Foundations of Quantum Mechanics

By the mid-1800's, many of the classical macroscopic physical problems of the day had been solved. Advances in photographic technology, largely associated with electronic instruments, led to an increasing body of knowledge about the microscopic world.

As we know, by the early 1900's it was clear that classical mechanics was unable to describe several new phenomena in physics:

photoelectric effect
hydrogen emission spectra } small distances
black-body radiation

speed of light measurements } large velocities.
Maxwell's equations (invariances)

As is usual with the creative process, these problems were solved in a path which, in retrospect, was not entirely clean cut. But the ultimate solutions to the problems ("ultimate" meaning "as we understand them today") came from addressing some very basic questions:

What do we mean by "measurement"? (How are systems prepared? What do we do when we measure them?

What do we mean by "simultaneity"? (How do we synchronize clocks? How do we measure length?

For the next few days we deal with these elementary questions. We see that by dropping our intuitive classical assumptions about the answers to these questions, we are lead to a mechanics formulation which "explains" or "describes" many experimental measurements which classical mechanics failed to explain.

But there are still problems to be solved. Among them:

- the starting point for mechanics is still not unique; we have a number of postulates which we still cannot justify (i.e. they don't spring from logic).
- The numerical work still has an unjustified constant $\hbar$ in it. Why does $\hbar$ have the value that it does?

This course will attempt to:

- show the underlying hypotheses of quantum mechanics
- show the connection between the quantum and classical worlds
- work through a number of examples other than the traditional bound state problems.

Book: Ballentine "Quantum Mechanics" (Prentice-Hall, 1989)
Material to be covered: Chs. 1–11.

Exams
- 1 Midterm 25%
- 1 Final 50%
- Assignments 25%

Mathematical Review

Linear Vector Space

A linear vector space is a set of elements (called vectors) that is closed under
i) addition
ii) multiplication by a scalar.

E.g. Discrete vectors, which may be a column of complex numbers $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}$

E.g. Spaces of functions, say the space of a differentiable functions $e^{i\alpha}$, x^n ...

Characteristics:

- set of vectors $\{\varphi_n\}$ is linearly independent if $\sum_n c_n \varphi_n = 0$ only if $c_n = 0 \ \forall n$.

- dimension of lin. vec. space is equal to the maximum # of linearly independent vectors (which form the basis for the space).

- inner product defined by
  i. $(\varphi, \psi) =$ a complex number
  ii) $(\varphi, \psi) = (\psi, \varphi)^*$
  iii) $(\varphi, \varphi) \geq 0$ $\quad (=0$ only if $\varphi = 0)$
  iv) $(\varphi, c_1\psi_1 + c_2\psi_2) = c_1(\varphi, \psi_1) + c_2(\varphi, \psi_2)$ linear in second argument

$$(c_1\psi_1 + c_2\psi_2, \varphi) = c_1^*(\psi_1, \varphi) + c_2^*(\psi_2, \varphi)$$ antilinear in first argument

From the above:
- Schwartz's inequality $|(\psi, \varphi)|^2 \leq (\psi,\psi)(\varphi,\varphi)$
- Triangle inequality $\|(\psi + \varphi)\| \leq \|\psi\| + \|\varphi\|$ where $\|\varphi\| = (\varphi,\varphi)^{1/2}$.

Corresponding to any linear vector space V is a dual space of linear functionals on V. A linear functional F assigns a scalar $F(\varphi)$ to each vector φ such that

$$F(a\varphi + b\psi) = a F(\varphi) + b F(\psi) \qquad \begin{array}{l} a, b \text{ scalars} \\ \varphi, \psi \text{ vectors.} \end{array}$$

The functionals themselves can form a linear space V′ if their sum is defined as

$$\underbrace{(F_1 + F_2)}_{\text{new functional}}(\varphi) = F_1(\varphi) + F_2(\varphi)$$

Riesz theorem There is a one-to-one correspondence between linear functionals F in V′ and vectors f in V such that

$$F(\varphi) = (f, \varphi)$$

where this is a fixed vector.

(Proof in Ballentine:)

Let ψ be an arbitrary vector in V, which can be expanded in the orthonormal basis set $\{\varphi_n\}$ by

$$\psi = \sum_n c_n \varphi_n .$$

From the def'n of F above:

$$F(\psi) = F\left(\sum_n c_n \varphi_n\right) = \sum_n c_n F(\varphi_n)$$

If we form f from the unique expression [although it depends on].

$$\underset{\text{vector}}{f} = \sum_n \underset{\text{scalars}}{[F(\varphi_n)]^*} \underset{\text{vectors}}{\varphi_n}$$

Then

$$(f, \psi) = \left(\sum_i [F(\varphi_i)]^* \varphi_i , \sum_j c_j \varphi_j\right) = \sum_i \sum_j F(\varphi_i) c_j \overset{\delta_{ij}}{(\varphi_i, \varphi_j)}$$

$$= \sum F(\varphi_i) c_i = F(\sum c_i \varphi_i) = F(\psi) \quad \text{QED.}$$

In Ballentine, pg 3, it is stated that the Schwarz & Triangle inequalities hold if and only if the vectors are linearly independent. I think that the equality holds only if the vectors are ~~positive~~ multiples. (Schwarz)
positive multiples (triangle).

In Margenau & Murphy, The proof of the Schwarz inequality on page 228 goes:

Choose $\gamma = \beta - \frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha$ where $\alpha \neq 0$.

$$\Rightarrow \quad 0 \leq \|\gamma\|^2 = \left(\beta - \frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha \,,\, \beta - \frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha \right)$$

equality only if $\gamma = 0$

$$= (\beta,\beta) - \frac{1}{\|\alpha\|^2}(\beta,\alpha)(\alpha,\beta) + \frac{(\beta,\alpha)^2}{\|\alpha\|^2} + \frac{(\beta,\alpha)^2\alpha^2}{\alpha^2}$$

$$= (\beta,\beta) - \frac{|(\alpha,\beta)|^2}{\|\alpha\|^2}$$

$$\Rightarrow \quad |(\alpha,\beta)|^2 \leq \|\alpha\|^2 \|\beta\|^2$$

equality if $\beta = \frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha$

So, if $\beta = c e^{i\varphi}\alpha$, then $\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha = \frac{c e^{-i\varphi}|\alpha|^2}{\|\alpha\|^2}\alpha = c e^{-i\varphi}\alpha$.

$$\Rightarrow \varphi = 0 \text{ or } \pi$$

For the triangle, if $\beta = -c\alpha$
$\|\alpha+\beta\|^2 = (1-c)^2\alpha^2 < (1+c^2)\alpha^2 = \alpha^2 + \beta^2$.

Dirac Notation

Dirac developed a simple shorthand which is often used (and abused) in physics. Represent the vectors φ as ket vectors $|\varphi\rangle$. The linear functionals are represented as bra vectors $\langle F|$. Then the numerical value of the functional $F(\varphi)$ is written

$$F(\varphi) = \langle F|\varphi\rangle.$$

From the Riesz theorem, $F(\varphi) = (f, \varphi)$, so we could say that $|F\rangle$ is just the ket vector corresponding to f.

In other words, there is more to forming $\langle\varphi|\psi\rangle$ than just multiplying two vectors together; it really is an operation to form a scalar. For continuous vectors $\varphi(x)$:

$$(\varphi, \psi) = \int \varphi^*(x)\psi(x)\, \underbrace{w(x)}_{\text{weight function for the integral.}}\, dx$$

Linear Operators

Functionals map vectors onto scalars
Operators map vectors onto vectors $\qquad \varphi = \underbrace{\hat{A}}_{\text{operator}}\psi$

* An operator $\hat{A}$ is fully defined by specifying its action on every vector in the vector space. To say that $\hat{A} = \hat{B}$ (operators) really means $\hat{A}\psi = \hat{B}\psi \;\; \forall\, \psi$, although usually the ψ's are omitted in writing $\hat{A}=\hat{B}$.

A linear operator satisfies

$$\hat{A}\underbrace{(c_1\psi_1 + c_2\psi_2)}_{\text{vector}} = \underbrace{c_1(\hat{A}\psi_1) + c_2(\hat{A}\psi_2)}_{\text{vectors.}}$$

Sums and products of operators: $\qquad (\hat{A}+\hat{B})\psi = \hat{A}\psi + \hat{B}\psi$
$\qquad \hat{A}\hat{B}\psi = \hat{A}(\hat{B}\psi).$

Let us introduce operators into the Dirac notation.

$\hat{A}|\psi\rangle \Rightarrow \hat{A}$ acting on vector ψ, giving another vector.
This new vector can be used to form an inner product $\langle\varphi|\hat{A}|\psi\rangle$.
Note that the notation $\langle\ |\ |\ \rangle$ really does imply the symmetry

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

To prove this, consider the functional approach. The bra vector $\langle\varphi|$ is a linear functional on the set of ket vectors $|\psi\rangle$,

$$F_\varphi(\psi) = (\varphi, \psi)$$

ψ is arbitrary
φ is a particularly chosen vector, which corresponds to F_φ via the Riesz theorem.

(so $\langle\varphi| = F_\varphi(\ )$, in a sense). $\overline{\overline{\text{scalar}}}$.

Suppose we operate on $\langle\varphi|$ with $\hat{A}$, using the representation

$$\hat{A}F_\varphi(\psi) = F_\varphi(\hat{A}\psi) \qquad \forall\psi.$$

This has to operate on a vector

This equation defines a $\underline{\text{NEW}}$ functional $AF_\varphi(\ )$, i.e. it satisfies the usual functional equation $F_\chi(\psi) = (\chi, \psi)$:

$$F_\chi(\ ) = \hat{A}F_\varphi(\ )$$

Since χ is uniquely determined by φ and $\hat{A}$, we define an operator A^+ so that $\chi = \hat{A}^+\varphi$, and

$$\hat{A}F_\varphi = F_{A^+\varphi}$$

Then:

$$= F_\varphi(\hat{A}\psi) \qquad = (\hat{A}^+\varphi, \psi)$$
$$= (\varphi, \hat{A}\psi)$$

$$\longrightarrow (\hat{A}^+\varphi, \psi) = (\varphi, \hat{A}\psi) \quad \forall\varphi,\psi.$$

This is just the usual definition of an adjoint $\hat{A}^+$ of the operator $\hat{A}$

It follows that $\langle\varphi|\hat{A}^\dagger|\psi\rangle^* = \langle\psi|\hat{A}|\varphi\rangle. \quad \forall \varphi, \psi.$

Other interesting properties of adjoint operators are:

$$(c\hat{A})^\dagger = c^* A^\dagger$$
$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger.$$

It is useful to define an operator from the bra (vector) and kets (functional):

$$|\psi\rangle\langle\varphi|$$ ← This is sometimes called an outer product.

It is an operator in the sense that it produces a vector when it operates on a vector:

$$(|\psi\rangle\langle\varphi|)|\lambda\rangle = \underbrace{|\psi\rangle}_{\text{vector}} \underbrace{\langle\varphi|\lambda\rangle}_{\text{scalar}}$$

If $|\psi\rangle\langle\varphi|$ is an operator, what is its adjoint $|\psi\rangle\langle\varphi|^\dagger$?

From above:

$$\langle\alpha| \underbrace{|\psi\rangle\langle\varphi|^\dagger}_{\text{operator}} |\beta\rangle^* = \langle\beta| \underbrace{|\psi\rangle\langle\varphi|}_{\text{operator}} |\alpha\rangle$$
$$= \langle\beta|\psi\rangle\langle\varphi|\alpha\rangle$$
$$= (\langle\beta|\psi\rangle^* \langle\varphi|\alpha\rangle^*)^*$$
$$= (\langle\psi|\beta\rangle\langle\alpha|\varphi\rangle)^*$$
$$= (\langle\alpha|\underbrace{\varphi\rangle\langle\psi}_{\text{operator}}|\beta\rangle)^*.$$

$$\therefore \quad |\psi\rangle\langle\varphi|^\dagger = |\varphi\rangle\langle\psi|$$

It is therefore tempting to write $|\psi\rangle^\dagger = \langle\psi|$, but neither $|\psi\rangle$ nor $|\psi\rangle$ are operators so they don't have an adjoint.

If $|\varphi\rangle$'s are discrete $|\varphi_i\rangle$ and form a complete orthonormal set, then we can write for any vector $|x\rangle$: $\quad |x\rangle = \sum_i c_i |\varphi_i\rangle \Rightarrow \langle\varphi_j|x\rangle = \sum_i c_i \langle\varphi_j|\varphi_i\rangle$

$$\Rightarrow c_i = \langle\varphi_i|x\rangle.$$

$$\therefore \sum_i |\varphi_i\rangle\langle\varphi_i| = \hat{I}.$$

Operators which satisfy $\langle \varphi|\hat{A}|\psi\rangle = \langle \psi|\hat{A}|\varphi\rangle^*$ are said to be self-adjoint or Hermitian. Some properties:

Theorem 1: If $\langle \psi|\hat{A}|\psi\rangle = \langle \psi|\hat{A}|\psi\rangle^*$ $\forall \psi$ then $\langle \varphi_1|\hat{A}|\varphi_2\rangle = \langle \varphi_2|\hat{A}|\varphi_1\rangle^*$ and $\hat{A} = \hat{A}^\dagger$ $\forall \varphi_1, \varphi_2$ (proof in Ballentine).

If the result of operating $\hat{A}$ on $|\varphi\rangle$ yields a scalar "a" times $|\varphi\rangle$,

$$\hat{A}|\varphi\rangle = a|\varphi\rangle$$

then we say $|\varphi\rangle$ is an eigenvector with eigenvalue "a" of the operator $\hat{A}$.
The eigenvalues & the adjoint operator $\hat{A}$ satisfy $\langle\varphi|A^\dagger = a^*\langle\varphi|$

Theorem 2: If $\hat{A}$ is a Hermitian operator, then all of its eigenvalues are real.

Theorem 3: Eigenvectors corresponding to distinct eigenvalues of a Hermitian operator must be orthogonal.

Application: it is interesting to note that we can define an operator $\hat{A}$ in terms of its eigenvalues and eigenvectors: $\hat{A}|\varphi_i\rangle = a_i|\varphi_i\rangle$. If $|\varphi_i\rangle$ form a complete set, then

$$\hat{A} = \sum_i a_i \underbrace{|\varphi_i\rangle\langle\varphi_i|}_{\text{operator.}}$$

So that

$$\hat{A}|\varphi_j\rangle = \sum_i a_i |\varphi_i\rangle\langle\varphi_i|\varphi_j\rangle = a_j|\varphi_j\rangle \text{ as required.}$$

We can even form functions of operators in the same way:

$$f(\hat{A}) = \sum_i f(a_i)|\varphi_i\rangle\langle\varphi_i|$$

Spectral Theorem (Theorem 4).

The outer product $\hat{P}_i = |\varphi_i\rangle\langle\varphi_i|$ is an example of a projection operator: it projects out one component of a basis set and annihilates all others. A projection operator satisfies

$$\hat{P}_i^2 = \hat{P}_i \qquad \left(\begin{aligned} &\text{or, here} \\ &(|\varphi_i\rangle\langle\varphi_i|)^2 = |\varphi_i\rangle\langle\varphi_i|\varphi_i\rangle\langle\varphi_i| \\ &\qquad = |\varphi_i\rangle\langle\varphi_i| \end{aligned}\right)$$

and

$$\sum_i \hat{P}_i = \hat{I}$$

The spectral theorem states that to each self-adjoint operator $\hat{A}$ there corresponds a unique family of projection operators $E(\lambda)$ for real λ (where λ is an eigenvalue):

$E(\lambda)$ has the properties:

a) if $\lambda_1 < \lambda_2$, $E(\lambda_1)E(\lambda_2) = E(\lambda_2)E(\lambda_1) = E(\lambda_1)$.
(i.e. it projects out the lower eigenvalue)

b) if $\epsilon > 0$, then $E(\lambda+\epsilon)|\psi\rangle \to E(\lambda)|\psi\rangle$ as $\epsilon \to 0$

c) $E(\lambda)|\psi\rangle \to 0$ as $\lambda \to -\infty$

d) $E(\lambda)|\psi\rangle \to |\psi\rangle$ as $\lambda \to +\infty$

e) $\int_{-\infty}^{\infty} \lambda \, dE(\lambda) = \hat{A}$

This integral is defined by Eq. 1-35 in Ballentine

One can show that $E(\lambda)$ has the formal representation

$$E(\lambda) = \int_{-\infty}^{\lambda} |q\rangle\langle q| \, dq \qquad \left(\text{like } P_i = |\varphi_i\rangle\langle\varphi_i|\right)$$

Theorem 5 If $\hat{A}$ and $\hat{B}$ are self-adjoint operators, each of which possesses a complete set of eigenvectors, and if $\hat{A}\hat{B} = \hat{B}\hat{A}$, then there exists a complete set of vectors which are eigenvectors of both $\hat{A}$ & $\hat{B}$

Theorem 6

Any operator that commutes with all members of a complete commuting set of operators must be a function of the operators in that set.