# INTRODUCTION

Functional analysis is a branch of mathematics which uses the intuitions and language of geometry in the study of functions. The classes of functions with the richest geometric structure are called Hilbert spaces, and the theory of these spaces is the core around which functional analysis has developed. One can begin the story of this development with Descartes' idea of algebraicizing geometry. The device of using co-ordinates to turn geometric questions into algebraic ones was so successful, for a wide but limited range of problems, that it dominated the thinking of mathematicians for well over a century. Only slowly, under the stimulus of mathematical physics, did the perception dawn that the correspondence between algebra and geometry could also be made to operate effectively in the reverse direction. It can be useful to represent a point in space by a triple of numbers, but it can also be advantageous, in dealing with triples of numbers, to think of them as the co-ordinates of points in space. This might be termed the geometrization of algebra: it enables new concepts and techniques to be derived from our intuition for the space we live in. It is regrettable that this intuition is limited to three spatial dimensions, but mathematicians have not allowed this circumstance to prevent them from using geometric terminology in handling $n$-tuples of numbers when $n > 3$. In the context of $\mathbb{R}^n$ one routinely speaks of points, spheres, hyperplanes and subspaces. Though such language comes to seem very natural to us, it still depends on analogy, and we must have recourse to algebra and analysis to verify that our analogies are valid and to determine which analogies are useful.

Once the geometric habit of mind was established in relation to $\mathbb{R}^n$ it was natural to extend it to other common objects of mathematics which enjoy a similar linear structure, such as functions and infinite sequences of real numbers. This is a bolder leap into the unknown, and we must expect that

Cambridge University Press
978-0-521-33717-5 - An Introduction to Hilbert Space
Nicholas Young
Excerpt
More information



our intuition for physical space will prove a shakier guide than it was for $\mathbb{R}^n$. Indeed, the task of sorting out the right basic concepts in the geometry of infinite-dimensional spaces preoccupied leading analysts for some decades around the turn of the century. Thereafter the geometric viewpoint proved its worth, and came to provide the backdrop for the greater part of modern work in differential and integral equations, quantum mechanics and other disciplines to which mathematics is applied.

The study of differential and integral equations arising in physics was one of the main impulses to the emergence of functional analysis. A precursor of the subject can be seen in attempts by several mathematicians to treat such equations as limits in some sense of finite systems of equations. This approach had fair success, particularly in the hands of Hilbert, and it still has plenty of life in the domain of numerical analysis. Suppose, for example, one wishes to solve the integral equation

$$\int_0^1 K(x, y) f(y)\, \mathrm{d}y = g(x).$$

Here $K$ and $g$ are known continuous functions on $[0, 1] \times [0, 1]$ and $[0, 1]$ respectively, and one is looking for a continuous solution $f$. It seems natural to approximate this system by the finite system

$$\sum_{j=0}^{n-1} K\left(\frac{i}{n}, \frac{j}{n}\right) f_{jn} \cdot \frac{1}{n} = g\left(\frac{i}{n}\right),$$

$i = 0, 1, \ldots, n-1$. Assuming that this system of $n$ linear equations in the $n$ unknowns $f_{0n}, \ldots, f_{n-1,n}$ has a unique solution, one might expect that, for large $n$, $f_{jn}$ ought to be close to $f(j/n)$, at least under further conditions on $K$ and $g$.

Hilbert was by no means the first to use this device. Fourier himself was led to introduce Fourier series in a rather similar way. In studying the conduction of heat he encountered the differential equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

subject to certain boundary conditions. By the method of the separation of variables he derived the solution

$$V(x, y) = \sum_{m=1}^{\infty} a_m \mathrm{e}^{-(2m-1)x} \cos(2m-1)y,$$

where the coefficients $a_m$ are determined by the infinite system of linear equations

$$\sum_1^\infty a_m = 1,$$

$$\sum_1^\infty (2m-1)^2 a_m = 0,$$

$$\sum_1^\infty (2m-1)^4 a_m = 0,$$

$$\ldots$$

Fourier handled these by taking the first $k$ equations and truncating them to $k$ terms. This gives a $k \times k$ system which has a solution $a_1^{(k)}, \ldots, a_k^{(k)}$. On letting $k \to \infty$ Fourier obtained the desired solution of the infinite system.

Although this trick often worked, it has its dangers. Consider the infinite system

$$\begin{aligned} x_1 + x_2 + x_3 + \cdots &= 1, \\ x_2 + x_3 + \cdots &= 1, \\ x_3 + \cdots &= 1, \\ \ldots & \end{aligned}$$

No choice of the $x_j$ will satisfy this system, yet Fourier's limiting procedure would yield the apparent solution $x_j = 0$ for all $j$.

By virtue of powerful technique and a perception of what was important, Hilbert was able to make great contributions using this idea. Nevertheless, mathematicians came to regard the method as inadequate. It is clumsy and notationally complicated. The procedure of passage to the limit is difficult, and, indeed, it has been asserted that Hilbert did not always accomplish it correctly (see Reid, 1970). He himself did not arrive at the modern geometric viewpoint: Hilbert never used 'Hilbert space'. It was other mathematicians, particularly Erhardt Schmidt and Frigyes Riesz, who reflected on his results and discovered the right conceptual framework for them. Thereby they created a simpler, more elegant and more powerful theory. In this one does not try to reduce essentially infinite-dimensional questions to finite-dimensional geometry and then 'let $n \to \infty$': instead one develops the geometry of the objects of analysis as they naturally occur, using the familiar finite-dimensional geometry rather as a guide and analogy.

Cambridge University Press
978-0-521-33717-5 - An Introduction to Hilbert Space
Nicholas Young
Excerpt
More information

# 1

# *Inner product spaces*

Some important metric notions such as length, angle and the energy of physical systems can be expressed in terms of the *inner product* $(x, y)$ of vectors $x, y \in \mathbb{C}^n$. This is defined by

$$(x, y) = \sum_{i=1}^{n} x_i \bar{y}_i, \tag{1.1}$$

where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, and $\bar{y}_i$ is the complex conjugate of $y_i$. We wish to construct an infinite-dimensional version of this inner product. The most obvious attempt is to consider the space $\mathbb{C}^{\mathbb{N}}$ of all complex sequences indexed by $\mathbb{N}$. This is a complex vector space in a natural way, but it is not clear how we can extend the notion of inner product to it. If we replace the finite sum in (1.1) by an infinite one then the series will fail to converge for many pairs of sequences. We therefore restrict attention to a subspace of $\mathbb{C}^{\mathbb{N}}$.

***1.1 Definition*** $\ell^2$ denotes the vector space over $\mathbb{C}$ of all complex sequences $x = (x_n)_{n=1}^{\infty}$ which are square summable, that is, satisfy

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty,$$

with componentwise addition and scalar multiplication, and with inner product given by

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n, \tag{1.2}$$

where $x = (x_n)$, $y = (y_n)$. □

'Componentwise' means the following: if $x = (x_n)$, $y = (y_n) \in \ell^2$ and $\lambda \in \mathbb{C}$ then

$$x + y = (x_n + y_n)_{n=1}^{\infty},$$
$$\lambda x = (\lambda x_n)_{n=1}^{\infty}.$$

Let us check that this definition of inner product does make sense. Using the Cauchy–Schwarz inequality we find, for $k \in \mathbb{N}$,

$$\begin{aligned}\sum_{n=1}^{k} |x_n \bar{y}_n| &= \sum_{n=1}^{k} |x_n||y_n| \\ &\leqslant \left\{\sum_{n=1}^{k} |x_n|^2\right\}^{1/2} \left\{\sum_{n=1}^{k} |y_n|^2\right\}^{1/2} \\ &\leqslant \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{1/2} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{1/2}.\end{aligned}$$

If $(x_n)$ and $(y_n)$ are square summable sequences then the latter expression is a finite number independent of $k$. Thus the series (1.2) converges absolutely, and so $(x, y)$ is defined by (1.2) as a complex number for any $x, y \in \ell^2$.

It is obvious that $\ell^2$ is closed under scalar multiplication but less so that it is closed under addition: we defer the proof of this to Exercise 1.12 below.

Let us make precise what it means to say that $\mathbb{C}^n$ and $\ell^2$ are spaces with an inner product.

**1.2 *Definition*** An *inner product* (or *scalar product*) on a complex vector space $V$ is a mapping

$$(\cdot, \cdot): V \times V \to \mathbb{C}$$

such that, for all $x, y, z \in V$ and all $\lambda \in \mathbb{C}$,

(i) $(x, y) = (y, x)^-$;

(ii) $(\lambda x, y) = \lambda(x, y)$;

(iii) $(x + y, z) = (x, z) + (y, z)$;

(iv) $(x, x) > 0$ when $x \neq 0$.

*An inner product space* (or *pre-Hilbert space*) is a pair $(V, (\cdot, \cdot))$ where $V$ is a complex vector space and $(\cdot, \cdot)$ is an inner product on $V$. □

It is routine to check that the formulae (1.1) and (1.2) do define inner products on $\mathbb{C}^n$ and $\ell^2$ in the sense of Definition 1.2. There are many other inner product spaces which arise in analysis, most of them having inner products defined in terms of integrals.

**1.3 *Exercise*** Show that the formula

$$(f, g) = \int_0^1 f(t)\overline{g(t)}\, \mathrm{d}t$$

defines an inner product on the complex vector space $C[0, 1]$ of all continuous $\mathbb{C}$-valued functions on $[0, 1]$, with pointwise addition and scalar multiplication. □

***1.4 Exercise*** Show that the formula

$$(A, B) = \text{trace}(B^* A)$$

defines an inner product on the space $\mathbb{C}^{m \times n}$ of $m \times n$ complex matrices, where $m, n \in \mathbb{N}$ and $B^*$ denotes the conjugate transpose of $B$. □

The conditions (ii) and (iii) in the definition of inner product are often summarized by the statement that $(\cdot, \cdot)$ is linear in the first argument. It follows from the definition that it is also *conjugate linear* in the second argument: this means that it satisfies (i) and (ii) of the following.

***1.5 Theorem*** For any $x, y, z$ in an inner product space $V$ and any $\lambda \in \mathbb{C}$,

(i) $(x, y+z) = (x, y) + (x, z)$;

(ii) $(x, \lambda y) = \bar{\lambda}(x, y)$;

(iii) $(x, 0) = 0 = (0, x)$;

(iv) if $(x, z) = (y, z)$ for all $z \in V$ then $x = y$.

*Proof.* (i) Using Definition 1.2(i) and (iii) we have

$$\begin{aligned}(x, y+z) &= (y+z, x)^- \\ &= [(y, x) + (z, x)]^- \\ &= (y, x)^- + (z, x)^- \\ &= (x, y) + (x, z).\end{aligned}$$

The proof of (ii) is similar. To prove (iii) put $\lambda = 0$ in (ii).

(iv) If $(x, z) = (y, z)$ then

$$\begin{aligned}0 &= (x, z) + (-1)(y, z) \\ &= (x, z) + (-y, z) = (x - y, z).\end{aligned}$$

If this holds for all $z \in V$ then in particular it holds when $z = x - y$; thus $(x - y, x - y) = 0$. By 1.2(iv) it follows that $x - y = 0$. □

## 1.1 Inner product spaces as metric spaces

In the familiar case of $\mathbb{R}^3$ the magnitude $|\boldsymbol{u}|$ of a vector $\boldsymbol{u}$ is equal to $(\boldsymbol{u}, \boldsymbol{u})^{1/2}$, and the Euclidean distance between points with position vectors $\boldsymbol{u}, \boldsymbol{v}$ is $|\boldsymbol{u} - \boldsymbol{v}|$. We copy this to introduce a natural metric in an inner product space.

***1.6 Definition*** The *norm* of a vector $x$ in an inner product space is defined to be $(x, x)^{1/2}$. It is written $\|x\|$. □

Thus, for $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ we have

$$\|x\| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2},$$

while for $f \in C[0, 1]$, with the inner product described in Exercise 1.3,

$$\|f\| = \left\{\int_0^1 |f(t)|^2 \, dt\right\}^{1/2}.$$

***1.7 Exercise*** Let $x = (1/n)_{n=1}^{\infty} \in \ell^2$. Show that $\|x\| = \pi/\sqrt{6}$. What is $\|I_n\|$ where $I_n \in \mathbb{C}^{n \times n}$ is the identity matrix and the inner product of Exercise 1.4 is used? □

***1.8 Theorem*** For any $x$ in an inner product space $V$ and any $\lambda \in \mathbb{C}$

(i) $\|x\| \geqslant 0$; $\|x\| = 0$ if and only if $x = 0$;

(ii) $\|\lambda x\| = |\lambda| \|x\|$.

*Proof.* (ii)

$$\begin{aligned}\|\lambda x\| &= (\lambda x, \lambda x)^{1/2} = \{\lambda\bar{\lambda}(x, x)\}^{1/2} \\ &= |\lambda| \|x\|.\end{aligned}$$ □

One knows that in $\mathbb{R}^3$ $(x, y)$ is $\|x\| \|y\|$ times the cosine of an angle, from which it follows that $|(x, y)| \leqslant \|x\| \|y\|$. This relation continues to hold in a general inner product space.

***1.9 Theorem*** For $x, y$ in an inner product space $V$,

$$|(x, y)| \leqslant \|x\| \|y\|, \tag{1.3}$$

with equality if and only if $x$ and $y$ are linearly dependent.

(1.3) is known as the *Cauchy–Schwarz inequality*.

*Proof.* Suppose first that $x$ and $y$ are linearly dependent – say $x = \lambda y$ where $\lambda \in \mathbb{C}$. Then both sides of (1.3) equal $|\lambda| \|y\|^2$, and so (1.3) holds with equality.

Now suppose that $x$ and $y$ are linearly independent: we must show that (1.3) holds with strict inequality. For any $\lambda \in \mathbb{C}$, $x + \lambda y \neq 0$ and therefore

$$\begin{aligned}0 &< (x + \lambda y, x + \lambda y) \\ &= (x, x + \lambda y) + (\lambda y, x + \lambda y) \\ &= (x, x) + (x, \lambda y) + (\lambda y, x) + (\lambda y, \lambda y) \\ &= \|x\|^2 + \bar{\lambda}(x, y) + \lambda(x, y)^- + |\lambda|^2 \|y\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}\{\bar{\lambda}(x, y)\} + |\lambda|^2 \|y\|^2.\end{aligned}$$

Pick a complex number $u$ of unit modulus such that $\bar{u}(x, y) = |(x, y)|$. On putting $\lambda = tu$ we deduce that, for any $t \in \mathbb{R}$,

$$0 < \|x\|^2 + 2|(x, y)|t + \|y\|^2 t^2.$$

This can only happen if the real quadratic on the right hand side has negative discriminant: that is,

$$4|(x,y)|^2 - 4\|x\|^2\|y\|^2 < 0,$$

which yields the desired conclusion

$$|(x,y)| < \|x\|\|y\|. \qquad \square$$

***1.10 Exercise*** Prove that, for any $f \in C[0, 1]$,

$$\left|\int_0^1 f(t)\sin \pi t \,\mathrm{d}t\right| \leqslant \frac{1}{\sqrt{2}}\left\{\int_0^1 |f(t)|^2 \,\mathrm{d}t\right\}^{1/2},$$

and describe the functions $f$ for which equality holds. $\square$

The following relation is known as the *triangle inequality*.

***1.11 Theorem*** For any $x, y$ in an inner product space $V$,

$$\|x + y\| \leqslant \|x\| + \|y\|.$$

*Proof.* We have (compare the proof of Theorem 1.9)

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}(x,y) + \|y\|^2 \\ &\leqslant \|x\|^2 + 2|(x,y)| + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned} \qquad \square$$

***1.12 Exercise*** By applying 1.11 to $\mathbb{C}^k, k = 1, 2, \ldots$, show that $\ell^2$ is closed under addition. $\square$

***1.13 Theorem*** (the parallelogram law) For vectors $x, y$ in an inner product space,

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

*Proof.* We have

$$\|x+y\|^2 = \|x\|^2 + (x,y) + (y,x) + \|y\|^2,$$
$$\|x-y\|^2 = \|x\|^2 - (x,y) - (y,x) + \|y\|^2.$$

Add. $\square$

The reader is recommended to draw a diagram in order to see the reason for the name of this relation.

We have defined the norm in an inner product space in terms of the inner product; it has some significance that if we know how to calculate the norm of any vector then we can recover the inner product. This is because of the following result, called the *polarization identity*, the proof of which is a simple exercise.

***1.14 Theorem*** For any $x, y$ in an inner product space,

$$4(x, y) = \|x+y\|^2 - \|x-y\|^2 + \mathrm{i}\|x+\mathrm{i}y\|^2 - \mathrm{i}\|x-\mathrm{i}y\|^2. \qquad \square$$

Note that the polarization identity can be written

$$(x, y) = \frac{1}{4}\sum_{n=0}^{3} \mathrm{i}^n \|x + \mathrm{i}^n y\|^2.$$

***1.15 Exercise*** Let $H, K$ be inner product spaces and let $U: H \to K$ be a linear mapping such that $\|Ux\| = \|x\|$ for all $x \in H$. Prove that

$$(Ux, Uy) = (x, y) \qquad \text{for all } x, y \in H. \qquad \square$$

In the coming chapters we shall need a supply of examples on which to try out the concepts and results we shall meet. So far the only inner product spaces we know are $\mathbb{C}^n$, $\ell^2$, $\mathbb{C}^{m \times n}$ and $C[0, 1]$. Calculations in the first three are often simple, but do not exhibit effectively the interaction of inner product properties with other mathematical structure. On the other hand, $C[0, 1]$ has the disadvantage that a general continuous function is a rather intangible entity. An ideal source of exercises is to be found in inner product spaces of rational functions, relatively concrete objects for which the reader will already have some intuition. Calculation in these spaces is often quite easy, particularly when use is made of basic complex analysis (principally Cauchy's integral formula, the residue theorem and some elementary facts about power series expansions). They illustrate the power of Hilbert space methods in complex analysis and lead on to the topic of the last quarter of the book, where we shall see that spaces of rational functions are of practical importance in a range of engineering applications.

***1.16 Examples*** $RL^2$ denotes the space of rational functions which are analytic on the unit circle

$$\partial\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\},$$

with the usual addition and scalar multiplication and with the inner product

$$(f, g) = \frac{1}{2\pi \mathrm{i}} \int_{\partial\mathbb{D}} f(z)\overline{g(z)} \frac{\mathrm{d}z}{z}, \qquad (1.4)$$

the integral being taken anti-clockwise round $\partial\mathbb{D}$.

$RH^2$ is the subspace of $RL^2$ consisting of those rational functions which are analytic on the closed unit disc clos $\mathbb{D}$, where

$$\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\},$$

with inner product given by (1.4).

Thus a rational function (i.e. a ratio of two polynomials with complex coefficients) belongs to $RL^2$ provided it has no pole of modulus 1 and belongs to $RH^2$ if it has no pole of modulus less than or equal to 1. The spaces $L^2$ and $H^2$ we shall meet later: the prefix '$R$' stands for 'rational'. This is a convention used in electrical engineering.

Let us check that (1.4) does define an inner product on $RL^2$. Axioms (i) to (iii) in Definition 1.2 are immediate; we must show that $(f,f)>0$ when $f \neq 0$. On parametrizing $\partial\mathbb{D}$ by $z=e^{i\theta}$, $-\pi<\theta\leqslant\pi$, we can re-write (1.4) in the form

$$(f,g)=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,d\theta,$$

so that, in particular

$$(f,f)=\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{i\theta})|^2\,d\theta.$$

Since $f(e^{i\theta})$ is continuous on $[-\pi,\pi]$, the right hand side is positive unless $f=0$.

To illustrate how Cauchy's integral formula simplifies calculations in $RL^2$ let us work out the inner product of the functions

$$f(z)=\frac{1}{z-\alpha}, \qquad g(z)=\frac{1}{z-\beta}$$

where $|\alpha|<1$, $|\beta|<1$.

$$(f,g)=\frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{1}{z-\alpha}\cdot\frac{1}{\bar{z}-\bar{\beta}}\cdot\frac{dz}{z}.$$

Since $z\bar{z}=1$ on $\partial\mathbb{D}$, this yields

$$(f,g)=\frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{1}{z-\alpha}\cdot\frac{1}{1-\bar{\beta}z}\,dz$$
$$=\frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{h(z)}{z-\alpha}\,dz$$

where $h(z)=(1-\bar{\beta}z)^{-1}$. Since $|\beta|<1$, $h$ is analytic on clos $\mathbb{D}$, and so Cauchy's integral formula applies to give

$$(f,g)=h(\alpha)$$
$$=\frac{1}{1-\bar{\beta}\alpha}. \qquad \square$$

Another worthy inner product space is $W[a,b]$, described in Problem 1.2. This is a representative of a whole class of inner product spaces central to the theory of differential equations.

We turn next to the development of the properties of the metric $\|x-y\|$ in an inner product space. In fact it will pay us to begin with a more general type of space.