NOTES ON HILBERT SPACE

by Prof. C-I Tan
Department of Physics
Brown University

1 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.

1.1 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], ie,

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

This space will be denoted as $\mathcal{L}_2$. One can directly verify that $\mathcal{L}_2$ is a complex vector space, ie, f and $g \in \mathcal{L}_2$, then $\alpha f + \beta g \in \mathcal{L}_2$ for α and β, etc. We will return to clarify in what sense one can "visualize a function as a vector".

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

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

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

$$\|f\| \equiv \sqrt{\langle f|f \rangle} \tag{3}$$

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

(3) Define a "distance" between two functions by

$$d(f,g) \equiv \|f - g\| \tag{4}$$

This turns $\mathcal{L}_2$ into a metric space. We shall explain the fact that $\mathcal{L}_2$ is a Hilbert space next.

2 I. Hilbert Space is Complete

2.1 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., $\{3, 3.1, 3.14, 3.141, 3.1415, \ldots\}$. 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 $\mathcal{L}_2$, defined by 1, as a metric space with a "distance between functions f and g" defined by $\|f - g\|$, is **complete**. That is, the "limiting function" of any Cauchy sequence of functions in $\mathcal{L}_2$ is also in $\mathcal{L}_2$. Therefore, $\mathcal{L}_2$ 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 $\mathcal{L}_2$ functions defined over $(-\infty, \infty)$].

(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", 1, is much less stringent than continuity. Many functions which are not continuous nevertheless satisfy 1. It can be shown that these "additional" functions can always be thought of as limits of Cauchy sequences of continuous functions. We "complete" $\mathcal{L}_2$ by adding to the set of continuous functions these "limiting points." That is: $\mathcal{L}_2$ 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 $f \in \mathcal{L}_2$ 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 $\mathcal{L}_2$.

(7) Let us end with an example using Fourier series. Let us consider $\mathcal{L}_2$ defined over the interval $[-\pi, \pi]$. Consider a periodic function $f(x)$. Assume that $f(x)$ is continuous and differentiable except at a point $x_o \in [-\pi, \pi]$, where $f(x)$ has a finite discontinuity. Let us

2

consider the following sequence of functions, $\{\mathcal{S}(x)\}, N = 0, 1, 2, \cdots$, defined by

$$\mathcal{S}(x) = \tfrac{1}{\sqrt{2\pi}} \sum_{n=-N}^{N} a_n e^{inx} \qquad (5)$$

where

$$a_n = \tfrac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} dx a_n e^{n\pi x} f(x) \qquad . \qquad (6)$$

for all integers n. Note that each $\{\mathcal{S}(x)\}$ is **continuous** at $\{\mathcal{S}(x)\}$ where $f(x)$ has a discontinuity. For any small number ϵ , we can always choose N large enough so that

$$\|f - \{\mathcal{S}(x)\}\| < \epsilon \qquad for N' \geq N \qquad (7)$$

That is, to the accuracy ϵ, f can be approximated by a continuous function, $\{\mathcal{S}(x)\}$, although $f(x)$ is not.

The set of $\{\mathcal{S}(x)\}$ is a Cauchy sequence. Therefore, the space of continuous functions is **not complete**. However, the space $\mathcal{L}_2$ is. Indeed, it is precisely in the sense of 7 that we understand the statement that every periodic function of period 2π can be represented by a **Fourier series**):

$$f(x) = \tfrac{1}{\sqrt{2\pi}} \sum_{n=-N}^{N} a_n e^{inx} \qquad (8)$$

with a_n given by 6.

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

3.1 DEFINITION:

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

$$\text{an orthonormal system if } \int_a^b dx \phi_n^*(x) \phi_m(x) = \delta_{n,m} \qquad (9)$$

3.2 DEFINITION:

A system of functions, $\{\phi_n(x)\}$, defined on [a,b] and belonging to $\mathcal{L}_2$, is said to be complete if there exists no functions different from zero in $\mathcal{L}_2$ which is orthogonal to all functions $\phi_n(x)$.

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

$$U_n(x) \equiv \frac{1}{\sqrt{L}} e^{\frac{2\pi nx}{L}} \qquad . \qquad (10)$$

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

$$\int_{-L/2}^{L/2} dx U_n^*(x) U_m(x) = \frac{1}{L} \int_{-L/2}^{L/2} dx e^{\frac{2\pi(n-m)x}{L}} = \delta_{n,m} \tag{11}$$

For each function $\mathcal{F}$ in $\mathcal{L}_2$, consider the Fourier series representation

$$\mathcal{F}(x) = \sum_{n=-\infty}^{\infty} C_n U_n(x) \tag{12}$$

$$C_n = \int_{-L/2}^{L/2} dx U_n^*(x) \mathcal{F}(x) \tag{13}$$

The fact that any function in $\mathcal{L}_2$ can be expanded in this basis can be used to show that this set is complete. By substituting 10 into 12, one finds that the completeness can be expressed mathematically as

$$\sum_{n=-\infty}^{\infty} U_n^*(x) U_n(x') = \frac{1}{L} \sum_{n=-\infty}^{\infty} e^{\frac{2\pi n(x'-x)}{L}} = \delta(x - x') \tag{14}$$

3.3 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 12. Instead of writing out 12 explicitly, we can also represent the function f by a "column vector", with components: $\cdots, C_{-n}, \cdots, C_{-2}, C_{-1}, C_o, C_1, C_2, \cdots$.

$$\begin{pmatrix} \vdots \\ C_{-n} \\ \vdots \\ C_{-2} \\ C_{-1} \\ C_o \\ C_1 \\ C_2 \\ \vdots \\ C_n \\ \vdots \end{pmatrix} \tag{15}$$

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, 2, and the norm, 4, as follows: Consider two functions, $F(x)$ and F'(x), described by components $F(x)$ and $F'(x)$ respectively in this standard basis. It follows from 11 that

$$\langle F|F'\rangle \equiv \int_{-L/2}^{L/2} dx F^*(x)F'(x) = \sum_{n=-\infty}^{\infty} C_n^* C_n'$$

$$\|f\|^2 \equiv \langle f|f\rangle = \sum_{n=-\infty}^{\infty} |C_n|^2$$

(16)

(3) In this representation, the "vector" nature of the 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: $\{|U_n\rangle\}$, $-\infty < n < \infty$, 8 can be written as:

$$|f\rangle = \sum_{n=-\infty}^{\infty} C_n |U_n\rangle,$$

(17)

with

$$C_n = \langle U_n|\mathcal{F}\rangle$$

(18)

The orthonormality condition in a Dirac notation becomes

$$\langle U_n|U_m\rangle = \delta_{n,m}$$

(19)

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

$$\sum_{n=-\infty}^{\infty} |U_n\rangle\langle U_n| = \hat{I}$$

(20)

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 $\mathcal{L}_2, \hat{M} : \mathcal{L}_2 \to \mathcal{L}_2$, can be associated with an infinite by infinite matrix as follows:

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

(21)

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

3.4 Coordinate Basis for $\mathcal{H}_1$

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

$$\textbf{Orthonormality}: \quad \langle x|x'\rangle \equiv \delta(x - x')$$

(22)

5

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

$$\textbf{Coordinate Representation}: \quad |f\rangle \equiv \int_a^b dx F(x)|x\rangle \tag{23}$$

It follows from 22 and 23 that $f(x)$ can be recovered by "dotting" a bra vector $\langle x|$ into Eq. 23, ie,

$$f(x) = \langle x|f\rangle \tag{24}$$

That is, we have achieved a representation where the original function $f(x)$ in $\mathcal{L}_2$ 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 24 and 23, is

$$\langle f|g\rangle \equiv \int_a^b \int_a^b dx dx' f*(x)g(x') = \int_a^b dx f*(x)g(x) \tag{25}$$

in agreement with the original definition, 23. Since every vector in can be expressed by 23, they form a complete basis:

$$\textbf{Completeness}: \quad \int_a^b dx|x\rangle\langle x| = \hat{I} \tag{26}$$

3.5   COMMENTS:

(1) We claim that $|x_o\rangle$ is an eigenstate of our position operator, $\hat{x}$, ie,

$$\hat{x}|x_o\rangle = x_o|x_o\rangle \tag{27}$$

for $x_o \in [a,b]$. Eq. 27 follows if one writes down the components of $|x_o\rangle$, ie,

$$|x_o\rangle = \int_a^b dx \delta(x - x_o)|x\rangle \tag{28}$$

That is, $\delta(x - x_o)$ is the coordinate wavefunction for the eigenvector of $\hat{x}$ with an eigenvalue x_o, $\quad \hat{x}\delta(x - x_o) = x_o\delta(x - x_o)$. By applying 24 to 28, one simply reproduces the orthogonality condition, 22.

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

$$\hat{M} = \int_a^b dx \int_a^b dx' M(x,x')|x\rangle M(x,x')\langle x'| = \int_a^b dx \int_a^b dx' M(x,x')|x\rangle\langle x'| \tag{29}$$

This representation follows by considering $\hat{M} = \hat{I}\hat{M}\hat{I}$ and using 26 twice. For instance,

$$\hat{x} = \int_a^b dx \int_a^b dx' x \delta(x - x')|x\rangle\langle x'|, \quad \hat{p} = \int_a^b dx \int_a^b dx'\{-\imath\hbar\delta(x - x')\frac{\partial}{\partial x'}\}|x\rangle\langle x'| \tag{30}$$

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. 30, one has a single-sum representation:

$$\hat{x} = \int_a^b dx x|x\rangle\langle x|, \quad \hat{p} = \int_a^b dx\{-\imath\hbar\frac{\partial}{\partial x}\}|x\rangle\langle x| \tag{31}$$

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

$$\hat{x} \rightarrow x\delta(x - x') \qquad \hat{p} \rightarrow \{-\imath\hbar\delta(x - x')\frac{\partial}{\partial x}\} \tag{32}$$

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

$$\hat{x} \rightarrow x \qquad \hat{p} \rightarrow \{-\imath\hbar\frac{\partial}{\partial x}\} \tag{33}$$

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

$$\hat{p}|f\rangle = \int_a^b \int_a^b dx dx'\{-\imath\hbar\delta(x - x')\frac{\partial}{\partial x}\}f(x')|x'\rangle \tag{34}$$

explicitly, it is much simpler to accept the notation

$$\hat{p}f(x) = -\imath\hbar\frac{\partial}{\partial x}f(x). \tag{35}$$

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

(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

$$|U_N\rangle = \int_{-L/2}^{L/2} dx U_N(x)|x\rangle \qquad U_N(x) = \langle x|U_N\rangle \tag{36}$$

With this understanding, by sandwiching 20 between $\langle x'|$ and $|x\rangle$, it leads to 14, which proves 20, as promised.