages 48-49 and the importance of interference terms arising when we don't disturb (observe!) the two beams of Ag ions that occur when we sum over states is an *especially important* feature of quantum mechanics.

4.) <u>**Measurement of Observables in QM**</u>

First, carefully read McIntyre's remarks in the opening paragraph of Section 2.3. I can't say it better than he does: In quantum mechanics, a single measurement is meaningless! For a state that's a single eigenstate (evec) of an operator observable A, there is always measurement error (statistical and systematic) to contend with. For the more usual state $|\psi\rangle$ that's a superposition of such evecs, apart from any measurement error there are "quantum fluctuations". These will be accounted for in the *expected value* A or the *expectation value* or average value $\langle A\rangle_\psi$ of A in the state $|\psi\rangle$ and the *standard deviation* or *root-mean-square (r.m.s.) deviation* ΔA, both defined next.

For a spin-$\frac{1}{2}$ system — i.e. a <u>*large*</u> number of identically prepared systems — in the state $|\psi\rangle = c_+|+\rangle + c_-|-\rangle$ where $c_\pm = \langle\pm|\psi\rangle$, the probability of obtaining $a_\pm = \pm\frac{1}{2}\hbar$ in a measurement of S_z is

$$\mathcal{P}_\pm = |c_\pm|^2 = |\langle\pm|\psi\rangle|^2 \tag{28}$$

and the expected value or average value or predicted mean value (choose your favorite terminology) after many measurements of S_z is, by completeness of $\{|+\rangle, |-\rangle\}$,

$$\begin{aligned}\langle S_z\rangle_\psi &= \tfrac{1}{2}\hbar\,\langle\psi|+\rangle\langle+| + (-\tfrac{1}{2}\hbar)\langle\psi|-\rangle\langle-|\psi\rangle \\ &= \langle\psi|S_z\big(|+\rangle\langle+| \,+\, |-\rangle\langle-|\big)|\psi\rangle \equiv \langle\psi|S_z|\psi\rangle.\end{aligned} \tag{29}$$

More generally, for any observable A,

$$\langle A\rangle_\psi = \sum_{i=1}^{n} a_i\mathcal{P}_{a_i} \equiv \langle\psi|A|\psi\rangle. \tag{30}$$

The "typical" spread in results of repeated measurement of A in identically prepared states $|\psi\rangle$ is called the standard deviation or r.m.s. deviation, ΔA, and is defined as the *square root* of the *mean* of the *square* of the deviations from the mean, or average, value:

$$(\Delta A)_\psi \equiv \sqrt{\langle (A - \langle A\rangle_\psi)^2\rangle_\psi} = \sqrt{\langle A^2\rangle_\psi - \langle A\rangle_\psi^2}\,, \tag{31}$$

as shown in the text. We'll usually drop the subscript ψ, as the state $|\psi\rangle$ is understood, and simply write

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}\,. \tag{32}$$

N.B.: It is always understood that $|\psi\rangle$ is *normalized*. Otherwise, we have to divide all these expressions for $\langle A\rangle$ and ΔA by $||\psi||^2 = \langle\psi|\psi\rangle$. Here is a little theorem for you to prove:

Theorem: The standard deviation $\Delta A \geq 0$, with equality holding if only if $|\psi\rangle$ is an eigenvector of A: $A|\psi\rangle = a|\psi\rangle$.

That $\Delta A > 0$ unless $A|\psi\rangle = a|\psi\rangle$ is obvious from its definition, Eq. (31). Don't use that. First, prove it directly for a 2-state system — an illuminating calculation. For the general case, insert a complete set of states in A^2. This theorem gives meaning to "quantum fluctuations". In a classical measurement of a quantity A, modulo experimental error, there is no "spread" in the result of its repeated measurement(s) on identically prepared systems. But, in a quantum state, $\Delta A > 0$ unless $A|\psi\rangle = a|\psi\rangle$. The contribution of all other states than $|\psi\rangle$ to ΔA is a purely quantum effect.

5.) **Commuting Observables and CSCO's**

A measurement of S_z with a particular outcome, say $+\frac{1}{2}\hbar$, followed by a measurement of S_x "destroys" the measurement of S_z. When S_z is measured again, it is equally likely that it will be $-\frac{1}{2}\hbar$ as it will be $\frac{1}{2}\hbar$. We say that the measurements of S_z and S_x are *incompatible*. Equivalently, we cannot write a state vector that is an eigenstate of *both* S_z and S_x or any other component of $\mathbf{S}$. This, too, is a strictly quantum effect. The reason for this is that the *commutator* $[S_z, S_x] \equiv S_zS_x - S_xS_z \neq 0$. This result is generally true in QM and is summarized in this theorem:

Theorem: Two *observables* A and B have a simultaneous set, i.e., a common set of eigenvectors if and only if they commute, $[A, B] = 0$.

Proof: Suppose $[A, B] = 0$. Let $\{|a_i\rangle, i = 1, 2, \ldots\}$ denote a complete orthonormal set of eigenvectors of A, with $A|a_i\rangle = a_i|a_i\rangle$. Consider the state $B|a_i\rangle$. Then

$$A(B|a_i\rangle) = BA|a_i\rangle = a_iB|a_i\rangle \tag{33}$$

is also an eigenvector of A with the same ev a_i. If a_i is nondegenerate (we'll argue later that it can't be, but never mind that now), then $B|a_i\rangle$ is proportional to $|a_i\rangle$, that is, $B|a_i\rangle = b|a_i\rangle$. So, it is an eigenvector of B too. If a_i is d-fold degenerate, with evecs $\{|a_i, j\rangle,\ j = 1, 2, \ldots, d\}$, then $B|a_i, j\rangle$ is again an evec of A with ev a_i. Now, however, $B|a_i, j\rangle$ is in general a linear combination of the evecs of A with ev a_i. Consider the matrix elements $\langle a_i, j|B|a_i, k\rangle$. Since B is hermitian,

$$\langle a_i, j|B|a_i, k\rangle^* = \langle a_i, k|B^\dagger|a_i, j\rangle = \langle a_i, k|B|a_i, j\rangle. \tag{34}$$

That is, in this subspace of the d degenerate evecs of operator A, the operator B is represented by a *hermitian* matrix and, so it can be diagonalized with eigenvectors that are d linearly independent — and orthonormal — combinations of the $|a_i, j = 1, 2, \ldots, d\rangle$. Call them $|a_i, b_j\rangle$ where $B|a_i, b_j\rangle = b_j|a_i, b_j\rangle$ as well as $A|a_i, b_j\rangle = a_i|a_i, b_j\rangle$. $\frac{1}{2}$ QED

Conversely, suppose that $A|a_i, b_j\rangle = a_i|a_i, b_j\rangle$ and $B|a_i, b_j\rangle = b_j|a_i, b_j\rangle$ for the complete orthonormal set of basis states in the vector space in which A and B act. Then it is easy to prove that $[A, B] = 0$. All you need is the fact that, if all the matrix elements of an operator $\mathcal{O}$ equal zero, then $\mathcal{O} = 0$. QED

Let's look at this from the point of view of measurement (see, e.g., Fig. 2.9 in the text). If the ev a_i is d-fold degenerate, a measurement of A giving a_i results in a state that is, in general, a linear superposition of the corresponding evecs of A, $\{|a_i, j\rangle,\ j = 1, 2, \ldots, d\}$. There must be *something* to distinguish these states, to *label* them, something other than just an index j. There is; it's another observable, a hermitian operator B which is compatible with A. A subsequent measurement of B will return an ev, b_j, and leave us in an eigenstate of B which we can call $|a_i, b_j\rangle$. Now, either this eigenstate is unique (up to a phase; it's normalized, remember), or it's not, i.e., there is further degeneracy. If it's unique, then the state is completely specified. According to QM postulate 1, that it is a simultaneous eigenstate of A and B with eigenvalues a_i and b_j is *all* the information we can obtain about the state. If it's not unique, i.e., it is *degenerate*, then there must be another observable, C, which is compatible with A and B and whose measurement will further specify the state. When this process is done, when there is no further degeneracy to "lift", we have what we call a *"complete set of commuting observables"*, or CSCO's, A, B, $C, \ldots$. Once we know them, their common eigenstates $\{|a_i, b_j, c_k, \ldots\rangle;\ i = 1, 2 \ldots d_i; j = 1, 2, \ldots, d_j;$ etc.$\}$ — corresponding to a set of nondegenerate ev's a_i, b_j, $c_k, \ldots$ — each contain all the information we can have about such a state, and they can be used to form a complete, orthonormal basis for all the states of our quantum system.

6.) **The Uncertainty Principle**

Let A and B be two hermitian operators corresponding to observables $\mathcal{A}$ and $\mathcal{B}$ of a quantum system. Suppose the system is in the (normalized) state $|\psi\rangle$. The spread in measurements of A and B are, as usual, $\Delta A = \sqrt{\langle A^2\rangle_\psi - \langle A\rangle_\psi^2}$ and $\Delta B = \sqrt{\langle B^2\rangle_\psi - \langle B\rangle_\psi^2}$. We can say that ΔA and ΔB are the *uncertainty* in a measurement of A and B. Then, the "uncertainty principle" is

$$\Delta A\, \Delta B \geq \tfrac{1}{2}|\langle\psi|[A, B]|\psi\rangle|\,. \tag{35}$$

N.B.: If A and B are hermitian, then their commutator is *antihermitian*, $[A, B]^\dagger = -[A, B]$, and its expectation value in $|\psi\rangle$ is imaginary or zero.

It is sometimes said that this product of uncertainties refers to measurements made at the same time. That is not necessary; it can just as well refer to measurements made sequentially, e.g., first A, then B (without an intervening measurement to complicate things). Note that if, e.g., $\Delta A = 0$ and $[A, B] \neq 0$, then ΔB must be infinite!

Proof: Define the hermitian operators $\alpha = A - \langle A\rangle_\psi$ and $\beta = B - \langle B\rangle_\psi$. Then $(\Delta A)^2 = \langle\alpha^2\rangle_\psi$ and $(\Delta B)^2 = \langle\beta^2\rangle_\psi$. For the operator $\gamma = \alpha + i\lambda\beta$, where λ is a real number, the expectation value of the positive semi-definite operator $\gamma^\dagger\gamma$ is

$$\langle\gamma^\dagger\gamma\rangle_\psi = (\Delta A)^2 + i\lambda\langle\psi|[A, B]|\psi\rangle + \lambda^2(\Delta B)^2 \geq 0 \tag{36}$$

for every real number λ. The equality holds if and only if there is one real solution. Therefore, the *discriminant* of this quadratic form cannot be positive:

$$\big(i\langle\psi|[A, B]|\psi\rangle\big)^2 - 4(\Delta A)^2(\Delta B)^2 \leq 0. \qquad \text{QED} \tag{37}$$

From the three matrices representing S_x, S_y and S_z in the S_z-basis, we deduce the famous commutation relations of angular momentum:

$$[S_a, S_b] = i\hbar\, \epsilon_{abc}\, S_c \equiv i\hbar \sum_{c=1}^{3} \epsilon_{abc}\, S_c\,, \tag{38}$$

where $a, b, c = \text{any of(!)}\, 1, 2, 3 = x, y, z$ and ϵ_{abc} is the totally antisymmetric Levi-Civita tensor with $\epsilon_{123} = +1$. In this equation, I employed the "summation convention" that twice-repeated indices are summed over. These commutation relations are independent of the representation of the spin-$\frac{1}{2}$ operators and they are true for <u>*any allowed value*</u> of angular momentum, not just $\frac{1}{2}\hbar$.

The square of the total spin-$\frac{1}{2}$ operator (or the square of the angular momentum operator for any allowed value) is

$$\boldsymbol{S}^2 = S_a S_a \equiv S_1^2 + S_2^2 + S_3^2 = S_x^2 + S_y^2 + S_z^2. \tag{39}$$

Using the angular momentum commutation relations, the very useful identity (prove it!),

$$[A, BC] = [A, B]C + B[A, C], \tag{40}$$

and the antisymmetry of ϵ_{abc}, you can easily prove that $\boldsymbol{S}^2$ and any <u>*single*</u> component of $\boldsymbol{S}$ are compatible hermitian operators:

$$[S_a, \boldsymbol{S}^2] = 0, \qquad (a = 1, 2, 3 = x, y, z). \tag{41}$$

For the three spin- $\frac{1}{2}$ matrices, you can show that

$$\boldsymbol{S}^2 = \begin{pmatrix} \frac{3}{4}\hbar^2 & 0 \\ 0 & \frac{3}{4}\hbar^2 \end{pmatrix} = \tfrac{3}{4}\hbar^2 \mathbf{1}, \tag{42}$$

where $\mathbf{1}$ is the 2×2 unit matrix. It is not difficult to prove that any matrix that commutes with the spin-$\frac{1}{2}$ matrices for S_x, S_y and S_z must be proportional to $\mathbf{1}$.

So, we see that $\boldsymbol{S}^2$ and S_z, say, are compatible operators. Does that mean that the eigenstates of S_z are degenerate? No, because $\boldsymbol{S}^2$ is redundant for a spin-$\frac{1}{2}$ system; it has the same value for $|+\rangle$ and $|-\rangle$ and, therefore for all states in the vector space of a spin-$\frac{1}{2}$ system. (This statement applies only to quantum systems that have only one allowed value of angular momentum.)

A final comment on angular momentum (for now): From the angular momentum commutation relations in Eq. (38) it is possible to derive all <u>allowed</u> values of angular momentum and all allowed values of <u>one</u> of its components. We will do this later in the semester — I hope.

I leave the rest of the discussion of Chapter 2 to the text.