NOTES ON HILBERT SPACE
PHYSICS 141 (2003)

I. Hilbert Space

DEFINITION: A *Hilbert space* is an inner product space which, as a metric space, is complete.

We will not present an exhaustive "mathematical" discussion of this subject. Rather, by using examples and analogies, hopefully you will feel more "at ease" with "Hilbert space" at the end of this short discussion.

REMARKS:

(1) Consider the space of functions, $\{f : [a, b] \to \mathcal{C}\}$, where f is a "square-integrable" complex-valued function on the real interval $[a, b]$, *i.e.*,

$$\int_a^b dx |f(x)|^2 < \infty. \tag{I.1}$$

This space will be denoted as L_2. One can directly verify that L_2 is a complex vector space, *i.e.*, if f and $g \in L_2$, then $\alpha f + \beta g \in L_2$ for α and β complex, *etc.* We will return to clarify in what sense one can "visualize a function as a vector".

(2) By introducing a "dot-product",

$$\langle f|g\rangle \equiv \int_a^b dx f^*(x) g(x), \tag{I.2}$$

this complex vector space becomes an **inner product space**. This inner product provides us with a positive definite "norm" for each vector,

$$||f|| \equiv [\langle f|f\rangle]^{1/2}. \tag{I.3}$$

($||f|| = 0$ if and only if $f = 0$.)

(3) Define a "distance" between two functions by $d(f, g) \equiv ||f - g||$. This turns L_2 into a **metric space**. We shall explain the fact that L_2 is a **Hilbert space** next.

II. Hilbert Space is Complete

COMMENTS:

(1) In dealing with real numbers, it is natural to first start working with integers. One then finds that one needs rationals, and finally irrationals. To formally introduce "irrationals", one can first introduce a notion of distance, $d(x, y)$, between two numbers x and y, which of course can be chosen to be the absolute value $|x - y|$. (With a notion of distance, the space of numbers becomes a "metric space".) Now one can talk about an infinite sequence of numbers, *e.g.*, $\{3, 3.1, 3.14, 3.141, 3.1415, \cdots\}$. Note that (i) the differences between any pairs of numbers further down the sequence become smaller and smaller. (This is an example of a "Cauchy sequence".) (ii) As finite decimals, every term in the sequence is a rational number. Now, as you can guess, the limit of this sequence is π, and it is **not** a rational number.

(2) A metric space is **not complete** if it contains Cauchy sequences whose limit points are not contained in the space. As the example above shows, the space of rational numbers, with the usual notion of distance, is **not** a complete metric space. By including all irrationals, the space of all real numbers is **complete.** Mathematically, one states that the real numbers, $\mathcal{R}$, is the **completion** of the set of rational numbers.

(3) We all agree "irrational numbers exist" and they must be included in any sensible usage of numbers. However, for physics, it is also clear that, for a given problem, it is sufficient to work with rationals only, *e.g.* by agreeing to always work to the twentieth decimals. Another way of stating this fact is that "one can always approximate an irrational number by a rational number to any degree of accuracy one desires." Mathematically, one states that the set of rational numbers is **dense** in $\mathcal{R}$.

(4) It is a mathematical fact that the space L_2, defined by Eq. (I.1), as a metric space with a "distance between functions f and g" defined by $d(f, g) \equiv ||f - g||$, is **complete**. That is, the "limiting function" of any Cauchy sequence of functions in L_2 is also in L_2. Therefore, L_2 is a *Hilbert space*, which will also be denoted by $\mathcal{H}_1$. [Note: The notation $\mathcal{H}_2$ will be used to denote the Hilbert space of L_2 functions defined over $(-\infty, \infty)$, which will be discussed in **"Notes on Fourier Tranform"**].

(5) Let us clarify the situation by drawing an analogy with the case of real numbers. The role of rationals is now played by continuous functions on the interval $[a, b]$. However,

the condition of "square-integrability", (I.1), is much less stringent than continuity. Many functions which are not continuous nevertheless saitisfy (I.1). It can be shown that these "additional" functions can always be thought of as limits of Cauchy sequences of continuous functions. We "complete" L_2 by adding to the set of continuous functions these "limiting points". That is: L_2 is the **completion** of the space of continuous functions on the interval $[a, b]$, with respect to a distance defined by $d(f, g) \equiv ||f - g||$.

(6) Just like the situation with real numbers, any function $f \in L_2$ can always be approximated by continuous functions, to any desired degree of accuracy. Mathematically, one states that the set of continuous functions on $[a, b]$ is **dense** in L_2.

(7) Let us end with an example using Fourier series. let us consider L_2 defined over the interval $[-\pi, \pi]$. Consider a periodic function $f(x)$. Assume that $f(x)$ is continuous and differentiable except at a point $x_0 \epsilon [-\pi, \pi]$, where $f(x)$ has a finite discontinuity. Let us consider the following sequence of functions, $\{S_N(x)\}$, $N = 0, 1, 2, \cdots$, defined by

$$S_N(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^{N} c_n e^{inx}, \tag{II.1}$$

where

$$c_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} dx\, e^{-inx} f(x), \tag{II.2}$$

for all integers n. Note that each $S_N(x)$ is **continuous** at x_0 where $f(x)$ has a discontinuity.

For any small number ϵ, we can always choose N large enough so that

$$||f - S_{N'}|| < \epsilon, \tag{II.3}$$

for $N' \geq N$. That is, to the accuracy ϵ, f can be approximated by a continuous function, S_N, although $f(x)$ is not.

The set of $\{S_N\}$ is a Cauchy sequence. Therefore, the space of continuous functions is **not complete**. However, the spcae L_2 is. Indeed, it is precisely in the sense of (II.3) that we understand the statement that every periodic function of period 2π can be represented by a **Fourier series**:

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{II.4}$$

with c_n given by (II.2).

III. Standard Complete Orthonormal Basis for $\mathcal{H}_1$

DEFINITION: A system of functions, $\{\phi_n(x)\}$, defined on $[a,b]$ is said to be an orthonormal system if

$$\int_a^b \phi_n^*(x)\phi_m(x)dx = \delta_{n,m}. \tag{III.1}$$

DEFINITION: A system of functions, $\{\phi_n(x)\}$, defined on $[a,b]$ and belonging to L_2, is said to be **complete** if there exists no functions different from zero in L_2 which is orthogonal to all functions $\phi_n(n)$.

Without a loss of generality, let us now shift the interval $[a,b]$ to $[-L/2, L/2]$. Consider the set of basis functions, $\{U_n(x)\}$, $-\infty < n < \infty$, defined by

$$U_n(x) \equiv \frac{1}{\sqrt{L}} e^{\frac{2\pi i n x}{L}}. \tag{III.2}$$

We have previously pointed out that this set forms an orthonormal set:

$$\textbf{Orthonormality}: \quad \int_{-L/2}^{L/2} dx\, U_n^*(x)U_m(x) = \frac{1}{L}\int_{-L/2}^{L/2} dx\, e^{\frac{2\pi i(m-n)x}{L}} = \delta_{n,m}. \tag{III.3}$$

For each function F in L_2, consider the Fourier series representation

$$F(x) = \sum_{n=-\infty}^{\infty} C_n U_n(x), \tag{III.4}$$

$$C_n = \int_{-L/2}^{L/2} dx\, U_n^*(x)F(x). \tag{III.5}$$

The fact that any function in L_2 can be expanded in this basis can be used to show that this set is **complete**.

[**Note: The notion of a complete basis set** is **not** the same as the notion of a **complete** matric space.]

By substituting (III.5) into (III.4), one finds that the completeness can be expressed mathematically as

$$\textbf{Completeness Relation}: \quad \sum_{n=-\infty}^{\infty} U_n^*(x)\, U_n(x') = \frac{1}{L} \sum_{n=-\infty}^{\infty} e^{\frac{2\pi i n (x'-x)}{L}} = \delta(x - x'). \qquad \text{(III.6)}$$

COMMENTS:

(1) This set of **complete orthonormal** basis functions will be referred to as our **standard basis**. In terms of this standard basis, every function F can be expanded as in (III.4). Instead of writing out (III.4) explicitly, we can also represent the function F by a "column vector", with components: $\cdots, C_{-n}, \cdots, C_{-2}, C_{-1}, C_0, C_1, C_2, \cdots$. (To save space, let's not write it out as a "column" here.) That is, in this basis, we can associate a function with a vector with infinite many "components". In this representation, one "sees" a function as a column vector, just like the situation with a finite dimensional vector space.

(2) To draw a closer analogy with our "standard representation" for a finite dimensional vector space, let's re-write the inner-product, (I.2), and the norm, (I.3), as follows: Consider two functions, $F(x)$ and $F'(x)$, described by components $\{C_N\}$ and $\{C'_N\}$ respectively in this standard basis. It follows from (III.3) that

$$\langle F|F'\rangle \equiv \int_{-L/2}^{L/2} dx F^*(x) F'(x) = \sum_{n=-\infty}^{\infty} C_n^* C'_n, \qquad \text{(III.7)}$$

$$||f||^2 \equiv \langle f|f\rangle = \sum_{n=-\infty}^{\infty} |C_n|^2. \qquad \text{(III.8)}$$

(3) In this representation, the "vector" nature of the L_2 space is manifest! Instead of representing a vector by its "components", one can also represent a function, $F(x)$, by an abstract vector in a Dirac notation, $|F\rangle$. If we first introduce a Dirac notation for our standard basis vectors: $\{|U_n\rangle\}$, $-\infty < n < \infty$, Eq. (III.4) can be written as:

$$|F\rangle = \sum_{n=-\infty}^{\infty} C_n |U_n\rangle, \qquad \text{(III.9)}$$

with

$$C_n = \langle U_n | F \rangle \tag{III.10}$$

The orthonormality condition in a Dirac notation becomes

$$\textbf{Orthonormality}: \quad \langle U_n | U_m \rangle = \delta_{n,m}. \tag{III.11}$$

We will shortly show that the completeness condition in a Dirac notation is

$$\textbf{Completeness Relation}: \quad \sum_{n=-\infty}^{\infty} |U_n\rangle\langle U_n| = \hat{I}. \tag{III.12}$$

You should now compare these representations with those for a finite dimensional vector space, and convince yourself that these two sets are formally identical.

(4) You should also convince yourself that every **linear operator** on L_2, $\hat{M} : L_2 \longrightarrow L_2$, can be associated with an infinite by infinite matrix as follows:

$$\hat{M} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} M_{m,n} |U_m\rangle\langle U_n|. \tag{III.13}$$

What is the form of the matrix $\{M_{m,n}\}$ for the momentum operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$?

IV. Coordinate Basis for $\mathcal{H}_1$

It is convenient to **formally** define a set of **improper** basis vectors, $\{|x\rangle\}$, **labelled by a continuous index**, $x \in [a, b]$, satisfying **Dirac-delta** orthonormality condition as follows:

$$\textbf{Orthonormality}: \qquad \langle x|x'\rangle \equiv \delta(x - x'). \tag{IV.1}$$

In terms of this basis, it is now possible to express each abstract vector $|f\rangle$ in $\mathcal{H}_1$ as

$$\textbf{Coordinate Representation}: \qquad |f\rangle \equiv \int_a^b dx f(x)|x\rangle. \tag{IV.2}$$

It follows from (IV.1) and (IV.2) that $f(x)$ can be recovered by "dotting" a bra vector $\langle x|$ into Eq. (IV.2), *i.e.*,

$$f(x) = \langle x|f\rangle. \tag{IV.3}$$

That is, we have achieved a representation where the original function $f(x)$ in L_2 now appears as "components" of the abstract vector $|f\rangle$ in a coordinate basis. In particular, the inner product of two vectors, $|f\rangle$ and $|g\rangle$, through (IV.3) and (IV.2), is

$$\langle f|g\rangle \equiv \int_a^b \int_a^b dx dx' f^*(x) g(x') = \int_a^b dx f^*(x) g(x), \tag{IV.4}$$

in agreement with the original definition, (I.2).

Since every vector in $\mathcal{H}_1$ can be expressed by (IV.2), they form a complete basis:

$$\textbf{Completeness}: \qquad \int_a^b dx |x\rangle\langle x| = \hat{I}. \tag{IV.5}$$

[Eq. (IV.5) follows when one substitutes (IV.3) into (IV.2).]

COMMENTS:

(1) We claim that $|x_0\rangle$ is an eigenstate of our **position operator**, $\hat{x}$, *i.e.*,

$$\hat{x}|x_0\rangle = x_0|x_0\rangle. \tag{IV.6}$$

for $x_0 \in [a,b]$. Eq. (IV.6) follows if one writes down the components of $|x_0\rangle$, *i.e.*,

$$|x_0\rangle = \int_a^b dx\delta(x-x_0)|x\rangle. \tag{IV.7}$$

That is, $\delta(x-x_0)$ is the coordinate wavefunction for the eigenvector of $\hat{x}$ with an eigenvalue x_0, $\hat{x}\delta(x-x_0) = x_0\delta(x-x_0)$. By applying (IV.3) to (IV.7), one simply reproduces the orthogonality condition, (IV.1).

(2) In this representation, every linear operator can be expressed as

$$\hat{M} = \int_a^b dx \int_a^b dx' M(x,x')|x\rangle M(x,x')\langle x'| = \int_a^b dx \int_a^b dx' M(x,x')|x\rangle\langle x'|. \tag{IV.8}$$

This representation follows by considering $\hat{M} = \hat{I}\hat{M}\hat{I}$ and use (IV.5) twice. For instance,

$$\hat{x} = \int_a^b dx \int_a^b dx' x\delta(x-x')|x\rangle\langle x'|, \qquad \hat{p} = \int_a^b dx \int_a^b dx' \{-i\hbar\delta(x-x')\frac{\partial}{\partial x'}\}|x\rangle\langle x'|. \tag{IV.9}$$

Whenever $M(x,x')$ is proportional to $\delta(x-x')$, the operator $\hat{M}$ will be called a **local operator**. Instead of a "double-sum", Eq. (IV.9), one has a single-sum representation:

$$\hat{x} = \int_a^b dx x|x\rangle\langle x|, \qquad \hat{p} = \int_a^b dx\{-i\hbar\frac{\partial}{\partial x}\}|x\rangle\langle x|. \tag{IV.10}$$

For local operators, one often simply writes down the matrix $M(x,x')$ for the operator $\hat{M}$,

$$\hat{x} \longrightarrow x\delta(x-x') \qquad \text{and} \qquad \hat{p} \longrightarrow \{-i\hbar\delta(x-x')\frac{\partial}{\partial x'}\}.$$

Better yet, one can drop the $\delta(x-x')$ factor entirely

$$\hat{x} \longrightarrow x, \qquad \text{and} \qquad \hat{p} \longrightarrow \{-i\hbar\frac{\partial}{\partial x}\}$$

if it is understood that one is dealing with a local operator! For instance, instead of writing out

$$\hat{p}|f\rangle = \int_a^b\int_a^b dxdx'\{-i\hbar\delta(x-x')\frac{\partial}{\partial x'}\}f(x')|x'\rangle = \int_a^b dx\{-i\hbar\frac{\partial f(x)}{\partial x}\}|x\rangle$$

explicitly, it is much simpler to accept the notation

$$\hat{p}f(x) = -i\hbar\frac{\partial f(x)}{\partial x}.$$

[Convince yourself that the Hamiltonian operator $\hat{H}$ is local in a coordinate representation, and make sure you understand the proper interpretation of the expression $i\hbar\frac{\partial\psi(x,t)}{\partial t} = \hat{H}\psi(x,t)$.]

(3) Lastly, let's return to our "standard basis", $\{U_N(x)\}$. It is clear that the abstract vector, $|U_N\rangle$, is given by

$$|U_N\rangle = \int_{-L/2}^{L/2} dx U_N(x)|x\rangle, \qquad \text{and} \qquad U_N(x) = \langle x|U_N\rangle$$

With this understanding, by sandwiching (III.12) between $\langle x'|$ and $|x\rangle$, it leads to (III.6), which proves (III.12), as promised.