Interpretation of Quantum Theory **U. Waterloo PHY773, Winter 2005**
J. Emerson and R. Laflamme (Instructors) **M. Silva (Scribe)**
Postulates of Quantum Mechanics **Lectures 1 and 2**

Postulates of Standard Quantum Mechanics

Lecturer: Joseph Emerson

Lecture 1 Jan 4, 2005

Postulate I. A physical configuration (state) is described by a state operator that is non-negative (Hermitian) operator with unit trace. Rank-one projectors, $\hat{\rho} = |\psi\rangle\langle\psi|$, called pure states, correspond to states of maximal knowledge.

In many applications, it is adequate to specify the quantum state using only vectors $|\psi\rangle$, where these vectors are elements of a Hilbert space. A Hilbert space $\mathcal{H}$ is a linear vector space with an inner product defined on it, $(\psi, \phi) \in \mathbb{C}$, or in Dirac notation, $\langle\psi|\phi\rangle \in \mathbb{C}$. (We will see later that for a vector space to qualify as an infinite dimensional Hilbert space we must specify a further condition.)

The dimension of $\mathcal{H}$ is the maximum number of linearly independent vectors.

A linearly independent set of vectors spanning $\mathcal{H}$ is called a basis.

Any vector can be expressed as a linear combination of basis vectors, e.g., let $\{\phi_j\}$ be a basis of $\mathcal{H}$, $|\psi\rangle \in \mathcal{H}$, then $|\psi\rangle = \sum_j c_j |\phi_j\rangle$.

Example 1. *A linearly independent set of column vectors form a basis for a discrete Hilbert space.*

Example 2. *The space of differentiable functions can form a Hilbert space.*

An *inner product* is defined by the properties:

i) $(\psi, \phi) \in \mathbb{C}$
ii) $(\phi, \psi) = (\psi, \phi)^\star$ ($\star$ denotes complex conjugation)
iii) $(\phi, c_1\psi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2)$
iv) $||\psi||^2 = (\psi, \psi) >= 0$

In Dirac's notation (i) takes the form: $\langle\psi|\phi\rangle \in \mathbb{C}$

An orthonormal basis $\{\phi_j\}$ has

$$(\phi_j, \phi_i) = \langle\phi_j|\phi_i\rangle = \delta_{ij} \tag{1}$$

where δ_{ij} is the Kronecker delta-function.

Example 3. *For column vectors with elements,*

$$|\phi\rangle \equiv \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix}$$

$$|\psi\rangle \equiv \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix}$$

the inner product is expressed as follows:

$$\langle\psi|\phi\rangle = \sum_j b_j^\star a_j$$

Note that bra vectors (e.g. $\langle\psi|$) are elements of a dual space $\mathcal{H}^\dagger$, which consists of linear functionals mapping elements of the Hilbert space to complex scalars.

Example 4. *Let $\psi(x)$ and $\phi(x)$ be complex functions, then the inner product takes the form: $\langle\psi|\phi\rangle = \int d\mu(x)\psi^\star(x)\phi(x)$*

An infinite dimensional $\mathcal{H}$ has to be complete in the norm – that is, all vectors obtained from limits of Cauchy sequences are contained in $\mathcal{H}$. Given a Cauchy sequence $\{\psi_m\}$, $||\psi_m\rangle - |\psi_n\rangle| \to 0$ as $m, n \to \infty$, $|\psi\rangle = lim_{m\to\infty}|\psi_m\rangle \in \mathcal{H}$, and $\|\psi\|^2 < \infty$.

An important case of a Hilbert space is $L^2(a,b)$, that is, the set of square integrable complex functions, or

$$\int_a^b dx|\psi(x)|^2 < \infty. \tag{2}$$

In practice it is convenient to make use of non-square integrable and generalized functions which do not fit in the Hilbert space framework, for example,

$$\psi(x) = \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(-i\frac{px}{\hbar}\right). \tag{3}$$

and the Dirac 'delta-function':

$$\delta(x - x_o) = \langle x|x_o\rangle, \tag{4}$$

defined by the conditions,

$$\begin{aligned} \int dx\delta(x - x_o) &= 1 \\ \int dx\delta(x - x_o)f(x) &= f(x_o) \end{aligned}$$

To accommodate these elements we can use the *rigged Hilbert space* formalism and treat the following inner products as well-defined:

$$\langle x|x_0\rangle = \delta(x - x_0) \tag{5}$$
$$\langle p|p_0\rangle = \delta(p - p_0) \tag{6}$$

A state operator $\hat{\rho}$ must be non-negative. An operator is non-negative iff $\langle\mu|\hat{\rho}|\mu\rangle \geq 0$ for all $|\mu\rangle \in \mathcal{H}$.

Note that this guarantees that the state operator $\hat{A}$ is Hermitian ($\hat{A} = \hat{A}^\dagger$), where the adjoint operation will be defined later.

For a normalized state operator $\hat{\rho}$ ($\mathrm{tr}(\hat{\rho}) = 1$) there are three equivalent definition of purity (maximal knowledge):

1) $\hat{\rho}^2 = \hat{\rho}$, which means that ρ is projector.
2) $\mathrm{tr}(\hat{\rho}^2) = 1$.
3) $\hat{\rho} = |\psi\langle\rangle\psi|$, defining a projector onto a one-dimensional subspace of $\mathcal{H}$.

The *trace* of an operator $\hat{A}$ is defined by

$$\mathrm{tr}(\hat{A}) = \sum_j \langle \phi_j | \hat{A} | \phi_j \rangle \tag{7}$$

where $\{|\phi_j\rangle\}$ is any (convenient) normalized orthogonal basis.

Postulate II. Each physical observable is represented by a Hermitian operator $\hat{O}$. Let $\hat{O}$ be a Hermitian operator with eigenvalues λ_l and eigenvectors $|\lambda_l\rangle$.

a) **The set of possible observable outcomes is determined from $\{\lambda_l\}$.**
b) **The probability of outcome λ_l, is given by $\mathrm{Pr}(\lambda_l) = \mathrm{tr}(|\lambda_l\rangle\langle\lambda_l|\hat{\rho})$ when the physical configuration is described by $\hat{\rho}$.**

Postulate 2.a) is responsible for the novel structural aspects of quantum theory. Operators with discrete spectra are ”quantized” (in the sense that they are discretized). Examples of this are the atomic energy levels, angular momentum, and electromagnetic radiation can only exchange discrete amounts of energy with some systems (i.e. ”photons”).

Postulate 2.b) provides the statistical/probabilistic/indeterministic character of quantum predictions. It is known as **the Born rule**.

Lecture 2 Jan 6, 2005

Hermitian operators are linear operators, i.e. $\hat{O}(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1(\hat{O}|\psi_1\rangle) + c_2(\hat{O}|\psi_2\rangle)$, that are self-adjoint $\hat{A}^\dagger = \hat{A}$. The adjoint operator $\hat{A}^\dagger$ is defined by the condition $(\hat{A}^\dagger\phi, \psi) = (\phi, A\psi)$ for all $\psi, \phi \in \mathcal{H}$. (In Dirac notation this definition is expressed as: $\langle\phi|\hat{A}^\dagger|\psi\rangle = (\langle\psi|\hat{A}|\phi\rangle)^*$.)

Of course operators do not necessarily commute, i.e., it maybe that $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0$, where we have just defined the *commutator* of operators A and B.

Hermitian operators are always diagonalizable, with eigenvectors and eigenvalues defined by the condition, $\hat{O}|\lambda_l\rangle = \lambda_l|\lambda_l\rangle$.

An important representation of a Hermitian operator is its *spectral decomposition*. In the case of a discrete spectrum this is,

$$\hat{O} = \sum_l \lambda_l |\lambda_l\rangle\langle\lambda_l|$$

where the projectors are $P_l = |\lambda_l\rangle\langle\lambda_l|$. For a continuous spectrum this is,

$$\hat{O} = \int d\lambda \lambda |\lambda\rangle\langle\lambda|$$

where $|\lambda\rangle$ may not be a Hilbert space vector but an element of a rigged Hilbert space (following Dirac's (1930) approach). An alternate approach is the generalized spectral decomposition proposed by von Neumann (1932).

The spectral decomposition allows us to define unambiguously a function of an operator

$$f(\hat{O}) = \sum_\lambda f(\lambda)|\lambda\rangle\langle\lambda|. \tag{8}$$

We can also compute expectation of an observable given by an operator via

$$\langle\hat{O}\rangle = \int d\lambda \langle\psi|\hat{O}|\psi\rangle = \mathrm{tr}(\hat{O}|\psi\rangle\langle\psi|). \tag{9}$$

The spectral decomposition is useful also for calculating probabilities according to the Born rule,

$$\mathrm{Pr}(\lambda \in [a,b]) = \mathrm{tr}\left[\int_a^b d\lambda |\lambda\rangle\langle\lambda||\psi\rangle\langle\psi|\right]. \tag{10}$$

Another useful result is the *(spectral) resolution of the identity*,

$$\hat{1} = \sum_l |\lambda_l\rangle\langle\lambda_l|. \tag{11}$$

Postulate III.

Dynamical transformations are generated by unitary operators,

$$\begin{aligned} |\psi(s_2)\rangle &= U(s_2, s_2)|\psi(s_1)\rangle & (12)\\ \hat{\rho}(s_2) &= U(s_2, s_1)\hat{\rho}(s_1)U^\dagger(s_2, s_1) & (13) \end{aligned}$$

By definition, an operator $\hat{U}$ is unitary if $\hat{U}\hat{U}^\dagger = \hat{1}$, which guarantees it can always be expressed in the form $\hat{U}(s) = \exp(is\hat{A})$, where $\hat{A}$ is a Hermitian operator.

Historically, Dirac identified the operator algebras associated with position, momentum and angular momentum by quantizing the classical Poisson bracket. This is not an ideal approach since we want quantum mechanics to be an independent theory based on its own set of postulates, and not dependent on a classical theory for its formulation.

T. Jordan proposed a different approach (Jordan, 1975) where the unitary transformation and the associated algebra of the generators $\{\hat{A}\}$ are identified from the considering the properties of the continuous Galilean space-time symmetries (generating displacements in position, rotations, and time-translations). Thus one is able to deduce that

$$\begin{aligned} \text{position translation} &\rightarrow \exp(-ia\hat{p}) \rightarrow [\hat{x}, \hat{p}] = iC & (14)\\ \text{rotation about } \hat{n} &\rightarrow \exp(-i\theta\hat{n}\cdot\hat{J}) \rightarrow [J_i, J_j] = iCJ_k = -[J_j, J_i] & (15) \end{aligned}$$

where C is an unspecified constant, that needs to be determined experimentally – this is the fundamental constant we call $\hbar$. From the time translation we deduce *Schrödinger's equation*,

$$\begin{aligned} |\psi(t_2)\rangle &= \hat{U}(t_2, t_1)|\psi(t_1)\rangle & (16)\\ \hat{H} &= i\hbar\left(\frac{d}{dt}\hat{U}\right)U^\dagger & (17)\\ \hat{H} &= -i\hbar\hat{U}\left(\frac{d}{dt}\hat{U}\right)^\dagger & (18)\\ \text{thus}\quad i\hbar\frac{d}{dt}|\psi(t)\rangle &= \hat{H}|\psi(t)\rangle. & (19) \end{aligned}$$

See Peres (1995) and Ballentine (1998).

We can view transformations also in the Heisenberg picture, where observables evolve instead of the state, $\hat{A}(t_2) = \hat{U}^\dagger(t_2, t_1)\hat{A}(t_1)\hat{U}(t_2, t_1)$.

Uncertainty Principle

The *uncertainty principle* due to Heisenberg (1925) states that at best

$$\delta x\delta p \simeq \hbar$$

where δx denotes the resolution for determining x for an individual system in a single experimental trial. There is, in principle, no limitation to how small the resolution can be for determining either x or p, but not both in the same setup. This result is motivated by *the Heisenberg microscope* which shows that simultaneous determination of position and momentum for a single particle is limited because of Einstein's relation $E = h\nu = \hbar\omega$ for the light as probe. The wave-particle duality for the light quanta places a limit on the resolution-disturbance tradeoff, as opposed to the classical picture where the energy of the light probe is proportional to intensity of the wave and there is no limit on the trade-off.

Heisenberg's uncertainty principle is often confused with Robertson's uncertainty principle. Define,

$$\Delta\hat{A}^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$$

to be the variance over an ensemble of measurement outcomes for the observable A. Then we can prove the generalized uncertainty inequality (Robertson, Phys. Rev 34, 163, 1929),

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|,$$

in which the uncertainty is non-zero if the two observables do not commute. In the case of position and momentum, this means we get $\Delta\hat{x}\Delta\hat{p} \geq \frac{\hbar}{2}$. This result is mathematically rigorous, unlike Heisenberg's result, and conceptually the two are quite distinct. One refers to simultaneous measurements on a single system, whereas the other refers to variances of statistics for ensembles of measurements where only one operator is measured on each individual syste. Note that normally different experimental set-ups are required for measuring complementary observables.

These uncertainty relations form the basis for the idea of complementarity introduced by Bohr, as will be discussed later.

Mixed States

We have so far only considered pure states. A general state operator $\hat{\rho}$ can be defined by the conditions:

i) $\mathrm{tr}(\hat{\rho}) = 1$.
ii) $\langle u|\hat{\rho}|u\rangle \geq 0$ for all $|u\rangle \in \mathcal{H}$.

A state for which $\mathrm{tr}(\hat{\rho}^2) < 1$ is called a *mixed state*. For a finite dimensional system, the purity is bounded by $\frac{1}{\dim(\mathcal{H})} \leq \mathrm{tr}(\hat{\rho}^2) \leq 1$.

A mixed state can be 'created' in two ways: (i) as a convex combination of pure states due to classical ignorance (this is called *a proper mixture*), or (ii) by ignoring any quantum correlations between the system and some ancillary system, leaving only partial information about the system state (this is called *an improper mixture*).

Note that there is an infinite number of different convex combinations of pure states for the same mixed state (this is called *the ambiguity of mixtures*.)