6.453 Quantum Optical Communication
Spring 2009

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.453 Quantum Optical Communication

Lecture Number 2
Fall 2008

Jeffrey H. Shapiro

Date: Tuesday, September 9, 2008

Dirac-notation Quantum Mechanics.

Introduction

Last time you were introduced to—teased with, really—three examples of how quantum optical communication has distinctly non-classical features: quadrature noise squeezing, polarization entanglement, and teleportation. In this lecture, we begin laying the foundation for understanding all three of these phenomena, and more. Our task is to present the essentials of Dirac-notation quantum mechanics. No prior acquaintance with this material is assumed. There are three fundamental notions that we must establish: state, time evolution of the state, and measurements. The first two will be completed in this lecture; the last will spill over into Lecture 3. Moreover, although these three concepts are easily stated, they will be accompanied by a variety of notational and mathematical details that will comprise most of today's lecture.

Quantum Systems and Quantum States

Slide 3 defines a quantum system and the state of a quantum system. The first definition—that of a quantum system—requires no explanation. There are several points to be made, however, about the definition of the *state* of a quantum system. First, let us remember what it means to be the state of a classical system. We'll do so by means of two examples from classical physics, one from mechanics, and one from circuit theory. After that, we'll review—and perhaps extend—what you know about vector spaces and linear operations on vectors. Here we will use the Dirac notation, but we also exhibit two special cases that will help illustrate the points being made.

The State of a Point Mass

The state, at time t_0, of an m-kg point mass that is moving in three-dimensional space under the influence of an applied force is its position, $\vec{r}(t_0)$, and its momentum, $\vec{p}(t_0)$. The state contains all information about the behavior of the mass prior to time

t_0 that is relevant to predicting its behavior for $t > t_0$. In particular, if the applied force, $\vec{F}(t)$, is known for $t_0 \le t \le t_1$, then $\vec{x}(t_1)$ and $\vec{p}(t_1)$ can be found by solving

$$\frac{\mathrm{d}\vec{p}(t)}{\mathrm{d}t} = \vec{F}(t) \quad \text{and} \quad m\frac{\mathrm{d}\vec{r}(t)}{\mathrm{d}t} = \vec{p}(t), \quad \text{for } t_0 \le t \le t_1, \tag{1}$$

subject to the initial conditions that the position and momentum at time t_0 be $\vec{x}(t_0)$ and $\vec{p}(t_0)$, respectively.

The State of an RLC Circuit

Consider the parallel RLC circuit shown in Fig. 1. The state of this circuit at time $t = t_0$ can be taken to be the charge on its capacitor, $Q(t) = Cv(t)$, and the flux through its inductor, $\Phi(t) = Li_L(t)$.[1]

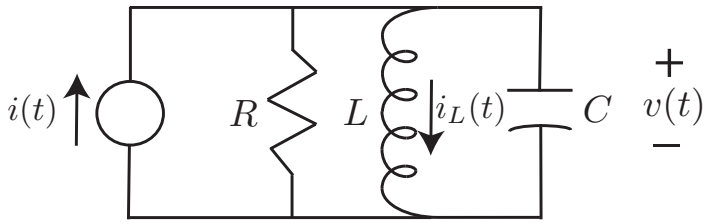

Figure 1: The state of this parallel RLC circuit can be taken to be the charge on its capacitor, $Q(t) = Cv(t)$, and the flux through its inductor, $\Phi(t) = Li_L(t)$.

To find the state at some later time, we can use Kirchhoff's current law and Kirchhoff's voltage law—plus the v-i relations for the three circuit elements—to show that

$$RLC\frac{\mathrm{d}^2v(t)}{\mathrm{d}t^2} + L\frac{\mathrm{d}v(t)}{\mathrm{d}t} + Rv(t) = RL\frac{\mathrm{d}i(t)}{\mathrm{d}t}, \quad \text{for } t \ge t_0, \tag{2}$$

which can be solved, given $i(t)$ for $t_0 \le t \le t_1$ and the initial conditions

$$v(t_0) = \frac{Q(t_0)}{C} \quad \text{and} \quad \left.\frac{\mathrm{d}v(t)}{\mathrm{d}t}\right|_{t=t_0} = \frac{i(t_0)}{C} - \frac{v(t_0)}{RC} - \frac{\Phi(t_0)}{LC}, \tag{3}$$

to obtain $v(t_1)$ and $\mathrm{d}v(t)/\mathrm{d}t|_{t=t_1}$. These, in turn, allow us to find

$$Q(t_1) = Cv(t_1) \quad \text{and} \quad \Phi(t_1) = Li_L(t_1) = Li(t_1) - L\frac{v(t_1)}{R} - LC\left.\frac{\mathrm{d}v(t)}{\mathrm{d}t}\right|_{t=t_1}, \tag{4}$$

proving that knowledge of $\{Q(t_0), \Phi(t_0)\}$ and $\{i(t) : t_0 \le t \le t_1\}$ is sufficient to determine $\{Q(t_1), \Phi(t_1)\}$.

[1]Because C and L are known constants, it is equivalent to say that $v(t_0)$ and $i_L(t_0)$ comprise the state at time t_0. Alternatively, we can take $v(t_0)$ and $\mathrm{d}v(t)/\mathrm{d}t|_{t=t_0}$ to be the state.

Vector Spaces

A vector space is a set of elements (vectors), which we'll denote $\{|\cdot\rangle\}$, and complex numbers (scalars) with vector addition and scalar multiplication defined and obeying:

- Vector addition is *closed*. If $|x\rangle$ and $|y\rangle$ are elements of a vector space, then so too is $|x+y\rangle \equiv |x\rangle + |y\rangle$.
- Vector addition is *commutative*: $|x\rangle + |y\rangle = |y\rangle + |x\rangle$.
- Vector addition is *associative*: $(|x\rangle + |y\rangle) + |z\rangle = |x\rangle + (|y\rangle + |z\rangle)$.
- There exists an identity element, $|0_a\rangle$, such that $|x\rangle + |0_a\rangle = |x\rangle$.
- There exists an additive inverse element, $|{-}x\rangle$, such that $|x\rangle + |{-}x\rangle = |0_a\rangle$.
- Scalar multiplication is *closed*. If $|x\rangle$ is a vector and c is a scalar, then $|cx\rangle \equiv c|x\rangle$ is also a vector.
- Scalar multiplication is *distributive*: $(c_1{+}c_2)|x\rangle = c_1|x\rangle{+}c_2|x\rangle$, and $c(|x\rangle{+}|y\rangle) = c|x\rangle + c|y\rangle$.
- There is an identity scalar, 1, such that $1|x\rangle = |x\rangle$.
- There is a zero scalar, 0, such that $0|x\rangle = |0_a\rangle$.

As we progress through this lecture's general mathematical development, we shall carry along the two running examples that we now introduce.

Example 1: N-D Real Euclidean Space
The elements of N-D real Euclidean space, $\mathcal{R}^N$, are conveniently represented as column vectors,

$$|x\rangle = \mathbf{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \tag{5}$$

where the $\{x_n\}$ and the scalars are real numbers. That the preceding vector space properties are satisfied by $\mathcal{R}^N$ should be familiar to you from your linear algebra prerequisite for 6.453.

Example 2: Complex-valued, Square-integrable Time Functions on $[0,T]$
The complex-valued, square-integrable time functions, $|x\rangle = \{x(t) : 0 \le t \le T\}$, form a vector space $\mathcal{L}_2[0,T]$. Here, by square-integrable, we mean that

$$\int_0^T \mathrm{d}t\, |x(t)|^2 < \infty. \tag{6}$$

You should verify that $\mathcal{L}_2[0,T]$ has the properties we have listed for a vector space.

Inner Product Spaces

An inner product space is a vector space on which an inner product (dot product) is defined. If $|x\rangle$ and $|y\rangle$ are elements of an inner product space, their inner product, denoted $\langle x|y\rangle$ is a complex number. In Dirac terminology, $|x\rangle$ is a ket vector, and $\langle x|$, which is the adjoint of this ket, is called a bra vector. The bra $\langle x|$ and the ket $|y\rangle$ then form a bra-ket, which is the inner product $\langle x|y\rangle$ of the vectors $|x\rangle$ and $|y\rangle$. Inner products have the following properties.

- Inner products are conjugate symmetric: $\langle x|y\rangle = \langle y|x\rangle^*$.
- If c_1 and c_2 are complex numbers and $|c_1 x + c_2 y\rangle = c_1|x\rangle + c_2|y\rangle$, then $\langle c_1 x + c_2 y|z\rangle = c_1^*\langle x|z\rangle + c_2^*\langle y|z\rangle$.
- The length of a vector $|x\rangle$, given by $\|x\| \equiv \sqrt{\langle x|x\rangle}$, is non-negative and equals zero if and only if $|x\rangle = |0_a\rangle$.
- Inner products satisfy the Schwarz inequality,

$$|\langle x|y\rangle| \leq \sqrt{\langle x|x\rangle\langle y|y\rangle}, \tag{7}$$

  where equality occurs if and only if $|x\rangle = c|y\rangle$ for some scalar c.
- Inner products satisfy the Triangle Inequality,

$$\|x + y\| \leq \|x\| + \|y\|, \tag{8}$$

  where equality occurs if and only $|x\rangle = c|y\rangle$ for some *non-negative* scalar c.

These properties can be illustrated by our two running examples as follows.

Example 1: N-D Real Euclidean Space
The bra vector associated with (5) is its transpose[2]

$$\langle x| = \mathbf{x}^T \equiv \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}, \tag{9}$$

and the inner product between $|x\rangle$ and $|y\rangle$ in $\mathcal{R}^N$ is

$$\langle x|y\rangle = \mathbf{x}^T\mathbf{y} \equiv \sum_{n=1}^{N} x_n y_n. \tag{10}$$

This inner product example and its properties should be familiar from your linear algebra background.

[2]If we had used complex scalars, instead of real scalars, for the elements of $\mathbf{x}$, then its adjoint would have been the conjugate transpose.

Example 2: Complex-valued, Square-integrable Time Functions on $[0, T]$
The bra vector associated with $|x\rangle = \{x(t) : 0 \le t \le T\}$ is $\langle x| = \{x^*(t) : 0 \le t \le T\}$, and the inner product for $x(t)$ and $y(t)$ in $\mathcal{L}_2[0, T]$ is

$$\langle x|y\rangle \equiv \int_0^T \mathrm{d}t\, x^*(t)y(t). \tag{11}$$

You should verify that this definition satisfies the properties we have listed for an inner product. Moreover, the Schwarz inequality,

$$\left|\int_0^T \mathrm{d}t\, x^*(t)y(t)\right|^2 \le \int_0^T \mathrm{d}t\, |x(t)|^2 \int_0^T \mathrm{d}t\, |y(t)|^2, \tag{12}$$

with equality if and only if $x(t) = cy(t)$, should be familiar from your linear systems class.

Hilbert Spaces

A Hilbert space is a complete inner product space. An inner product space is complete if every Cauchy sequence converges. Let $\{|x_n\rangle : 1 \le n < \infty\}$ be a sequence of vectors. This sequence is a Cauchy sequence if and only if for every $\delta > 0$ there is an N such that

$$\|x_n - x_m\| = \sqrt{\langle x_n - x_m | x_n - x_m\rangle} < \delta \quad \text{for all } n, m > N. \tag{13}$$

The sequence $\{|x_n\rangle : 1 \le n < \infty\}$ converges if and only if there is a vector $|x\rangle$ such that for every $\delta > 0$ there is an N such that

$$\|x_n - x\| = \sqrt{\langle x_n - x | x_n - x\rangle} < \delta \quad \text{for all } n > N. \tag{14}$$

All convergent sequences are Cauchy, but the converse need not be true. For example, consider the set of rational numbers, $\{x = p/q : p, q = \text{integers}\}$. A Cauchy sequence of rational numbers may converge to an irrational number, hence the set of rational numbers is not complete. Both of our running examples, $\mathcal{R}^N$ and $\mathcal{L}_2[0, T]$, are complete, and hence their inner product spaces are Hilbert spaces.

Time Evolution

Slide 4 gives the first of our three axioms for quantum mechanics: it specifies how the state of an *isolated* quantum system—one that does *not* interact with an external environment—evolves in time. There, we have stated equivalent formulations for this evolution, one based on a unitary operator and the other based directly on the Schrödinger equation. To establish comfort with the former, let's review some theory for linear operators on vector spaces.

Linear Operators

Let $\mathcal{H}_\mathcal{S}$ be the Hilbert space of states for some quantum system $\mathcal{S}$. An operator, $\hat{T}$, that maps $\mathcal{H}_\mathcal{S}$ into $\mathcal{H}_\mathcal{S}$ has the property that for every $|x\rangle \in \mathcal{H}_\mathcal{S}$ there is some $|y\rangle \in \mathcal{H}_\mathcal{S}$ such that $|y\rangle = \hat{T}|x\rangle$. The operator $\hat{T}$ is a *linear* operator if it obeys the superposition principle, i.e.,

$$\hat{T}(c_1|x\rangle + c_2|y\rangle) = c_1(\hat{T}|x\rangle) + c_2(\hat{T}|y\rangle). \tag{15}$$

At this juncture it is worthwhile to define the *adjoint*, $\hat{T}^\dagger$, of a linear operator of $\hat{T}$. The adjoint operator obeys

$$\langle y|(\hat{T}|x\rangle) = \langle x|(\hat{T}^\dagger|y\rangle)^*, \quad \text{for all } |x\rangle, |y\rangle. \tag{16}$$

Once more, it is worth examining these properties in the context of our two running examples.

Example 1: N-D Real Euclidean Space
A linear operator, $\hat{T}$, that maps $\mathcal{R}^N$ into $\mathcal{R}^N$ is an $N \times N$ matrix of real numbers

$$\hat{T} = \mathbf{T} \equiv \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1N} \\ T_{21} & T_{22} & \cdots & T_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ T_{N1} & T_{N2} & \cdots & T_{NN} \end{bmatrix}, \tag{17}$$

and $\mathbf{y} = |y\rangle = \hat{T}|x\rangle = \mathbf{T}\mathbf{x}$ is found by matrix-vector multiplication,

$$y_n = \sum_{m=1}^{N} T_{nm} x_m. \tag{18}$$

It is now easy to see that the adjoint operator, $\hat{T}^\dagger$, associated with $\hat{T}$ is the transpose of the $\mathbf{T}$ matrix, viz.,

$$\hat{T}^\dagger = \mathbf{T}^T \equiv \begin{bmatrix} T_{11} & T_{21} & \cdots & T_{N1} \\ T_{12} & T_{22} & \cdots & T_{N2} \\ \vdots & \vdots & \vdots & \vdots \\ T_{1N} & T_{2N} & \cdots & T_{NN} \end{bmatrix}. \tag{19}$$

Example 2: Complex-valued, Square-integrable Time Functions on $[0, T]$
A linear operator, $\hat{T}$ that maps $\mathcal{L}_2[0, T]$ into $\mathcal{L}_2[0, T]$ is a complex-valued function of two time variables, $T(t, u)$, and $|y\rangle = \hat{T}|x\rangle$ is found from the superposition integral,

$$y(t) = \int_0^T \mathrm{d}u\, T(t, u) x(u). \tag{20}$$

Here, in order to ensure that $y(t)$ is square integrable, $T(t,u)$ must satisfy a regularity condition, e.g.,

$$\int_0^T \mathrm{d}t \int_0^T \mathrm{d}u \, |T(t,u)|^2 < \infty. \tag{21}$$

The adjoint operator, $\hat{T}^\dagger$, associated with $\hat{T} = T(t,u)$ is $\hat{T}^\dagger = T^*(u,t)$, i.e.,

$$\hat{T}^\dagger |y\rangle = \int_0^T \mathrm{d}u \, T^*(u,t) y(u). \tag{22}$$

In our development and application of Dirac-notation quantum mechanics we will need to know about some special classes of linear operators.

- A linear operator is said to be Hermitian, i.e., self-adjoint, if it satisfies $\hat{T}^\dagger = \hat{T}$.
- The identity operator, $\hat{I}$, has the property that $\hat{I}|x\rangle = |x\rangle$ for all $|x\rangle$.
- The inverse of a linear operator, denoted $\hat{T}^{-1}$, is such that $\hat{T}^{-1}\hat{T} = \hat{T}\hat{T}^{-1} = \hat{I}$. BUT, not all linear operators have inverses.
- A linear operator $\hat{U}$ is unitary if $\hat{U}^{-1} = \hat{U}^\dagger$. Unitary operators have the property that they preserve lengths:

  $$\|\hat{U}|x\rangle\|^2 = (\langle x|\hat{U}^\dagger)(\hat{U}|x\rangle) = \langle x|(\hat{U}^\dagger\hat{U})|x\rangle = \langle x|\hat{I}|x\rangle = \langle x|x\rangle = \|x\|^2. \tag{23}$$

  You can make yourself comfortable with the manipulations performed in these equations by comparing them with the corresponding results for the vector space $\mathcal{R}^N$:

  $$\|\mathbf{U}\mathbf{x}\|^2 = (\mathbf{U}\mathbf{x})^T(\mathbf{U}\mathbf{x}) = \mathbf{x}^T\mathbf{U}^T\mathbf{U}\mathbf{x} = \mathbf{x}^T\mathbf{I}\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2. \tag{24}$$

Unitary operators also preserve inner products, i.e.,

$$(\hat{U}|x\rangle)^\dagger(\hat{U}|y\rangle) = \langle x|(\hat{U}^\dagger\hat{U})|y\rangle = \langle x|y\rangle \quad \text{for all } |x\rangle, |y\rangle. \tag{25}$$

The physical importance of unitary operators in Axiom 1 should now be apparent. A ket that represents a finite-energy state of a quantum system at time t_0 has unit length. If that system is isolated—so that its evolution is unitary—then its state at some later time t_1 will also have unit length. Mathematically, a unitary operation is a rotation of coordinates, perhaps augmented by inverting some of the axes. You should check that in $\mathcal{R}^2$ the operator

$$\mathbf{U} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \tag{26}$$

is both unitary—so that $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix—*and* a rotation of coordinates by θ.

Observables and Quantum Measurements

Slide 5 presents the second and third of our three axioms for quantum mechanics. An *observable*, i.e., a measurable dynamical variable of a quantum system is represented by an Hermitian operator with a complete set of eigenkets. For our classical point mass, observables would include the position and momentum vectors and the energy. For our classical RLC circuit, observables would include all the voltages and currents in the circuit, as well as the the energies stored in the inductor and the capacitor. Before we are ready to make use of these axioms, we should review eigenkets and eigenvalues, both in a general setting and for our two running examples.

Eigenkets and Eigenvalues

Let $\hat{O}$ be an observable. Because $\hat{O}$ is Hermitian, it has eigenkets $\{|o\rangle\}$ and associated eigenvalues $\{o\}$ that obey

$$\hat{O}|o\rangle = o|o\rangle, \tag{27}$$

i.e., the applying the operator to one of its eigenkets results in scalar multiplication—by the associated eigenvalue—of that eigenket. It is conventional to label eigenkets by their associated eigenvalues.

Example 1: N-D Real Euclidean Space
For the vector space $\mathcal{R}^N$, this eigenket-eigenvalue relation becomes

$$\mathbf{O}\mathbf{o} = o\mathbf{o}, \tag{28}$$

which can be rearranged to read

$$(\mathbf{O} - o\mathbf{I})\mathbf{o} = \mathbf{0}, \quad \text{where } \mathbf{I} \text{ is the identity matrix, and } \mathbf{0} \text{ is the zero vector.} \tag{29}$$

Thus, for there to be a non-trivial, $\mathbf{o} \neq \mathbf{0}$, solution, then o must satisfy the characteristic equation

$$\det(\mathbf{O} - o\mathbf{I}) = 0. \tag{30}$$

For $\mathbf{O}$ a real, symmetric matrix, there are N real roots to this equation, although some may be degenerate. Once the eigenvalues have been determined, the eigenkets are found by using those values in the eigenket-eigenvalue relation.

Example 2: Complex-valued, Square-integrable Time Functions on $[0, T]$
For the vector space $\mathcal{L}_2[0, T]$, the eigenket-eigenvalue relation is the Fredholm integral equation

$$\int_0^T \mathrm{d}u\, O(t, u)o(u) = oo(t), \quad \text{for } 0 \leq t \leq T. \tag{31}$$

The identity operator for $\mathcal{L}_2[0, T]$ is the impulse (Dirac delta) function, $\delta(t - u)$, because

$$\int_0^T \mathrm{d}u\, \delta(t - u)x(u) = x(t), \quad \text{for } 0 \leq t \leq T. \tag{32}$$

Here are some fundamental properties of eigenkets and eigenvalues that we shall need and which you will explore on Problem Set 1.

- The eigenvalues are real valued.
- The eigenkets associated with distinct eigenvalues are orthogonal, i.e., if o and o' are distinct eigenvalues of $\hat{O}$, then their associated eigenkets satisfy $\langle o|o'\rangle = 0$.
- Eigenkets can be normalized to have unit length, i.e., we can assume that $\langle o|o\rangle = 1$.
- If there are M linearly independent eigenkets that have the same eigenvalue, then these can be converted into M orthonormal eigenkets that have this eigenvalue.

Outer Product Notation and its Uses

Suppose that $|x\rangle$ and $|y\rangle$ are kets in a Hilbert space $\mathcal{H}_{\mathcal{S}}$. Then it should be self-evident that the *outer* product, $|x\rangle\langle y|$ is a linear operator that maps $\mathcal{H}_{\mathcal{S}}$ into $\mathcal{H}_{\mathcal{S}}$. In particular, for any $|w\rangle, |z\rangle \in \mathcal{H}_{\mathcal{S}}$ and $|c_1 w + c_2 z\rangle \equiv c_1|w\rangle + c_2|z\rangle$ we have that

$$(|x\rangle\langle y|)|c_1 w + c_2 z\rangle = |x\rangle(c_1\langle y|w\rangle + c_2\langle y|z\rangle), \tag{33}$$

where $\langle y|w\rangle$ and $\langle y|z\rangle$ are scalars.

Outer products give us some very useful operator representations. For $\hat{O}$ an observable with a discrete (or even countable) set of orthonormal eigenkets $\{|o_n\rangle\}$ and associated eigenvalues $\{o_n\}$, we have that

$$\hat{O} = \sum_n o_n |o_n\rangle\langle o_n|, \tag{34}$$

as you will show on Problem Set 1. If the eigenkets are complete, then any $|x\rangle \in \mathcal{H}_{\mathcal{S}}$ can be represented as a linear combination of these eigenkets:

$$|x\rangle = \sum_n x_n |o_n\rangle, \tag{35}$$

where the coefficients $\{x_n\}$, depend on $|x\rangle$. Because the eigenkets have been taken to be orthonormal, we have that these coefficients can be found from projection onto the eigenkets:

$$x_n = \langle o_n|x\rangle. \tag{36}$$

It then follows that the eigenkets resolve the identity operator in the sense that

$$\hat{I} = \sum_n |o_n\rangle\langle o_n|, \tag{37}$$

which is something that you will also prove on Problem Set 1. As usual, it's worth grounding our abstract notions by referring them to the running examples of $\mathcal{R}^N$ and $\mathcal{L}_2[0,T]$.

Example 1: N-D Real Euclidean Space
The standard orthonormal basis for $\mathcal{R}^N$ is $\{\mathbf{1}_n : 1 \leq n \leq N\}$, where $\mathbf{1}_n$ has its nth element equal to unity and all others equal to zero. Then, it should be clear that

$$\mathbf{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad \text{has} \quad x_n = \mathbf{1}_n^T \mathbf{x}, \tag{38}$$

and the $N \times N$ identity matrix satisfies

$$\mathbf{I} = \sum_{n=1}^{N} \mathbf{1}_n \mathbf{1}_n^T. \tag{39}$$

Furthermore, if $\mathbf{U}$ is any real-valued, $N \times N$ unitary matrix, then

$$\mathbf{e}_n \equiv \mathbf{U}\mathbf{1}_n \quad \text{for } 1 \leq n \leq N, \tag{40}$$

defines another orthonormal basis for $\mathcal{R}^N$.

Example 2: Complex-valued, Square-integrable Time Functions on $[0,T]$
The complex sinusoids comprise an orthonormal basis for $\mathcal{L}_2[0,T]$, viz.,

$$\phi_n(t) \equiv \frac{e^{j2\pi nt/T}}{\sqrt{T}} \quad \text{for } -\infty < n < \infty, \tag{41}$$

satisfy

$$\int_0^T \mathrm{d}t\, \phi_n^*(t)\phi_m(t) = \delta_{nm} \equiv \begin{cases} 1, & \text{for } n = m \\ 0, & \text{for } n \neq m, \end{cases} \tag{42}$$

and any $x(t) \in \mathcal{L}_2[0,T]$ can be represented in the Fourier series

$$x(t) = \sum_{n=-\infty}^{\infty} x_n \frac{e^{j2\pi nt/T}}{\sqrt{T}}, \quad \text{for } 0 \leq t \leq T, \tag{43}$$

where

$$x_n = \int_0^T \mathrm{d}t\, \phi_n^*(t)x(t) = \frac{1}{\sqrt{T}} \int_0^T \mathrm{d}t\, x(t)e^{-j2\pi nt/T}. \tag{44}$$

We also have that the identity operator for $\mathcal{L}_2[0,T]$ has the following series representation:

$$\delta(t-u) = \sum_{n=-\infty}^{\infty} \phi_n^*(t)\phi_n(u) = \sum_{n=-\infty}^{\infty} \frac{e^{-j2\pi n(t-u)/T}}{T}, \quad \text{for } 0 \leq t, u \leq T. \tag{45}$$

Measurement Statistics

Axioms 3 and 3a point to an essential way in which quantum mechanics diverges from classical physics. When a measurement is made on a classical system whose state is known, then there is no limit to the precision of that measurement, i.e., there is no fundamental requirement that classical measurements be noisy. Such is not the case in quantum mechanics. Even if the state of the system is known, the outcome of measuring an observable is, in general, a random variable. The state of the system and the observable that has been chosen for measurement determine the statistics of the resulting outcome according to the prescription given on Slides 5 and 6, for the cases of countable and uncountable eigenvalues, respectively. In both cases, the measurement outcome will be one of the eigenvalues, and the measurement statistics are obtained by projection of the state onto the associated eigenkets. Because calculating and manipulating the statistics of quantum measurements are so important to what we will cover this semester, we need to take the time now to pin down the fundamental ideas.

Consider an observable $\hat{O}$ with distinct, discrete eigenvalues $\{o_n\}$. The associated orthonormal eigenkets, $\{|o_n\rangle\}$, obey

$$\langle o_n|o_m\rangle = \delta_{nm}. \tag{46}$$

If we measure this observable when the system is in state $|\psi\rangle$, then Axiom 3 states that the outcome o_n will occur with the following probability

$$\Pr(o_n) = |\langle o_n|\psi\rangle|^2, \tag{47}$$

Let's see that this axiom consistent with probability theory, which holds that probabilities must lie between 0 and 1, and that summing the probabilities of all possible disjoint outcomes must equal 1. That the probabilities in (47) are non-negative follows immediately from its right-hand side being the squared magnitude of an inner product. That these probabilities do not exceed 1 follows from the Schwarz inequality,

$$|\langle o_n|\psi\rangle|^2 \leq \langle o_n|o_n\rangle\langle\psi|\psi\rangle, \tag{48}$$

and $|o_n\rangle, |\psi\rangle$ both having unit length. To show that the total probability is 1, we argue as follows:

$$\sum_n \Pr(o_n) = \sum_n |\langle o_n|\psi\rangle|^2 = \sum_n \langle\psi|o_n\rangle\langle o_n|\psi\rangle = \langle\psi| \left(\sum_n |o_n\rangle\langle o_n|\right) |\psi\rangle \tag{49}$$

$$= \langle\psi|\hat{I}|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{50}$$

The situation is more complicated for observables that have a continuum of eigenvalues.[3] To see what is involved, let $\hat{O}$ be an observable whose eigenvalues are

[3]Examples include position and momentum, and, as we will see in considerable depth later, the quadrature components of the electromagnetic field.

$\{o : -\infty < o < \infty\}$ and non-degenerate. The eigenket-eigenvalue relation now leads to *infinite-length* eigenkets that satisfy the orthonormality relation

$$\langle o|o'\rangle = \delta(o - o'), \tag{51}$$

from which it follows that

$$\hat{O} = \int_{-\infty}^{\infty} \mathrm{d}o\, o|o\rangle\langle o|, \tag{52}$$

and

$$\hat{I} = \int_{-\infty}^{\infty} \mathrm{d}o\, |o\rangle\langle o|, \tag{53}$$

so that for any $|x\rangle \in \mathcal{H}_{\mathcal{S}}$ we get

$$|x\rangle = \int_{-\infty}^{\infty} \mathrm{d}o\, x(o)|o\rangle, \quad \text{with} \quad x(o) \equiv \int_{-\infty}^{\infty} \mathrm{d}o\, \langle o|x\rangle. \tag{54}$$

According to Axiom 3a, when we measure this observable, with the system being in state $|\psi\rangle$, the probability *density* for getting the value o is

$$p(o) = |\langle o|\psi\rangle|^2, \quad \text{for } -\infty < o < \infty. \tag{55}$$

To check that this probability density specification is consistent with classical probability theory, we note that $p(o) \geq 0$ and

$$\int_{-\infty}^{\infty} \mathrm{d}o\, p(o) = \int_{-\infty}^{\infty} \mathrm{d}o\, |\langle o|\psi\rangle|^2 = \int_{-\infty}^{\infty} \mathrm{d}o\, \langle\psi|o\rangle\langle o|\psi\rangle \tag{56}$$

$$= \langle\psi \left(\int_{-\infty}^{\infty} \mathrm{d}o\, |o\rangle\langle o| \right) |\psi\rangle = \langle\psi|\hat{I}|\psi\rangle = \langle\psi|\psi\rangle = 1, \tag{57}$$

and consistency is proven.

A few final comments and we will be done for today. First, we note that if—and only if—the quantum state is an eigenket of the observable that is measured do we get a *non-random* outcome. in particular, for an observable with discrete eigenvalues, if $|\psi\rangle = |o_m\rangle$, then Axiom 3 shows that

$$\Pr(o_n) = |\langle o_n|\psi\rangle|^2 = |\langle o_n|o_m\rangle|^2 = |\delta_{nm}|^2 = \delta_{nm}. \tag{58}$$

Now, because the eigenkets of an observable whose eigenvalues form a continuum are of infinite length, it is impossible to avoid randomness when this observable is measured and the state of the system has finite energy. Finally, we point to the Projection Postulate, which appears on Slide 6. This postulate, which provides the means for studying what happens to a quantum system *after* a measurement has been made, is found in standard introductions to quantum mechanics. However, it will *not* play

a strong role in our work, because we will be concerned with photodetection measurements, which are invariably annihilative, i.e., the photons—which comprised the quantized light field that was measured—get destroyed in the measurement process. Nevertheless, the Projection Postulate is worth citing because—for those quantum systems and measurements to which it applies—it makes clear why we cannot use repeated measurements to circumvent the fundamental randomness inherent in Axioms 3 and 3a.

The Road Ahead

Believe it or not, you have now seen almost all of the foundations of quantum mechanics that we will need for the entire semester. Next lecture we complete this foundational work by: continuing our work on quantum measurements; converting the Schrödinger picture of quantum mechanics—which is what we have been doing so far—into the equivalent (but more convenient for quantum optics) Heisenberg picture; and deriving the Heisenberg uncertainty principle. After that, we will be ready to tackle the quantum harmonic oscillator, which we will later learn can represent a single mode of the electromagnetic field.