20 Hilbert Space

Fourier expansion is quite parallel to the expansion of a vector into a linear combination of basis vectors in a finite dimensional vector space. However, function spaces are generally very different from finite dimensional vector spaces. To understand Fourier expansion more intuitively, it is convenient to introduce an infinite dimensional vector space in which our knowledge of finite dimensional vector spaces can be used almost 'freely.' This is the Hilbert space.

Key words: Hilbert space, scalar product, completeness, l_2, L_2, H^2, Cauchy-Schwartz inequality, bra-ket, dual space, K-vector space, orthonormal basis, Gram-Schmidt orthonormalization, generalized Fourier expansion, orthogonal projection, Bessel's inequality, Parseval's equality

Remember:
(1) Hilbert space is an infinite dimensional vector space in which we can define an angle between vectors (**20.3**).
(2) Understand Gram-Schmidt orthonormalization geometrically (**20.16**).
(3) Fourier expansion is a orthogonal decomposition in a Hilbert space (**20.14**).
(4) Be familiar with the bra-ket notation (**20.21-24**).
(5) Understand the formal expression of Green's functions (**20.28**).

20.1 Vector space. Let V be a set such that any (finite) linear combination of its elements with coefficients taken from a field K is again in V. V is called a *K-vector space.* K may be $\boldsymbol{R}$ or $\boldsymbol{C}$. A $\boldsymbol{R}$-vector space is called a *real vector space* and a $\boldsymbol{C}$-vector space is called a *complex vector space.* For example, the set $C^0([0,1])$ of continuous real functions on the interval $[0,1]$ is a real vector space. The set of analytic functions on the unit disc is a complex vector space.

Examples.
(1) The set of all the real polynomials of degree n forms a real vector space.
(2) The totality of continuous functions on $[a,b]$ is a vector space (with respect to the ordinary $+$ and $\times$).
(3) The totality of sequences $\{x_i\}$ converging to zero is a vector space, if we introduce $+$ as $\{x_i\}+\{y_i\}=\{x_i+y_i\}$ and scalar multiplication by $c\{x_i\}=\{cx_i\}$.

20.2 Infinite dimensional space. Consider the set $C^0([0,1])$ of all the continuous functions on $[0,1]$. x^n cannot be written as a linear combination of $1, x, x^2, \cdots, x^{n-1}$ for any n. Thus this function space is obviously infinite dimensional, if we wish to define the 'dimension' of the space as in the ordinary vector space by counting the necessary number of components to specify a vector uniquely. Another approach may be to refer to the interpretation of $f(x)$ as the x-component of a vector f as in functional differentiation (→**3.7**, **20.21**).[289]

Infinite dimensionality causes special difficulties in convergence. For example, the boundedness of a sequence does not guarantee the existence of a convergent subsequence. For example, consider,

$$(1,0,\cdots,),(0,1,0,\cdots),(0,0,1,0,\cdots),\cdots. \tag{20.1}$$

Discussion.
Infinite dimensional spaces have important peculiar features.
(1) We cannot define a 'uniform volume.' More precisely, there is no uniform measure (=volume) μ (→**19a**) such that for the unit cube C (of infinite dimension) $\mu(C) = 1$ with the translational symmetry (i.e., even if we translate an object, its volume does not change), and the additivity ($\mu(A \cup B) = \mu(A) + \mu(B)$, if $A \cap B = \emptyset$). If such a μ were to exists, then the volumes of most bounded sets are 0 or ∞.[290] Therefore, we cannot define the concept of 'almost everywhere' (→**19.5**).[291]
(2) Compactness and boundedness are distinct. Compactness means (→**A1.25**): if a set A is covered by a family of open sets, then A can already be covered by a finite subset of the family. If the space dimension is finite, this is equivalent to the open boundedness (the Heine-Borel theorem). However, this is obviously untrue for infinite dimensional space: to cover a unit open ball we need infinitely many open balls of radius 1/2. This distinction of compactness and boundedness in infinite dimensional space makes functional analysis much more difficult. A bounded operator and a compact operator are distinct (→**34C.9**).

20.3 Hilbert space. An infinite dimensional vector space V, which is complete (see below) with respect to the norm (→**3.3** footnote) defined

[289]In this case one might feel that the dimension is uncountable (→**17.15**(3)). However, usually we do not pay the minute details of the functions, but pay attention to the equivalence classes of functions as individual elements (for example, we ignore the difference on measure zero sets (→**19.3**), so that often the dimension is countable. See Weierstrass' theorem **17.3**.

[290]Here, we are not discussing 'non-measurable' sets. We confine ourselves to the Borel sets. That is, we discuss the sets which can be constructed as joins and intersections of countable finite cubes. See **19a**.

[291]See B R Hunt, T Sauer, and J A Yorke, "Prevalence: a translational-invariant "almost-every" on infinite dimensional spaces," Bull. Amer. Math. Soc. **27**, 217 (1992). Addendum **28**, 306 (1993).

by the scalar product (see below) is called a *Hilbert space.*[292]
A *scalar product* is a bilinear functional of two vectors $f, g \in V$ denoted by the bracket product $\langle f|g\rangle$ satisfying

$$\begin{aligned}
\langle f|f\rangle &\geq 0, \ \langle f|f\rangle = 0 \iff f = 0, && (20.2)\\
\langle f_1 + f_2|g\rangle &= \langle f_1|g\rangle + \langle f_2|g\rangle, && (20.3)\\
\langle f|g_1 + g_2\rangle &= \langle f|g_1\rangle + \langle f|g_2\rangle, && (20.4)\\
\overline{\langle f|g\rangle} &= \langle g|f\rangle, && (20.5)\\
\langle af|g\rangle &= \overline{a}\langle f|g\rangle, \ \langle f|ag\rangle = a\langle f|g\rangle. && (20.6)
\end{aligned}$$

Here a is a constant scalr (i.e., an element in K). The *norm* in a Hilbert space is defined by $\|f\| = \sqrt{\langle f|f\rangle}$. '*Complete*' means that all the Cauchy sequences[293] do converge: in particular, if $\|f_n - g\| \to 0$, then actually $f_n \to g$.

Introduction of scalar product allows us to introduce the concept of angle between two vectors. We may say that an infinite dimensional space in which we can talk about not only lengths but also angles is a Hilbert space. In other words, in any vector spaces we can define magnitudes by a norm, but the concept of direction is not easy to visualize. To this end, we need a scalar product to introduce the angle between vectors.

Discussion.

(A) **Banach space.** A complete normed space is called a Banach space. It is more important in the study of PDE than the Hilbert space. L_1 ($\to$**19.8**) is a typical Banach space.

(B) **Euclidean space.** In these notes, Hilbert spaces are defined as infinite dimensional spaces. Hilbert spaces and finite dimensional vector spaces (with the ordinary scalar product) are sometimes called Euclidean spaces (written as E^d).

20.4 Who was Hilbert? [294] David Hilbert was born in 1862. He studied mainly at Königsberg, where he befriended Minkowski (who was already famous when he was a high school student. He died relatively young due to appendicitis). From 1895 until his retirement in 1930 he was a named professor at Göttingen. At the Second International Congress of Mathematicians in Paris in 1900, he presented the

[292]The definition of 'Hilbert space' can change slightly from book to book. Many authors include finite dimensional vector spaces. Here, following Kolmogorov and Fomin, we understand that a Hilbert space is always infinite dimensional (need not be countably so).

[293]A Cauchy sequence for a given norm $\| \ \|$ is a sequence $\{y_n\}$ such that $\|y_n - y_m\| \to 0$ as n and m go to infinity. If the sequence is a complex number sequence, then the norm is the usual modulus. We know that $\boldsymbol{C}$ is complete.

[294]See also C Reid, *Hilbert* (Springer, 1970).

famous 23 problems for the mathematics of twentieth century. He had a characteristic optimism that new discoveries would continuously be made and that these discoveries were necessary for the vitality of mathematics.

His scientific study covers vast area of mathematics, algebra, number theory, functional analysis (as one of the founders; the term 'spectrum' (→**34B**, **34C**) is due to him). His *Grundlagen der Geometrie* (based first on the lectures delivered in 1898-9; there are many versions, because he continued to improve the work) made an epoch.[295] He endeavored to make axiomatic systems more general; he believed that fundamental terms should not have a single privileged interpretation.

Hilbert's last two main scientific interests were theoretical physics and foundation of mathematics. His study of the Boltzmann equation was an important contribution.

He was the major proponent of Formalism, trying hard to prove the consistency of the axiomatic systems on which the modern mathematics is based on (→**17.18(5)**). This was shown to be untenable by Gödel. However, we must remember that Gödel's sharp result was possible because the problem was posed (formulated) unambiguously by the Hilbert school.

Hilbert died during the World War II (1943). The motto on his grave in Göttingen reads, "Wir müssen wissen, wir werden wissen."[296]

20.5 Examples.
(1) l_2-space. Let V be the totality of infinite sequences $\{c_n\} = \{c_1, \cdots, c_n, \cdots\}$ such that $\sum_n c_n^2 < +\infty$. If we introduce the natural linear structure $a\{c_n\} = \{ac_n\}$ and $\{a_n\} + \{b_n\} = \{a_n + b_n\}$ and the scalar product $\{a_n\} \cdot \{b_n\} = \sum a_n b_n$, then V is a Hilbert space, which is called the *l_2-space.*
(2) $L_2([a,b])$. Let V be the totality of square Lebesgue integrable (→**19.8**) functions (complex valued) on the interval $[a,b]$. Then, with the definition of the scalar product

$$\langle f|g\rangle \equiv \int_a^b dx \overline{f(x)} g(x) \tag{20.7}$$

V becomes a Hilbert space called the $L_2([a,b])$-space (→**20.19**).[297]
(3) $H^1([a,b])$. Let V be the totality of Lebesgue square integrable functions defined on $[a,b]$ whose first derivatives are also square integrable.

[295] Hilbert's axiomatization of Euclidean geometry is summarized in the book of Mac Lane quoted in Book Guide (p63 and on of the book).
[296] We must know; we will know.
[297] Some authors use L^2 and l^2 for L_2 and l_2.

If we introduce the following scalar product

$$\langle f|g\rangle \equiv \int_a^b dx\{\overline{f(x)}g(x) + \overline{f'(x)}g'(x)\}, \tag{20.8}$$

then V becomes a Hilbert space called the H^1-space.[298] The norm based on this scalar product is called in the context of wave equations the energy norm ($\rightarrow$**a1D.12**).

Discussion.
(A) **Theorem[Riesz-Fischer].** Let $\{|n\rangle\}$ be an orthonormal set (not necessarily a basis $\rightarrow$**21.10**) of a Hilbert space H. Then for any element $c = \{c_n\}$ of l_2 ($\rightarrow$**21.4**(1)), there is $|a\rangle \in H$ such that $\langle n|a\rangle = c_n$. □
In this sense, any separable ($\rightarrow$**21.11**) Hilbert space is isomorphic.
(B) $\{(2\pi(n^2+1))^{-1/2}e^{inx}\}$ is a complete orthonormal basis of $\boldsymbol{H}^1([-\pi,\pi])$.
(C) Let $u \in L_2[(-\pi,\pi)]$. A condition for $u \in H^1([-\pi,\pi])$ is that $\sum_{n\in \boldsymbol{Z}} n^2|c_n|^2 < \infty$, where c_n is the complex Fourier expansion coefficient ($\rightarrow$**17.1**.

Exercise.
Set up the Gram-Schmidt orthonormalization scheme ($\rightarrow$**20.16**) for the $H^1([-1,1])$-space. Apply it to $\{1, x, x^2, \cdots\}$ and obtain the first three polynomials. Compare them with the Legendre polynomials ($\rightarrow$**21A.5**, **21B.2**).

20.6 Parallelogram law and Pythagoras theorem. Let V be a Hilbert space and $x, y \in V$.
(1) **Parallelogram law.** $\|x+y\| + \|x-y\| = 2(\|x\|^2 + \|y\|^2)$.
(2) **Pythagoras' theorem.** If $\langle x|y\rangle = 0$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Discussion.
The parallelogram law is a necessary and sufficient condition that the vector space is an Euclidean space ($\rightarrow$**20.3**). To demonstrate this we have only to show that

$$\langle x, y\rangle \equiv \frac{1}{4}(\|x+y\| - \|x-y\|) \tag{20.9}$$

is a respectable scalar product ($\rightarrow$**20.3**). Demonstrating the linearity (20.6 $\wedge$) is not very easy. See Kolmogorov-Fomin.

From this we can show that ℓ_p-space defined by $\sum |c_n|^p < \infty$ is a Euclidean space only when $p = 2$. Also the vector space $C_{[a,b]}$ can never be an Euclidean space.

20.7 Cauchy-Schwartz inequality. Let V be a Hilbert space and $f, g \in V$. Then

$$|\langle f|g\rangle| \le \|f\|\,\|g\|. \tag{20.10}$$

To prove this assume $g \neq 0$, and g is normalized (without loss of generality). Make $h \equiv f - g\langle g|f\rangle$. $\langle h|h\rangle \ge 0$ implies the desired inequality.

[298]This is an example of the *Sobolev space* (Sergei L'vovich Sobolev, 1908-?).

This inequality tells us a very obvious fact that the modulus of cosine cannot be larger than 1. As is often the case, very obvious things tell us deep things. Heisenberg's uncertainty principle is a disguised version of $|\cos\theta| \le 1$ (→**32B.1**).

From this it is easy to derive the
Triangle inequality: $\|f+g\| \le \|f\| + \|g\|$.

Discussion.
This inequality allows us to show that + and scalar product are continuous for a Hilbert space. For example, $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

20.8 Bracket notation.
(1) **Ket**. In elementary algebra, we regard an element of a vector space a column vector $\boldsymbol{a}$. Dirac introduced a symbol $|f\rangle$ to denote an element f of a vector space, and called it a *ket*.
(2) **Dual space**. A map from a K-vector space (→**20.1**) V to a field K is called a linear map, if it satisfies the superposition principle (→**1.4**): $f(\alpha|a\rangle+\beta|b\rangle) = \alpha f(|a\rangle)+\beta f(|b\rangle)$. The totality V^* of these linear maps is again a K-vector space.
Exercise.
Demonstrate this statement.
This space V^* is called the *dual space* of V.
(3) **Scalar product**. In a finite dimensional vector space V, a scalar product is introduced as $\langle \boldsymbol{a}, \boldsymbol{b}\rangle = \boldsymbol{a}^*\boldsymbol{b}$.[299] Any linear map $f(\boldsymbol{a})$ from a K-vector space to K can be uniquely described as a scalar product $f(\boldsymbol{a}) = \langle \boldsymbol{b}, \boldsymbol{a}\rangle$ by choosing an appropriate vector $\boldsymbol{b}$.
Exercise.
Demonstrate the above statement. [It is convenient to use a basis vector set of V.]
This implies that if $\boldsymbol{a} \in V$, then $\boldsymbol{a}^* \in V^*$. That is, (at least for a finite dimensional vector space) we may identify the dual space as the vector space spanned by all the row vectors. We write the hermitian conjugate of a ket $|a\rangle$ as $\langle a|$, which is called a *bra*. We regard V^* the totality of bras.
Notation. The scalar product of $|a\rangle$ and $|b\rangle$ is written as $\langle a|b\rangle$.

20.9 How Dirac introduced brackets. The bra-ket notation was introduced by Dirac. See P. A. M. Dirac, *Principles of Quantum Mechanics* (Oxford UP, 1958). The book is a good example to demonstrate that mathematical depth and mathematical rigor can be different. In this book he introduces kets to describe the states of a quantum mechanical system after explaining superposition of states is required to understand the double slit interference experiment. What he claims

[299] * implies the hermitian conjugate. That is, $\boldsymbol{a}^*$ is the complex conjugate of the transposition of $\boldsymbol{a}$.

is that the state space of a quantum mechanical system is a vector space. Then, he says that for a given vector space, there is always another space, and introduces the space of bras as the dual vectors of kets.

20.10 Orthonormal basis, separability. A subset $\{e_j\}$ of a Hilbert space V is said to be an *orthonormal basis*, if $\langle e_i|e_j\rangle = \delta_{ij}$ and the subspace spanned by $\{e_j\}$ is dense[300] in V. If a Hilbert space has a countable dense set, then we say the Hilbert space is *separable*. Separable Hilbert spaces have countable orthonormal basis.

Discussion.
(A) $L_2(\boldsymbol{R}^3)$ is separable.
(B) An example of a non-separable Hilbert space is the totality of functions on $[0,1]$ such that they are nonzero only on a countably many points, and the square sum of these values is finite. The scalar product is defined by $\langle x, y\rangle = \sum x(t)y(t)$, where the sum is over all the countable points on which $x(t)y(t) \neq 0$. (from Kolmogorov-Fomin)
(C) Let $e_n = \{\delta_{nk}\}_{k\in\boldsymbol{N}}$. Then, $\{e_n\}_{n=0}^{\infty}$ is a complete orthonormal system of ℓ_2.

20.11 Bessel's inequality. Let $\{|e_n\rangle\}$ be an orthonormal set of a separable Hilbert space V. Then for $\forall|f\rangle \in V$

$$\sum_{n=1}^{\infty} |\langle e_n|f\rangle|^2 \leq \langle f|f\rangle. \tag{20.11}$$

[Demo]

$$\|f - \sum_{n=1}^{N} |e_n\rangle\langle e_n|f\rangle\|^2 = \langle f|f\rangle - \sum_{n=1}^{N} |\langle e_n|f\rangle|^2 \geq 0 \tag{20.12}$$

for any positive integer N. Hence, (20.11).□

20.12 Parseval's equality. Let $\{|e_n\rangle\}$ be an orthonormal basis of a separable Hilbert space V. Then, for $\forall|f\rangle \in V$

$$\sum_{n=1}^{\infty} |\langle e_n|f\rangle|^2 = \langle f|f\rangle. \tag{20.13}$$

Conversely, if (20.13) holds for $\forall|f\rangle \in V$, then $\{|e_n\rangle\}$ is an orthonormal basis of V. (This follows easily from $|S[f]\rangle = |f\rangle$ (see below **20.14**). This is a natural extension of Pythagoras' theorem **20.6**.)

Discussion.

[300] i.e., for any $f \in V$ there is a sequence $\{a_i\}$ such that $b_N = \sum_{i=1}^{N} a_i e_i$ converges to f in the norm as $N \to \infty$. That is, $\{e_i\}$ is complete (→**20.3**).

(A) Let $Q = \{|n\rangle\}$ be an orthonormal set of a Hilbert space. Q is an orthonormal basis, iff[301] $|a\rangle$ satisfying $\langle n|a\rangle = 0$ for all n is actually zero.
[Demo] If Q is an orthonormal basis, vanishing of all the Fourier coefficients implies that $|a\rangle = 0$. Suppose Q is not a basis. Then due to Bessel's inequality **21.12** and Parseval's equality **21.13** there is a nonzero vector $|b\rangle$ such that

$$\langle b|b\rangle > \sum_n |\langle n|b\rangle|^2. \tag{20.14}$$

Thanks to the Riesz-Fischer theorem (→**D20.5(1)**), there is a ket $|a\rangle$ such that

$$|a\rangle = \sum_n |n\rangle\langle n|b\rangle. \tag{20.15}$$

Since $\langle b|b\rangle > \langle a|a\rangle$, $|b\rangle - |a\rangle \neq 0$. However, $\langle n|b - a\rangle = 0$ for any n. That is, there is a ket $|c\rangle$ satisfying $\langle n|c\rangle = 0$ for all n but not zero. Hence, if there is no such ket $|c\rangle$, then Q must be a basis.
(B) **Rademacher functions**. Define $r_n(x)$ as $r_0(x) = 1$ and

$$r_n(x) \equiv 1 - 2x_n \tag{20.16}$$

where x_n is the number of the n-th binary place of x. $R' = \{r_n(x)\}_{n\in \boldsymbol{N}}$ is called the *Rademacher orthogonal function system.*
(1) Show that it is an orthonormal system for $L_2([0,1])$.
(2) Show, however, the system is not complete.
(3) Let R be the totality of functions made by multiplying finite number of functions in R'. Then, R is a complete orthonormal system for $L_2([0,1])$.

20.13 Generalized Fourier expansion. Let $\{|e_n\rangle\}$ be an orthonormal basis (→**20.10**) of a Hilbert space V. The following sum for $|f\rangle \in V$

$$|S[f]\rangle = \sum_{n=1}^{\infty} |e_n\rangle\langle e_n|f\rangle \tag{20.17}$$

is called the *generalized Fourier expansion* of f (cf. **20.24**). Due to the definition of the orthonormal basis, actually $|S[f]\rangle = |f\rangle$.[302] The expansion allows us to make a one to one map between any separable Hilbert space (→**20.8**) and the ℓ_2-space (→**20.3**). Hence, all the separable Hilbert spaces are isomorphic.[303]

20.14 Least square approximation and Fourier expansion. 20.11

[301] i.e., if and only if.
[302] This equality is in the L_2 sense (→**20.5**). When this equality is in the ordinary sense is a non-trivial question as we have seen in **17**.
[303] In these notes, we use the terminology 'Hilbert space' for infinite dimensional cases only.

tells us that the Fourier coefficients can be determined by the following minimization problem:

$$\min \|f - \sum_{n=0}^{N} c_n e_n\|. \tag{20.18}$$

That is, the generalized Fourier series gives the best approximation in the L_2-sense. This gives another reason why L_2 is a natural space to consider Fourier series (Fourier analysis in general) (→**19.18**).

20.15 Decomposition of unity. The main result of **20.12** can be abstracted as

$$1 \equiv \sum_n |e_n\rangle\langle e_n| \tag{20.19}$$

for an orthonormal basis $\{|e_n\rangle\}$ of a Hilbert space V. This formula is called a *decomposition of unity.*

20.16 Gram-Schmidt orthonormalization. Let V be a Hilbert space, and $\{|1'\rangle, |2'\rangle, \cdots\}$ be a set of linearly independent kets in V whose linear hull is dense in V (i.e., complete →**20.3**). Then, we can construct an orthonormal basis $\{|1\rangle, |2\rangle, \cdots\}$ of V out of these kets as follows. The procedure is called the *Gram-Schmidt orthonormalization.*

(1) $|1\rangle = |1'\rangle/|1'|$, where $|a|$ will denote $\sqrt{\langle a|a\rangle}$ in this entry.
(2) $|2\rangle = |2''\rangle/|2''|$, where $|2''\rangle = (1 - |1\rangle\langle 1|)|2'\rangle$.
(3) $|3\rangle = |3''\rangle/|3''|$, where $|3''\rangle = (1 - |1\rangle\langle 1| - |2\rangle\langle 2|)|3'\rangle$, etc.

This is a method to construct orthogonal polynomials from $1, x, x^2, x^3, \cdots$ (→**21A.2**).

20.17 Respect the order in the basis. Hilbert spaces may almost be treated as finite dimensional vector space. However, we must respect the ordering of the basis set. The (generalized) Fourier expansion is not absolutely convergent usually, so this is a very natural thing to respect.

20.18 Orthogonal projection. Let the k-th summand in (20.19) be $P_k \equiv |e_k\rangle\langle e_k|$. Then we have $P_iP_j = P_jP_i = \delta_{ij}P_i$. Especially, $P_iP_i = P_i$. These operators are hermitian, $P_k^* = P_k$.

If a linear operator P satisfies the *idempotency*, i.e., $P^2 = P$, then P is called a *projection* (or a projection operator).
If it is hermitian, then it is called an *orthogonal projection*: For a nonzero ket $|a\rangle$, let $|p\rangle \equiv P|a\rangle$ and $|q\rangle \equiv (1 - P)|a\rangle$. $\langle p|q\rangle = \langle a|P^*(1 - P)|a\rangle = \langle a|(P^* - P^*P)|a\rangle$. If P is hermitian, this vanishes. That is, $|p\rangle$ and $|q\rangle$ are orthogonal.

Discussion.
(A) **[What is P_1P_2?]** Let P_1 and P_2 be orthogonal projection operators. A necessary and sufficient condition for P_1P_2 to be a projection operator is that P_1 and

P_2 commute. Let $P_iV = V_i$, where V is a vector space on which these projection operators are defined. What is P_1P_2V?
(B) **[System reduction]**. We wish to study a nonlinear equation

$$\frac{du}{dt} = \mathcal{N}(u). \tag{20.20}$$

Here $\mathcal{N}$ is a nonliner functional (a map). Formally, orthogonal projections are used to reduce a complicated system. Suppose P is a projection to a space spanned by 'important variables' (say, slow variables). Let us write $Q = 1 - P$. We can formally rewrite

$$\frac{dPu}{dt} = P\mathcal{N}(Pu + Qu), \tag{20.21}$$

$$\frac{\partial Qu}{\partial t} = Q\mathcal{N}(Pu + Qu). \tag{20.22}$$

If we could solve the second equation for Qu for any Pu as $Qu = F(Pu)$, then the first member becomes

$$\frac{dPu}{dt} = P\mathcal{N}(Pu + F(Pu)). \tag{20.23}$$

In this way we can get rid of unwanted variables, and reduce the number of variables or the dimension of the space we work. The procedure is only formal, and the crucial point is how to choose P, and how to obtain F. This is a very active field of research now.

20.19 Space $L_2([a, b], w)$. Let $L_2([a, b], w)$ be the totality of the functions which are square integrable[304] with the weight w on the interval $[a, b]$:

$$L_2([a, b], w) \equiv \{f \mid \int_a^b |f(x)|^2 w(x)dx < \infty\}. \tag{20.24}$$

This set is a Hilbert space with the following definition of the scalar product

$$\langle f|g\rangle \equiv \int_a^b \overline{f(x)}g(x)\, w(x)dx. \tag{20.25}$$

When $w(x) \equiv 1$ we omit w and write $L_2([a, b])$ as in **20.5**. $L_2((-\infty, +\infty))$ is often written as L_2 or $L_2(\boldsymbol{R})$. The convergence with respect to the norm (called the *L_2-norm*) defined by $\|f\| = \sqrt{\langle f|f\rangle}$ is called the *L_2-convergence.* As we know from the theory of Lebesgue integrals (→**19.8**), we may freely change the values of the function on a measure zero set (→**19.3**). so that the convergence in this sense could be quite different from the ordinary sense of convergence (w.r.t the sup norm).

Discussion.
(A) **measure** (→**19a**). Mathematicians usually avoid to discuss the weight functions w, because w need not be an ordinary function (i.e., the density need not be

[304] Usually, 'integrable' means 'Lebesgue integrable' (→**19.8**).

well-behaved). Hence, instead of writing wdx we usually write $d\mu$, introducing a measure μ. Hence, more officially, it is better to call $L_2([a,b],w)$ as $L_2([a,b],\mu)$:

$$L_2([a,b],\mu) \equiv \{f \mid \int_a^b |f(x)|^2 d\mu(x) < \infty\}. \tag{20.26}$$

(B) **L_p-space**. The L_p-space ($p \geq 1$) is defined by the completion[305] of the following function set

$$\{\varphi \mid \|\varphi\|_p < +\infty\}, \tag{20.27}$$

where $\| \ \|_p$ is the L_p-norm defined a

$$\|f\|_p \equiv \left(\int |f|^p dx\right)^{1/p}. \tag{20.28}$$

L_p-space is a Banach space (→**20.3** Discussion), but not a Hilbert space except for $p = 2$, because the parallelogram law (→**20.6**) does not hold.

20.20 Dirac's "abuse" of symbols. As we have seen, in a Hilbert space[306] Dirac's bra-ket notation causes no mathematical problem and is quite useful. However, Dirac wished to unify not only the linear space spanned by normalizable states (physically, localized states →**34C.8**(4); this part is a Hilbert space) but also the space containing 'plane wave states' which cannot be normalized in the usual way.[307] The starting point of his formal approach is the following interpretation of an ordinary function as a vector with uncountably many components.

20.21 $f(x)$ as an x-component of a vector. It is not an unnatural idea to regard the i-th component of a vector $|v\rangle$ as a 'value' $v(i)$ of a function v defined on $\{1, 2, \cdots, n\}$, where n is the dimension of the vector space. Then, as we have already used the idea (→**3.7**), it is not outrageous to regard $f(x)$ as the 'x-component' of a vector $|f\rangle$. We know the i-th component of a vector v may be written as $v_i = \langle i|v\rangle$ using the basis vecor $|i\rangle$. Analogously, we write

$$f(x) = \langle x|f\rangle, \quad \overline{f(x)} = \langle f|x\rangle. \tag{20.29}$$

[We Thus we may regard a function as a vector in an infinite dimensional vector space spanned by position kets $\{|x\rangle : x \in [a,b]\}$. These position kets may be regarded as orthonormal vectors (→**20.10**).

[305] Completion means to add elements to make all the Cauchy sequences have unique limits.

[306] assuming separability (→**20.10**)

[307] Dirac wished to use the Hilbert space notation in a much wider class of spaces now called rigged Hilbert space.

20.22 Inner product of functions. It is natural to interpret summations over the coordinate indices as integrations (weighted with a function w as in **20.19**) over the independent variable x. Thus, it is natural to define the scalar product or *inner product* of two functions f and g defined on the same domain as

$$\langle f|g\rangle \equiv \int dx w(x)\langle f|x\rangle\langle x|g\rangle = \int dx w(x)\overline{f(x)}g(x). \tag{20.30}$$

20.23 Decomposition of unity. The formula (20.30) suggests that we can decompose unity (cf. **20.15**) as

$$\int |x\rangle w(x)dx\langle x| \equiv 1. \tag{20.31}$$

This suggests that we may interpret $\{|x\rangle\}$ as an "orthonormal basis." Often unity is written as the following operator:

$$1 = |x\rangle \int dx\, w(x)\langle x|. \tag{20.32}$$

20.24 Trigonometric expansion revisited. Let $V = L_2([-\pi, \pi])$ ($\rightarrow$**20.5**). Let us introduce the kets $|0\rangle, |n, c\rangle, |n, s\rangle$ such that

$$\langle x|0\rangle = \frac{1}{\sqrt{2\pi}}, \quad \langle x|n, c\rangle = \frac{1}{\sqrt{\pi}}\cos nx, \quad \langle x|n, s\rangle = \frac{1}{\sqrt{\pi}}\sin nx. \tag{20.33}$$

Then $\{|0\rangle, |1, c\rangle, |1, s\rangle, |2, c\rangle, |2, s\rangle, \cdots\}$ is an orthonormal basis, because it is a complete set for C^0-functions on $[-\pi, \pi]$, ($\rightarrow$**17.4**). The standard Fourier expansion **17.1** is

$$|f\rangle = |0\rangle\langle 0|f\rangle + \sum_{n=1}^{\infty}\{|n, c\rangle\langle n, c|f\rangle + |n, s\rangle\langle n, s|f\rangle\}. \tag{20.34}$$

[Here, the equality is in the L_2-sense.] Notice, again, that the equality in this formula is in the L_2-sense. Bessel's inequality ($\rightarrow$**20.11**) and Parseval's equality ($\rightarrow$**20.12**) adapted to the trigonometric function set are their original forms.

20.25 δ-function (with weight). We can formally write ($\rightarrow$**20.23**)

$$f(x) = \langle x|1|f\rangle = \int \langle x|y\rangle w(y)dy\langle y|f\rangle = \int f(y)\langle x|y\rangle w(y)dy. \tag{20.35}$$

Therefore, it is natural to introduce

$$\langle x|y\rangle = \delta_w(x - y) \tag{20.36}$$

such that

$$\begin{aligned} &\int \delta_w(x-y)w(y)dy = 1, \\ &\delta_w(x-y) = 0 \quad x \neq y. \end{aligned} \tag{20.37}$$

Obviously, δ_w is a generalization of δ ($\rightarrow$**14.5**). We should identify as

$$\delta_w(x-y) = \delta(x-y)/w(x). \tag{20.38}$$

Exercise.
Show (for $r' > 0$)

$$\delta(x-x')\delta(y-y')\delta(z-z') = \delta(r-r')\delta(\theta-\theta')\delta(\varphi-\varphi')/r^2 \sin\theta. \tag{20.39}$$

20.26 δ-function for curvilinear coordinates. (20.38) tells us that if we wish to use functions defined in terms of the O-$q^1q^2q^3$ coordinates which are orthogonal curvilinear ($\rightarrow$**2D.3**), then it is natural to choose the function space whose scalar product uses the weight function $w = h_1h_2h_3$ ($\rightarrow$**2D.8**). Thus it is convenient to define the position bra-ket with the normalization

$$\langle q^1, q^2, q^3 | q'^1, q'^2, q'^3 \rangle = \delta(q^1 - q'^1)\delta(q^2 - q'^2)\delta(q^3 - q'^3)/h^1h^2h^3. \tag{20.40}$$

For example, for the spherical coordinate system ($\rightarrow$**2D.5**)

$$\langle r, \theta, \varphi | r', \theta', \varphi' \rangle = \frac{\delta(r-r')\delta(\theta-\theta')\delta(\varphi-\varphi')}{r^2 \sin\theta}. \tag{20.41}$$

Exercise.
Write down the δ-function adapted to the elliptic cylindrical coordinates.

20.27 Delta function in terms of orthonormal basis. Since $\delta(x-y) = \langle x|y\rangle$ may be interpreted as $\langle x|1|y\rangle$, we may introduce the decomposition of unity **20.15** into this formula to obtain

$$\delta(x-y) = \sum_n e_n(x)^* e_n(y), \tag{20.42}$$

where $\{|e_n\rangle\}$ is an orthonormal basis, and $e_n(x) \equiv \langle x|e_n\rangle$.

20.28 Green's operator and Green's function – a formal approach. We have already seen the fundamental idea of Green in **1.8**, and know several examples of Green's functions ($\rightarrow$**15**, **16**). We wish to solve the following linear equation:

$$[Lu](z) = f(z) \tag{20.43}$$

with the homogeneous boundary condition. Let $\{|x\rangle\}$ be the position kets w.r.t. the Cartesian coordinates (→**20.21**). With the aid of the decomposition of unity (→**20.23**), we rewrite (20.43) as

$$\langle z|L|y\rangle \int dy \langle y|u\rangle = \langle z|f\rangle \tag{20.44}$$

or

$$\int dy L(z,y)u(y) = f(z), \tag{20.45}$$

where $L(x,y) \equiv \langle x|L|y\rangle$ (a sort of matrix element). If we can invert the 'matrix' $L(x,y)$, then we can solve this equation. In other words, if we can solve

$$LG = 1 \tag{20.46}$$

for G, then $u = Gf$ tanks to superposition (linearity). (20.46) reads

$$\int dy L(x,y)\langle y|G|z\rangle = \langle x|z\rangle = \delta(x-z). \tag{20.47}$$

G is called a *Green's operator*, and $G(x|y) \equiv \langle x|G|y\rangle$ is called a *Green's function.* Formally, $G = L^{-1}$, so that $G(x|y) = \langle x|L^{-1}|y\rangle$.

20.29 Eigenfunction expansion of Green's function – a formal approach. Suppose we know the eigenkets $\{|n\rangle\}$ of the operator L:

$$L|n\rangle = \lambda_n|n\rangle \tag{20.48}$$

If all the eigenvalues are non-zero, then formally

$$G(x,y) = \langle x|L^{-1}|y\rangle = \sum_n u_n(x)\lambda_n^{-1}\overline{u_n(y)}, \tag{20.49}$$

where $\langle x|n\rangle = u_n(x)$. Here we have assumed that the eigenkets of L make a complete orthonormal set. This is the Fourier decomposition formula for the Green's function. We can immediately see the symmetry of the Green's function: $G(x|y) = G(y|x)$ (→**16A.20**, **35.2**, **36.4**, **37.7**). We will later return to a more careful discussion (→**37**).