THE LEBESGUE INTEGRAL

Unit 12
Hilbert Space
THE LEBESGUE INTEGRAL

Unit 12
Hilbert Space

Prepared by the Course Team
Set Book

It is essential to have this book; the course is based on it and will not make sense without it.

This unit is based on material in Chapter 7 of the set book, but there are no set reading passages other than Theorem 7.5.2, page 191. For this reason the correspondence text of this unit is unusually long.

Bibliography

The following book is referred to quite frequently, and is useful though not essential.

Conventions

Before starting work on this text, please read the Guide to the Course.

The set book is referred to as *Weir*, and the above book *Calculus*, by M. Spivak, is referred to as *Spivak*. 

The Open University, Walton Hall, Milton Keynes, MK7 6AA.


First published as M331 1992.

Copyright © 1992 The Open University

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without written permission from the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd of 90 Tottenham Court Road, London, W1P 9HE.

Edited, designed and typeset by the Open University using the Open University TeX System.

Printed in the United Kingdom by Whitstable Litho Printers Limited, Whitstable, Kent.

ISBN 0 7492 2077 5

This text forms part of an Open University Fourth Level Course. If you would like a copy of *Studying with The Open University*, please write to the Central Enquiry Service, PO Box 200, The Open University, Walton Hall, Milton Keynes, MK7 6YX. If you have not already enrolled on the Course and would like to buy this or other Open University material, please write to Open University Educational Enterprises Ltd, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, United Kingdom.
Contents

Introduction 4

1 Inner product spaces 6
   1.1 Norms and inner products 6
   1.2 Orthonormal sets of vectors 10
   1.3 The Space $\ell^2$ 14

2 Hilbert space 18
   2.1 Infinite bases 18
   2.2 Fourier series 24
   2.3 Hilbert space 27

3 The Hilbert space $L^2(R)$ 32

4 Summary of the text 34
   4.1 Notation 35
   4.2 Glossary 35
   4.3 Results 35
   4.4 Further reading 38

5 Self-Assessment Questions 39
   5.1 Euclidean spaces 39
   5.2 Hilbert space 39

   Solutions to Self-Assessment Questions 40

6 Appendix 43
Introduction

In this unit we continue preparing for the discussion of Fourier series in Unit 13 by developing the right setting for this discussion, namely Hilbert Space. We shall define a Hilbert space, and show that in such a space we can develop an abstract theory of Fourier series. To see how to do this, we first briefly discuss the approach to Fourier series which pre-dated Lebesgue integration. Mathematicians were then restricted to deal with functions defined and continuous on the closed interval \([-x, x]\) except for a finite number of jump discontinuities. To be precise, this means that we can partition \([-x, x]\) into a finite number of closed intervals \([-x, a_1], [a_1, a_2], \ldots, [a_n, x]\) such that \(f\) is continuous on \((-x, a_1), (a_1, a_2), \ldots, (a_n, x)\) and all limits

\[
\lim_{\varepsilon \to 0^+} f(a_i - \varepsilon), \quad \lim_{\varepsilon \to 0^+} f(a_i + \varepsilon), \quad 0 \leq i \leq n + 1,
\]

exist. We have put \(a_0 = -x\) and \(a_{n+1} = +x\) for convenience. These two limits at \(a_i\) will be equal if and only if \(f\) is continuous there. A step function is a good example. Such functions are said to be piecewise continuous on \([-x, x]\), and they constitute a vector space which is denoted by \(PC[-x, x]\). One considers series expansions of such functions relative to a basis for \(PC[-x, x]\), the most usual basis comprising the trigonometric functions defined by

\[
\begin{align*}
\cos kx, & \quad k \in \mathbb{Z}^+, k \geq 1, \\
\sin kx, & \quad k \in \mathbb{Z}^+, k \geq 1.
\end{align*}
\]

(We have yet to make clear precisely what is meant by a basis in this context.)

Starting with a function \(f\), piecewise continuous on the interval \([-x, x]\), one computes the Fourier coefficients:

\[
\begin{align*}
\hat{a}_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx \\
\hat{a}_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx, \quad k \in \mathbb{Z}^+, k \geq 1 \\
\hat{b}_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx, \quad k \in \mathbb{Z}^+, k \geq 1.
\end{align*}
\]

The Fourier series for \(f\) is then defined to be the expression

\[
\frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx). \tag{2}
\]

Let us define a sequence of functions \(\{f_n\}\) as follows:

\[
\begin{align*}
f_0(x) &= \frac{1}{2} a_0, \\
f_{2n}(x) &= \frac{1}{2} a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \quad n \geq 1, \\
f_{2n+1}(x) &= f_{2n}(x) + a_{n+1} \cos((n+1)x), \quad n \geq 0,
\end{align*}
\]

for each \(x \in [-\pi, \pi]\). Thus \(\{f_n\}\) is the sequence of partial sums of the Fourier series for \(f\), and the relation between \(f\) and this series is connected with the convergence properties of the corresponding sequence \(\{f_n\}\).

One finds that, in general, the sequence \(\{f_n\}\) does not converge pointwise to \(f\), but that it does converge to \(f\) ‘in the mean’. That is, \(\{f_n\}\) converges to \(f\) in the norm defined by

\[
\|f\| = \left( \int_{-\pi}^{\pi} f^2 \, dx \right)^{1/2},
\]
meaning that
\[ \lim_{n \to \infty} \int_{-\pi}^{\pi} (f_n(x) - f(x))^2 \, dx = 0. \]

Before the work of Lebesgue, mathematicians were restricted to discussing Fourier series of piecewise continuous functions by the limitations of the integration theory they had available. Now that we have the Lebesgue integral at our disposal, we naturally seek to extend these ideas to a wider class of functions. This is readily done, but we leave the details to Unit 13. For the present, we observe that the norm defined above for the space \( PC[-\pi, \pi] \) resembles the norm which we introduced for the space \( L^2 \) in Unit 11.

**Convergence and normed spaces.**

We shall see in this unit that \( L^2 \) is indeed the appropriate space in which to extend our study of Fourier series. However, our main aim in this unit is to consider Fourier series in a very general context. To what extent can we describe something like a Fourier series in a general normed space? To try to answer this question, we begin by phrasing what we are doing with Fourier series in vector space terminology.

Now the norm
\[ f \mapsto \left( \int_{-\pi}^{\pi} f^2 \right)^{1/2} \quad (f \in PC[-\pi, \pi]) \]
derives from the inner product defined by
\[ (f, g) \mapsto \int_{-\pi}^{\pi} f \cdot g = \int_{-\pi}^{\pi} f \cdot g \quad (f, g \in PC[-\pi, \pi]) \]
(we shall write \( f \cdot g \) for the inner product of the functions \( f \) and \( g \)). Thus \( PC[-\pi, \pi] \) is a Euclidean space.

In this space we have a particular set of vectors (functions), namely
\[
\begin{align*}
\mathcal{E}_1 : x &\mapsto \frac{1}{\sqrt{2\pi}}, \\
\mathcal{E}_2 : x &\mapsto \frac{\sin x}{\sqrt{\pi}}, \\
\mathcal{E}_3 : x &\mapsto \frac{\cos x}{\sqrt{\pi}}, \\
\mathcal{E}_4 : x &\mapsto \frac{\sin 2x}{\sqrt{\pi}}, \\
\mathcal{E}_5 : x &\mapsto \frac{\cos 2x}{\sqrt{\pi}}, \ldots \\
&\quad \ldots \quad \ldots 
\end{align*}
\]

The set \( \{ \mathcal{E}_i : i = 1, 2, \ldots \} \) has the following property:
\[
\mathcal{E}_i \cdot \mathcal{E}_j = 0, \quad \text{if } i \neq j, \\
\mathcal{E}_i \cdot \mathcal{E}_i = 1, \quad i = 1, 2, \ldots.
\]

Such a set of vectors is said to be orthonormal. Notice that the Fourier coefficients in (1) are of the form \( (1/\sqrt{\pi}) \mathcal{E}_i \) for \( i \geq 2 \), and \( a_0 = (1/\sqrt{2\pi}) \mathcal{E}_1 \), and that the series (2) is equivalent to
\[ (f \cdot \mathcal{E}_1) \mathcal{E}_1(x) + (f \cdot \mathcal{E}_2) \mathcal{E}_2(x) + \cdots. \]

In this unit we shall not be interested in the particular vector space \( PC[-\pi, \pi] \), but we shall consider a general Euclidean space, \( E \), in which we have a countable orthonormal set of vectors \( \{ \mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \ldots \} \), and series of the form
\[ \sum_{i=1}^{\infty} a_i \mathcal{E}_i \quad (a_i \in \mathbb{R}). \]
We shall investigate the following questions.

(a) Under what conditions do we expect a series of the above type to converge?

(b) For any vector \( v \in E \), can we find coefficients \( a_i \) such that the series \( \sum_{i=1}^{\infty} a_i e_i \) converges to \( v \)?

(c) If \( \sum_{i=1}^{\infty} a_i e_i \) does converge to \( v \), are the coefficients \( a_i \) necessarily of the form \( v \cdot e_i \)?

The structure of this unit is as follows. We begin by discussing inner product spaces, and we obtain a criterion which tells us when a normed space is an inner product space. In Section 1 we shall introduce the space \( l^2 \) which is an infinite-dimensional analogue of the Euclidean space \( \mathbb{R}^k \). The elements of \( l^2 \) are sequences of real numbers, \( (a_1, a_2, \ldots) \), for which the series \( \sum_{i=1}^{\infty} a_i^2 \) converges. In Section 2 we shall discuss the concept of a basis of an infinite-dimensional normed space, and then look at generalized Fourier series in terms of such a basis (which we call a Schauder basis). We discover that the ideal setting for an abstract theory of Fourier series is a Euclidean space which is complete (as defined in Unit 11). Such a space is a Hilbert space. We then prove that every Hilbert space which is separable has a Schauder basis and is isomorphic to \( l^2 \). Finally we consider the familiar space \( L^2 \) and prove that this is a separable Hilbert space.

We feel that the material in this unit is interesting, advanced and challenging. It provides a tightly knitted scheme of analysis of the consequences of providing a vector space with an inner product. Do not expect to read straight through this unit and understand everything straight away, even if you have done so up to now. To ease your burden we have made a number of proofs optional but the usual advice about such things holds. Is the effort worth your while? We think so. One of the goals to set yourself is to see how mathematicians go about analysing a new structure. There are consequences of the hypotheses, examples, counter-examples, and considerations of existence and uniqueness. It is these last two items which will be hardest to assimilate, but persistence brings its rewards!

## 1 Inner product spaces

### 1.1 Norms and inner products

In Unit 11 we looked in some detail at various types of convergence and, in particular, at convergence in mean, that is, relative to a norm. Here we shall be more interested in the origin of these norms and, as we have seen in the Introduction that a norm can be defined from an inner product, we begin by considering inner products.

We have met inner products, previously, and we now recall their definition.

**Definition T.12.1.1** An inner product on a real vector space \( V \) is a function from \( V \times V \) to \( \mathbb{R} \) which associates with the pair of vectors \( (v, w) \) the real number \( v \cdot w \) in such a way that:

(a) \( (v_1 + v_2) \cdot w = v_1 \cdot w + v_2 \cdot w \quad (v_1, v_2, w \in V) \);

(b) \( v \cdot w = w \cdot v \quad (v, w \in V) \);

(c) \( (\lambda v) \cdot w = \lambda(v \cdot w) \quad (\lambda \in \mathbb{R}; v, w \in V) \);

(d) \( v \cdot v \geq 0 \), and \( v \cdot v = 0 \) if and only if \( v = 0 \) \((v \in V)\).

A real vector space supplied with an inner product is called a Euclidean space.

An immediate consequence of this definition is the following.
Lemma T.12.1.1 If \( V \) is a Euclidean space, then \( V \) is a normed space with respect to the norm defined by

\[
\|v\| = \sqrt{v \cdot v} \quad (v \in V).
\]

The verification that this is a norm satisfying the conditions \( N1, N2', N3 \) and \( N4 \) (in the Introduction to the previous unit) is straightforward, but it requires the Schwarz inequality which we prove in Exercise 1 below. For this reason we have left the verification to you, as SAQ 4.

There are many advantages in having a norm that derives from an inner product, not least of which is that we can introduce the concepts of orthogonal and orthonormal sets of vectors. However, there are many norms in which we are interested that do not derive from any inner product. For example, we shall see that convergence in \( L^1 \) is relative to a norm which does not derive from an inner product. It would seem useful, then, to have some way of distinguishing between those norms which do derive from an inner product and those which do not.

Theorem T.12.1.1 If \( V \) is a vector space with a norm derived from an inner product, then the Parallelogram Law holds; that is, for any \( v, w \in V \),

\[
\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + \|w\|^2.
\]

This corresponds to the theorem in \( \mathbb{R}^2 \) that says that the sum of the squares of the lengths of the diagonals of a parallelogram is twice the sum of the squares of the lengths of the two adjacent sides.

![Parallelogram Law](image)

Figure 1

Proof Since the norm is derived from an inner product, we have

\[
\|v + w\|^2 + \|v - w\|^2 = (v + w) \cdot (v + w) + (v - w) \cdot (v - w) \\
= 2(v \cdot v + w \cdot w) \\
= 2(\|v\|^2 + \|w\|^2).
\]

In fact the Parallelogram Law is not only a necessary condition for a norm to derive from an inner product; it is also a sufficient condition. Given that the Parallelogram Law holds in a normed space \( V \), if we put

\[
v \cdot w = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2),
\]

then we can show that this defines an inner product on \( V \). However, the proof is a little tricky and, as we can manage without this result, we omit it.

Example 1 In Unit 11 we saw that \( L^1 \) is a normed space with respect to the norm defined by

\[
\|f\| = \int |f| \quad (f \in L^1).
\]

We have claimed that this norm does not derive from an inner product. We can now justify this claim by exhibiting two functions \( f, g \in L^1 \) for which Equation (1) does not hold.
If we take $f = \chi_{[0,1]}$, $g = \chi_{[2,3]}$, then

$$\|f\| = \|g\| = 1 \quad \text{and} \quad |f + g| = |f - g| = f + g,$$

so

$$\|f + g\| = \|f - g\| = \|f\| + \|g\| = 2.$$

Hence,

$$\|f + g\|^2 + \|f - g\|^2 = 8$$

whilst

$$2(\|f\|^2 + \|g\|^2) = 4,$$

so the Parallelogram Law does not hold in the $L^1$-norm, and this norm cannot derive from an inner product.

Exercise 1

Show that Schwarz's inequality, namely

$$|x \cdot y| \leq \|x\| \times \|y\|,$$

holds in any Euclidean space. In many texts this is known as the Cauchy–Schwarz inequality.

Solution For any real number $\alpha$,

$$(x - \alpha y) \cdot (x - \alpha y) \geq 0,$$

that is,

$$\|x\|^2 - 2\alpha(x \cdot y) + \alpha^2\|y\|^2 \geq 0.$$

The corresponding quadratic equation in $\alpha$,

$$\|y\|^2\alpha^2 - 2(x \cdot y) + \|x\|^2 = 0,$$

has, therefore, either no real roots or a repeated real root. That is, the discriminant of this quadratic must be less than or equal to zero, so

$$[2(x \cdot y) - 4\|x\|^2 \times \|y\|^2] \leq 0.$$

This implies that

$$(x \cdot y)^2 \leq \|x\|^2 \times \|y\|^2,$$

or, because $\|x\|$, $\|y\|$ are positive,

$$|x \cdot y| \leq \|x\| \times \|y\|.$$

Exercise 2

Show that the norm defined by

$$\|f\|_3 = \left[ \int |f|^3 \right]^{1/3}$$

in $L^3$ is not derived from an inner product.

Solution The same choice of functions as in Example 1 shows that the Parallelogram Law does not hold in $L^3$. Taking $f = \chi_{[0,1]}$, $g = \chi_{[2,3]}$, we have

$$|f|^3 = f^3 = f \quad \text{and} \quad |g|^3 = g^3 = g,$$

so

$$\|f\|_3 = \|g\|_3 = 1.$$

Also

$$|f + g|^3 = |f - g|^3 = f + g,$$

so

$$\|f + g\|_3 = \|f - g\|_3 = 2^{1/3}.$$
Now \[\|f + g\|_3^2 + \|f - g\|_3^2 = 2(4)^{1/3}\]
and
\[2(\|f\|_3^2 + \|g\|_3^2) = 4.\]
Hence
\[\|f + g\|_3^2 + \|f - g\|_3^2 \neq 2(\|f\|_3^2 + \|g\|_3^2)\]
so the Parallelogram Law fails to hold in \(L^3\), and therefore the norm in \(L^3\) does not derive from an inner product.

**Exercise 3**

Let \(\ell^\infty\) denote the vector space of all bounded sequences of real numbers
\[a = (a_1, a_2, \ldots),\]
with no assumption of convergence, with the obvious componentwise addition and multiplication by scalars. (The superscript \(\infty\) is conventional and should not be taken to signify anything 'large'.)

Let
\[\|a\| = \sup\{|a_i|\}.\]
Verify that this defines a norm on \(\ell^\infty\) and that this norm does not derive from an inner product. (The boundedness condition means that, for each sequence \(a = (a_1, a_2, \ldots)\),
there is some positive real number \(K\) such that
\[-K \leq a_i \leq K \quad \text{for all } i.\]
The value of \(K\) will obviously vary from sequence to sequence.)

**Solution**

Checking that we have a norm is quite straightforward. The boundedness condition ensures that
\[\|a\| = \sup\{|a_i|\}\]
exists. Because each \(|a_i| \geq 0\), we must have
\[\sup\{|a_i|\} \geq 0,\]
so
\[\|a\| \geq 0.\]
The zero in the vector space, \(0 = (0, 0, \ldots)\), is clearly such that
\[\|0\| = 0,\]
and, conversely, if \(\|a\| = 0\), we have
\[\sup\{|a_i|\} = 0,\]
which forces \(|a_i|\) to be 0 for all \(i\). That is, each \(a_i = 0\), and so \(a = 0\). Also,
\[\|\lambda a\| = \sup\{|\lambda a_i|\} = \sup\{|\lambda| \times |a_i|\} = |\lambda| \sup\{|a_i|\} = |\lambda| \|a\|,\]
as required. Finally,
\[\|a + b\| = \sup\{|a_i + b_i|\} \leq \sup\{|a_i| + |b_i|\} \leq \sup\{|a_i|\} + \sup\{|b_i|\} = \|a\| + \|b\|.\]
Hence we have a norm on \(\ell^\infty\), and again all we have to do to show that this norm does not derive from an inner product is to produce two sequences in \(\ell^\infty\) for which the Parallelogram Law does not hold.
We take 
\[ a = (1, 0, 0, \ldots) \text{ and } b = (0, 1, 0, \ldots), \]
so that 
\[ a + b = (1, 1, 0, 0, \ldots) \text{ and } a - b = (1, -1, 0, 0, \ldots). \]
Clearly, 
\[ ||a|| = ||b|| = ||a + b|| = ||a - b|| = 1, \]
and 
\[ ||a + b||^2 + ||a - b||^2 = 2, \]
which is not equal to 
\[ 2(||a||^2 + ||b||^2) = 4, \]
so the Parallelogram Law does not hold.

### 1.2 Orthonormal sets of vectors

One advantage of an inner product on a vector space \( V \) is that it enables us to single out special sets of vectors in \( V \) which have particularly convenient properties.

**Definition T.12.1.2(1)**

(a) Two vectors \( u, v \) in a Euclidean space are orthogonal if \( u \cdot v = 0 \).
(b) A set \( A \) of vectors in a Euclidean space is orthogonal if any two distinct vectors \( u, v \) in \( A \) are orthogonal.
(c) A set \( A \) of vectors in a Euclidean space is orthonormal if \( A \) is an orthogonal set of vectors and \( ||v|| = 1 \) for all \( v \in A \). We call the condition \( ||v|| = 1 \) normalization, and say that \( v \) is normalized, or a unit vector.

In this course much of our interest is in infinite orthogonal or orthonormal sets of vectors; because many of the results we shall obtain mirror the results for finite sets, this is an appropriate place to recall and extend some of those results.

One important property of sets of vectors is that of linear independence. We now give a definition of this property which applies to both finite and infinite sets. This enables us to define what we mean by a Hamel basis for any vector space. This was the sort of basis considered in earlier courses, for example M203. In the next section we shall introduce another sort of basis, a Schauder basis, which is much more useful for infinite dimensional normed spaces. Note that, as a Hamel basis only involves finite sums, no questions of convergence arise.

**Definition T.12.1.1(2)** A set \( A \) of vectors in a vector space \( V \) is linearly independent if, for every finite subset \( \{v_1, v_2, \ldots, v_n\} \) of \( A \),
\[
\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0 \quad (\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R})
\]
implies
\[
\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.
\]
If \( A \) is not linearly independent we say that it is linearly dependent. If \( A \) contains the zero vector, it is linearly dependent. A one-point set \( A = \{v\}, v \neq 0 \), is linearly independent.
The span of a set $A$ of vectors in $V$ is the set of all finite linear combinations of vectors from $A$. The set $A$ is said to be spanning for $V$ if the span of $A$ coincides with $V$.

A Hamel basis for $V$ is a set $A$ of vectors in $V$ which is both spanning and linearly independent.

We can demonstrate a connection between linear independence and orthogonality. Recall that two vectors, $v, w$, are orthogonal if $v \cdot w = 0$.

**Lemma T.12.1.2** An orthogonal set of nonzero vectors in a Euclidean space $V$ is linearly independent.

**Proof** Let $A$ be an orthogonal set of nonzero vectors of $V$, and let \( \{b_1, b_2, \ldots, b_n\} \) be any finite non-empty subset of $A$. Suppose that

\[
\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n = 0.
\]

For each $i = 1, 2, \ldots, n$, we have

\[
0 = (\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n) \cdot b_i = \lambda_1 (b_1 \cdot b_i) + \lambda_2 (b_2 \cdot b_i) + \cdots + \lambda_n (b_n \cdot b_i) = \lambda_i ||b_i||^2
\]

since the $b_j, j = 1, 2, \ldots, n$, are orthogonal. Now, for each $i, b_i \neq 0$, so $||b_i|| \neq 0$ and therefore

\[
\text{if } \lambda_i ||b_i||^2 = 0, \text{ then } \lambda_i = 0.
\]

We have shown that

\[
\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n = 0
\]

implies

\[
\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0;
\]

that is, $A$ is linearly independent.

Another interesting situation we shall be looking at later is the Schauder basis extension of the case where we have an orthonormal Hamel basis in a finite-dimensional Euclidean space. Recall that a set $\{v_1, \ldots, v_n\}$ is orthonormal if $v_i \cdot v_j = 0$ if $i \neq j$ and $v_i \cdot v_i = 1$. In [M203] we proved part (a) of the following theorem for the space $\mathbb{R}^3$.

**Theorem T.12.1.2** Let $V$ be a real finite-dimensional Euclidean space with an orthonormal Hamel basis $\{e_1, e_2, \ldots, e_k\}$.

(a) Each $v \in V$ can be expressed in the form

\[
v = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_k e_k,
\]

where the coefficients are unique, and given by

\[
\lambda_i = v \cdot e_i, \quad i = 1, \ldots, k.
\]

(b) For each $v, w \in V$,

\[
v \cdot w = \sum_{i=1}^{k} (v \cdot e_i)(w \cdot e_i);
\]

in particular,

\[
||v||^2 = \sum_{i=1}^{k} (v \cdot e_i)^2.
\]
(c) If $\mathbb{R}^k$ is considered as a Euclidean space with respect to the standard inner product

$$ (a_1, a_2, \ldots, a_k) \cdot (b_1, b_2, \ldots, b_k) = \sum_{i=1}^{k} a_i b_i, $$

then there is an isomorphism $\theta : V \rightarrow \mathbb{R}^k$ given by

$$ \theta(v) = (v \cdot e_1, v \cdot e_2, \ldots, v \cdot e_k), $$

which preserves inner products, so that

$$ v \cdot w = \theta(v) \cdot \theta(w). $$

**Proof**

(a) Each vector $v \in V$ may be expressed as a sum of the form

$$ v = \lambda_1 e_1 + \cdots + \lambda_k e_k, $$

because $\{e_1, \ldots, e_k\}$ is a basis. But it is an orthonormal basis, so

$$ v \cdot e_1 = \lambda_1 e_1 \cdot e_1 + \cdots + \lambda_k e_k \cdot e_1 = \lambda_1, $$

$$ v \cdot e_2 = \lambda_1 e_1 \cdot e_2 + \lambda_2 e_2 \cdot e_2 + \cdots + \lambda_k e_k \cdot e_2 = \lambda_2, $$

and so on. This shows that the coefficients must be as given,

$$ \lambda_i = v \cdot e_i, $$

for all $i = 1, 2, \ldots, k$. At the same time, this computation proves the uniqueness of the $\lambda_i$.

(b) Write

$$ v = \sum_{i=1}^{k} (v \cdot e_i) e_i $$

in accordance with the result of part (a). Then

$$ v \cdot w = [(v \cdot e_1) e_1 + \cdots + (v \cdot e_k) e_k] \cdot w $$

$$ = (v \cdot e_1)(e_1 \cdot w) + \cdots + (v \cdot e_k)(e_k \cdot w). $$

Using the fact that

$$ w \cdot e_i = e_i \cdot w, \quad i = 1, 2, \ldots, k, $$

the first part of (b) is valid. Next put $v = w$ in this result, and the second part of (b) is clearly true.

(c) We define $\theta$ as given in the statement of part (c). First of all, $\theta$ is a linear mapping. To show this, consider any vectors $v, w$ in $V$ and any real numbers $a, b$. Then

$$ (av + bw) \cdot e_i = a(v \cdot e_i) + b(w \cdot e_i) $$

for all $i = 1, 2, \ldots, k$. Then

$$ \theta(av + bw) = (a(v \cdot e_1) + b(w \cdot e_1), \ldots, a(v \cdot e_k) + b(w \cdot e_k)) $$

$$ = (a(v \cdot e_1), \ldots, a(v \cdot e_k)) + (b(w \cdot e_1), \ldots, b(w \cdot e_k)) $$

$$ = a\theta(v) + b\theta(w). $$

Secondly, it is one-to-one (or one-one or injective). This means that if

$$ \theta(v) = \theta(w), $$

then $v = w$. Because $\theta$ is linear, this is equivalent to saying that

$$ \theta(v) = 0 $$

must imply $v = 0$. This is the case; for $\theta(v) = 0$ can be rewritten in the form

$$ (v \cdot e_1, \ldots, v \cdot e_k) = (0, \ldots, 0). $$
This is none other than the \( k \) equations
\[ v \cdot e_i = 0, \quad i = 1, 2, \ldots, k. \]
Using the second part of (b) above, this implies \( \|v\| = 0 \), so that \( v = 0 \) is the zero vector.

Thirdly, \( \theta \) is onto (or surjective). This means that if
\[ a = (a_1, a_2, \ldots, a_k) \]
is any vector in \( \mathbb{R}^k \), there is a vector \( v \) in \( V \) such that
\[ \theta(v) = a. \]
Given \( a \), then, define a vector \( v \) by setting
\[ v = a_1 e_1 + a_2 e_2 + \cdots + a_k e_k; \]

note that the \( a_i \) are fixed by the choice of the vector \( a \in \mathbb{R}^k \). Part (a) above requires that
\[ a_i = v \cdot e_i, \quad i = 1, 2, \ldots, k. \]
Hence,
\[ \theta(v) = (v \cdot e_1, \ldots, v \cdot e_k) = (a_1, \ldots, a_k) = a. \]

We have now shown that \( \theta \) is an isomorphism between \( V \) and \( \mathbb{R}^k \) as vector spaces. It remains to show that \( \theta \) preserves inner products. Now let \( v, w \) be vectors in \( V \). Consider the \( \mathbb{R}^k \)-inner product
\[ (v \cdot e_1, \ldots, v \cdot e_k) \cdot (w \cdot e_1, \ldots, w \cdot e_k) \]
\[ = (v \cdot e_1)(w \cdot e_1) + \cdots + (v \cdot e_k)(w \cdot e_k) = v \cdot w \] from part (b).

We have now completed the proof.

Let us note the following comments. Part (c) can be rephrased as saying that Euclidean space of dimension \( k \) is essentially unique, and in this sense we shall find that a separable Hilbert space is essentially unique. When we come to discuss Hilbert space in detail, we shall find that we are developing analogues of the parts of this theorem; in Section 1.3 we begin this process by describing an infinite-dimensional analogue of \( \mathbb{R}^2 \) and its standard inner product.

There is one further result on finite orthonormal sets of vectors which we shall need later in this unit; something very much like this was shown for \( \mathbb{R}^3 \) in the last unit on linear algebra in 1203.

**Lemma T.12.1.2(2)** Let \( \{e_1, e_2, \ldots, e_n\} \) be an orthonormal set of vectors in a Euclidean space \( V \), and let \( v \in V \). Then
\[ \left\| v - \sum_{i=1}^{N} \lambda_i e_i \right\| \geq \left\| v - \sum_{i=1}^{n} (v \cdot e_i)e_i \right\|, \]
whenever \( N \leq n \), and \( \lambda_1, \ldots, \lambda_N \) are any real numbers. In less formal language, this tells us that the 'best approximation' to \( v \) in terms of \( e_1, \ldots, e_n \) is \( \sum_{i=1}^{n} (v \cdot e_i)e_i \).

**Proof** Let
\[ K = \left\| v - \sum_{i=1}^{N} \lambda_i e_i \right\|^2 - \left\| v - \sum_{i=1}^{n} (v \cdot e_i)e_i \right\|^2. \]

It is sufficient to show that \( K \geq 0 \).
Exercise 4

(a) Let $A_1, A_2, \ldots, A_n$ be non-null measurable subsets of $\mathbb{R}$. When is the set of functions \( \{\chi_{A_1}, \chi_{A_2}, \ldots, \chi_{A_n}\} \) orthogonal in $L^2$?

(b) Let $f_n = \chi_{[-n,n]}$ and $A = \{f_1, f_2, f_3, \ldots\}$. Is $A$ orthogonal or orthonormal in $L^2$?

Solution

(a) Because the $A_i$ are non-null and measurable, the $\chi_{A_i}$ are nonzero vectors in $L^2$. In $L^2$ we have

\[
\chi_{A_i} \cdot \chi_{A_j} = \int \chi_{A_i} \chi_{A_j} = \int \chi_{A_i \cap A_j}.
\]

Therefore the given set of functions is orthogonal in $L^2$ if and only if $A_i \cap A_j$ is null whenever $i \neq j$.

(b) Similarly, in $L^2$ we have, for $m > n$, that

\[
f_n \cdot f_m = \int f_n f_m = \int f_n = 2n \neq 0.
\]

Therefore $A$ is not orthogonal, and so not orthonormal, in $L^2$.

1.3 The Space $\ell^2$

The space we are going to discuss next can be regarded as an infinite-dimensional analogue of the space $\mathbb{R}^k$. Therefore it is extremely important to understand its properties. In order to do so we shall make extensive use of certain basic facts about infinite series. We have gathered these together in an appendix at the end of the unit. If you do not know the proofs, a straightforward account will be found in K. Knopp, *Infinite Sequences and Series* (Dover Publications Inc. 1956).

If we want to emphasize that we have used a result $n$, say, from the appendix, we insert the symbol $[\text{App } n.]$.
Definition T.12.1.3  The space $\ell^2$ consists of all sequences of real numbers $a = (a_1, a_2, \ldots)$ for which the series $\sum_{i=1}^{\infty} a_i^2$ is convergent. Recall that this means that the sequence of partial sums $s_N = \sum_{i=1}^{N} a_i^2$ is convergent, $s_N \to s$. By convention we write $s = \sum_{i=1}^{\infty} a_i^2$.

Note two things. Firstly, the sequence $\{s_N\}$ is increasing, and secondly, it is sufficient for convergence that $\{s_N\}$ be a Cauchy sequence. We define addition and multiplication by a scalar on $\ell^2$ as follows:

$$a + b = (a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$

and

$$\lambda a = \lambda (a_1, a_2, \ldots) = (\lambda a_1, \lambda a_2, \ldots).$$

To check that $\ell^2$ is a vector space, all we have to check is closure with respect to these operations. For this we need the following result.

Lemma T.12.1.3  If the series $\sum_{i=1}^{\infty} a_i^2$, $\sum_{i=1}^{\infty} b_i^2$ are convergent, then the series $\sum_{i=1}^{\infty} a_i b_i$ is absolutely convergent.

Proof  For any $i \geq 1$, $0 \leq (|a_i| - |b_i|)^2$, and so $|a_i b_i| \leq \frac{1}{2} (a_i^2 + b_i^2)$. Thus

$$\sum_{i=1}^{n} |a_i b_i| \leq \frac{1}{2} \left( \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 \right) \leq \frac{1}{2} \left( \sum_{i=1}^{\infty} a_i^2 + \sum_{i=1}^{\infty} b_i^2 \right)$$

for all $n \geq 1$, and hence $\sum_{i=1}^{\infty} |a_i b_i|$ converges, as required [App 8].

We are now in a position to prove that $\ell^2$ is a vector space, and more besides.

Proposition T.12.1.3  $\ell^2$ is a Euclidean space, with inner product given by

$$a \cdot b = \sum_{i=1}^{\infty} a_i b_i, \quad (a, b \in \ell^2).$$

Proof  If $a, b \in \ell^2$, the series $\sum_{i=1}^{\infty} a_i^2$, $\sum_{i=1}^{\infty} a_i b_i$, $\sum_{i=1}^{\infty} b_i^2$ all converge, [App 9] the series

$$\sum_{i=1}^{\infty} (a_i + b_i)^2 = \sum_{i=1}^{\infty} (a_i^2 + 2a_i b_i + b_i^2)$$

also converges, and we conclude that $a + b \in \ell^2$. Clearly $\lambda a \in \ell^2$ whenever $\lambda \in \mathbb{R}$ and $a \in \ell^2$[App 9], and hence $\ell^2$ is a vector space.

Since $\sum_{i=1}^{\infty} a_i b_i$ is absolutely convergent for any $a, b \in \ell^2$, we can define $a \cdot b$ for any $a, b \in \ell^2$ by the above formula [App 11]. Standard properties of convergent series [Appendix] show us that

$$(a + b) \cdot c = \sum_{i=1}^{\infty} (a_i + b_i) c_i = \sum_{i=1}^{\infty} a_i c_i + \sum_{i=1}^{\infty} b_i c_i = a \cdot c + b \cdot c,$$

$$(\lambda a) \cdot b = \sum_{i=1}^{\infty} (\lambda a_i) b_i = \sum_{i=1}^{\infty} \lambda (a_i b_i) = \lambda \sum_{i=1}^{\infty} a_i b_i = \lambda (a \cdot b),$$

$$b \cdot a = \sum_{i=1}^{\infty} b_i a_i = \sum_{i=1}^{\infty} a_i b_i = a \cdot b,$$

for all $a, b, c \in \ell^2$ and $\lambda \in \mathbb{R}$. 

15
Moreover
\[ a \cdot a = \sum_{i=1}^{\infty} a_i^2 \geq 0 \]
for any \( a \in \ell^2 \). Also if \( a \in \ell^2 \) with \( a \neq 0 \), there is some \( j \geq 1 \) such that \( a_j \neq 0 \). Then \( \sum_{i=1}^{n} a_i^2 \geq a_j^2 > 0 \) for all \( n \geq j \), and hence \( a \cdot a = \sum_{i=1}^{\infty} a_i^2 \geq a_j^2 > 0 \). From this we see that if \( a \in \ell^2 \) and \( a \cdot a = 0 \), then \( a = 0 \).

Thus we have shown that \( \ell^2 \) is a Euclidean space.

**Corollary T.12.1.3** If \( a, b \in \ell^2 \) then
\[
\left| \sum_{i=1}^{\infty} a_i b_i \right| \leq \left( \sum_{i=1}^{\infty} a_i^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{\infty} b_i^2 \right)^{\frac{1}{2}}.
\]

**Proof** This is just the Cauchy-Schwarz inequality (Exercise 1) for the Euclidean space \( \ell^2 \) written out in full.

It is worth noting that we could have proved this form of the Cauchy-Schwarz inequality for \( \ell^2 \) directly, using standard theory of sequences and series. However, we chose not to do so because in many ways the above proof is much neater, and it also illustrates the power of abstract methods.

Knowing that \( \ell^2 \) is a Euclidean space, we can look for orthogonal and orthonormal sets of elements in \( \ell^2 \). An obvious set to consider is the following set, which we denote by \( \tilde{S} \). Let
\[
\begin{align*}
\tilde{e}_1 &= (1, 0, 0, \ldots), \\
\tilde{e}_2 &= (0, 1, 0, \ldots), \\
\vdots \\
\tilde{e}_i &= (0, 0, \ldots, 0, 1, 0, \ldots), \\
\end{align*}
\]
where \( \tilde{e}_i \) has 1 in the \( i \)-th place and 0 everywhere else. Then, clearly,
\[
\tilde{e}_i \cdot \tilde{e}_j = 0, \quad \text{if } i \neq j,
\]
\[
\tilde{e}_i \cdot \tilde{e}_i = 1,
\]
so \( \{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_i, \ldots\} \) is an orthonormal subset of \( \ell^2 \).

Our next result proves, in effect, that \( \tilde{S} = \{\tilde{e}_1, \tilde{e}_2, \ldots\} \) is a basis for \( \ell^2 \), but we do not state it in this form because we have not yet defined a Schauder basis of a normed space. (We shall do this in the next section.)

**Theorem T.12.1.3** If the vectors \( \tilde{e}_i \in \ell^2 \) are defined as above, and \( a \) is any vector in \( \ell^2 \), then the series
\[
\sum_{i=1}^{\infty} (a \cdot \tilde{e}_i) \tilde{e}_i
\]
converges to \( a \), in the norm derived from the inner product on \( \ell^2 \).

**Proof** Suppose \( a = (a_1, a_2, \ldots, a_n, \ldots) \). Then, for \( i = 1, 2, \ldots \),
\[
a \cdot \tilde{e}_i = a_i \quad \text{and} \quad a_i \tilde{e}_i = (0, 0, \ldots, a_i, \ldots).
\]
Now
\[
\left\| a - \sum_{i=1}^{n} a_i \tilde{e}_i \right\| = \|(0, 0, \ldots, 0, a_{n+1}, a_{n+2}, \ldots)\| = \left( \sum_{i=n+1}^{\infty} a_i^2 \right)^{\frac{1}{2}}.
\]
Since \( a \in \ell^2 \), \( \sum_{i=1}^{\infty} a_i^2 \) is convergent, so
\[
\sum_{i=n+1}^{\infty} a_i^2 \to 0 \quad \text{as} \quad n \to \infty.
\]
Therefore
\[
\| a - \sum_{i=1}^{n} (a \cdot e_i) e_i \| \to 0 \quad \text{as} \quad n \to \infty,
\]
hence \( \sum_{i=1}^{\infty} (a \cdot e_i) e_i \) converges to \( a \), as required.

This result is analogous to part (a) of Theorem T. 12.1.2. For the analogue of part (b), see Exercise 5. We shall be seeing a lot more of the space \( \ell^2 \) later in this unit.

**Exercise 5**

Let
\[
a = (a_1, a_2, \ldots), \quad b = (b_1, b_2, \ldots)
\]
be any vectors in \( \ell^2 \). Show that
(a) \( a \cdot b = \sum_{i=1}^{\infty} (a \cdot e_i)(b \cdot e_i) \)
and
(b) \( \|a\|^2 = \sum_{i=1}^{\infty} (a \cdot e_i)^2 \).

**Solution**

For \( i = 1, 2, \ldots \), we have \( a \cdot e_i = a_i \) and \( b \cdot e_i = b_i \). Therefore
\[
a \cdot b = \sum_{i=1}^{\infty} a_i b_i = \sum_{i=1}^{\infty} (a \cdot e_i)(b \cdot e_i).
\]
Taking \( a = b \), we obtain
(b) \( \|a\|^2 = a \cdot a = \sum_{i=1}^{\infty} (a \cdot e_i)^2 \).

**Exercise 6**

Let
\[
e'_1 = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0, \ldots \right),
e'_2 = \left( 0, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0, \ldots \right),
\]
\[
\ldots
\]
\[
e'_i = \left( 0, 0, \ldots, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0, \ldots \right) = \frac{1}{\sqrt{2}}(e_{2i-1} + e_{2i}),
\]
\[
\ldots
\]
Show that \( \{e'_1, e'_2, \ldots\} \) is an orthonormal set, and that the statement of Theorem T.12.1.3. is false if we replace the orthonormal set \( \{e_1, e_2, \ldots\} \) by \( \{e'_1, e'_2, \ldots\} \).

**Solution**

It is straightforward to check that \( e'_1 \cdot e'_i = 1 \) and \( e'_1 \cdot e'_j = 0 \) if \( i \neq j \). To show that Theorem T.12.1.3 does not hold for \( \{e'_i\} \), it is sufficient to give an element \( a \in \ell^2 \) for which \( a \neq \sum_{i=1}^{\infty} (a \cdot e'_i)e'_i \).

The element \( a = (1, 0, 0, \ldots) \) will do, because \( a \cdot e'_1 = 1/\sqrt{2} \), and \( a \cdot e'_i = 0 \) if \( i > 1 \), so
\[
\sum_{i=1}^{\infty} (a \cdot e'_i)e'_i = (a \cdot e'_1)e'_1 = \left( \frac{1}{2}, \frac{1}{2}, 0, 0, \ldots \right) \neq a.
\]
2 Hilbert space

2.1 Infinite bases

In our search for the right setting in which to discuss Fourier series, we have been discussing Euclidean spaces which contain infinite orthonormal sets of vectors. In particular, we have seen that in $\ell^2$ we can express each vector $a$ as a series

$$\sum_{i=1}^{\infty} (a \cdot \hat{e}_i) \hat{e}_i,$$

where $\mathcal{S} = \{\hat{e}_1, \hat{e}_2, \ldots\}$ is the orthonormal set of vectors defined in Section 1.

However, we saw in Exercise 6 that the choice of the orthonormal set of vectors is crucial; it is not sufficient to take any orthonormal set in $\ell^2$ if we wish to express each element of $\ell^2$ as a series of the above form.

Our aim in this section is to give a name to, and develop some theory about, sets of orthonormal vectors with the desired property that every element of the Euclidean space has a series expansion of the above form in terms of the orthonormal vectors.

Any finite-dimensional Euclidean space $E$ has a Hamel basis, and we can apply the Gram–Schmidt orthogonalization process (see Weir: Theorem 7.5.2, page 191) to this Hamel basis to obtain an orthonormal Hamel basis for $E$. Theorem T.12.1.2 now comes into play, and so the theory of finite-dimensional Euclidean spaces is clear. For infinite-dimensional Euclidean spaces, the notion of Hamel basis is not particularly useful for our purposes; we need a new concept.

We begin by looking at normed spaces in general, and giving a definition of a new sort of basis, which is a generalization of the definition of a Hamel basis of a finite-dimensional vector space different from that given in Definition T.12.1.1(2).

**Definition T.12.2.1(1)** Let $N$ be a normed space and $\mathcal{S} = \{s_1, s_2, \ldots\}$ be a countable infinite subset of $N$. Then $\mathcal{S}$ is called a Schauder basis for $N$ if, for each $x \in N$, we can find unique real numbers $a_1, a_2, \ldots$ such that $\sum_{i=1}^{\infty} a_is_i$ converges in norm to $x$.

A little thought shows that this definition implies that any Schauder basis is linearly independent: the clue is in the uniqueness of the coefficients. Recall that for an infinite set of vectors, as $\mathcal{S}$ is, linear independence is defined to mean that any finite subset of vectors in $\mathcal{S}$ is linearly independent in the usual sense.

**Lemma T.12.2.1(1)** If $\mathcal{S} = \{s_1, s_2, \ldots\}$ is a Schauder basis in the normed space $N$, then $\mathcal{S}$ is linearly independent.

**Proof** Let $T = \{s_{i_1}, s_{i_2}, \ldots, s_{i_N}\}$ be a finite subset of $\mathcal{S}$, where $i_1 < i_2 < \ldots < i_N$. If $b_1, \ldots, b_N$ are real numbers with $\sum_{j=1}^{N} b_js_{i_j} = 0$, define real numbers $a_1, a_2, \ldots$, by setting $a_{i_j} = b_j$ for $1 \leq j \leq N$, and setting $a_i = 0$ for all other values of $i$. It is clear that

$$\sum_{i=1}^{n} a_is_i = \sum_{j=1}^{N} b_js_{i_j} = 0$$

for all $n > i_N$, and hence it follows that $\sum_{i=1}^{\infty} a_is_i$ converges in norm to 0. But it is clear that $\sum_{i=1}^{\infty} a_is_i$ converges in norm to 0, where $a_i = 0$ for all $i$. Hence it follows that $a_i = 0$ for all $i$, so that $b_1 = \ldots = b_N = 0$. Hence $T$ is linearly independent. Since this is true for any finite subset $T$ of $\mathcal{S}$, it follows that $\mathcal{S}$ is linearly independent.

Satisfying the uniqueness condition in the definition of a Schauder basis is crucial.

The following example, appropriate to Fourier series, illustrates this. The space in question is $L^2[-\pi, \pi]$ and

$$\mathcal{S} = \{1, x, \cos x, \sin x, \cos 2x, \sin 2x, \ldots\}.$$
This set is linearly independent, and, by a result we shall prove in Unit 13, any function in $L^2(-\pi, \pi)$ has an expansion in terms of $S$. However, the expansion is not unique. The function $f(x) = x$ has two expansions. One is $x = x$ and the other is in terms of the trigonometric functions only, without $x$. So $S$ is not a Schauder basis.

This was a trick example! In fact, the set

$$T = \{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots\}$$

is a Schauder basis for $L^2[-\pi, \pi]$, and the inclusion of the extra vector $x$ 'ruins things'. But you see from this how strong the requirement of uniqueness is.

We shall see later that, in a Euclidean space, if a Schauder basis $S = \{s_1, s_2, \ldots\}$ is orthonormal, then each coefficient $a_i$ in the series $\sum_{i=1}^{\infty} a_i s_i$ converging to $x$ must in fact be $x \cdot s_i$.

It is not very easy to give examples of Schauder bases, or even to show that they exist; however, we have seen in Theorem T.12.1.3 that the set $\{e_1, e_2, \ldots\}$ is a Schauder basis in $\ell^2$, and in Section 3 we shall see that $L^2$ contains a Schauder basis. Because of these difficulties, our next task is to develop an alternative way of deciding whether a normed space contains a Schauder basis.

Suppose that $S = \{s_1, s_2, \ldots\}$ is a Schauder basis in a normed space $N$ and $x \in N$; then we know that

$$x = \sum_{i=1}^{\infty} a_i s_i$$

for suitable choices of the $a_i$, convergence being with respect to the norm in $N$. That is, given $\varepsilon > 0$, there is an integer $K > 0$ such that

$$\|x - \sum_{i=1}^{n} a_i s_i\| < \varepsilon$$

for all $n \geq K$; in particular,

$$\|x - \sum_{i=1}^{K} a_i s_i\| < \varepsilon.$$

We can put this another way. If $\langle S \rangle$ is the set of all finite linear combinations of elements in $S$, that is, the span of $S$, then

$$s = \sum_{i=1}^{K} a_i s_i$$

is an element of $\langle S \rangle$, and $\|x - s\| < \varepsilon$. That is, for each $\varepsilon > 0$ we can find an element $s \in \langle S \rangle$ such that $\|x - s\| < \varepsilon$.

In other words, each $x \in N$ can be approximated arbitrarily closely by an element of $\langle S \rangle$. This suggests the following definition.

**Definition T.12.2.1(2)**

(a) Let $N$ be a normed space and $A$ be a subset of $N$; then $A$ is dense in $N$ if, for each $x \in N$ and $\varepsilon > 0$, there is an element $a$ in $A$ such that $\|x - a\| < \varepsilon$.

(b) A subset $C$ of $N$ is called a total set in $N$ if $\langle C \rangle$, the set of all finite linear combinations of elements from $C$, is dense in $N$.

(c) A normed space $N$ is called separable if it possesses a countable dense subset.

Our statement above can now be rephrased as follows: a Schauder basis in a normed space $N$ is a total subset of $N$. The concept of separability will be very useful later on, and we present a result now which will be of great use to us in identifying separable spaces. It involves some fairly delicate arguments concerning countability, arguments you most probably have not seen before. These are not really any more difficult than those you know from Weir, but are not properly part of this course. Rather they belong to set theory and logic. We have chosen to list the necessary set
theory results in the appendix, after the results for sequences and series. For proofs of these and further reading we suggest the following widely used text: P.R. Halmos, \textit{Naive Set Theory} (Springer-Verlag, 1987).

**Theorem T.12.2.1(1)** If \( N \) is a normed space containing a countably infinite total subset \( C \) (for example a Schauder basis), then \( N \) is separable.

**Proof.** (Optional) From the appendix we know that the \( n \)-fold Cartesian product \( \mathbb{Q}^n \) of the set \( \mathbb{Q} \) of rationals with itself is countable. Let \( C = \{c_1, c_2, \ldots\} \) be the countably infinite total subset of \( N \) referred to, and let \( \langle c_1, \ldots, c_n \rangle_\mathbb{Q} \) be the span of \( \{c_1, \ldots, c_n\} \), but using coefficients from \( \mathbb{Q} \) only: the \textit{rational span}. A typical element is of the form \( q_1c_1 + \cdots + q_nc_n \), with \( q_1, \ldots, q_n \in \mathbb{Q} \). We view this as a map
\[
g: \mathbb{Q}^n \to \langle c_1, \ldots, c_n \rangle_\mathbb{Q}, \quad g(q_1, \ldots, q_n) = q_1c_1 + \cdots + q_nc_n.
\]

This map is surjective, so the set \( \langle c_1, \ldots, c_n \rangle_\mathbb{Q} \) is countable, and this holds for any \( n \geq 1 \) [App 16].

The set \( C_\mathbb{Q} \) of all finite rational linear combinations of vectors from \( C \) can be written as
\[
C_\mathbb{Q} = \bigcup_{n \geq 1} \langle c_1, \ldots, c_n \rangle_\mathbb{Q},
\]
which is a countable union of countable sets. Therefore \( C_\mathbb{Q} \) is countable.

We now show that \( C_\mathbb{Q} \) is dense in \( N \), and so by Definition T.12.2.1(2), part (iii), \( N \) is separable. For any \( x \in N \) and any \( \varepsilon > 0 \), we may choose an element
\[
y = a_1c_1 + \cdots + a_nc_n \quad (a_i \in \mathbb{R})
\]
from \( \langle C \rangle \), which is within \( \varepsilon/2 \) of \( x \), as \( C \) is total
\[
||x - y|| < \varepsilon/2.
\]

Now \( \mathbb{Q} \) is dense in \( \mathbb{R} \), so we may choose rational numbers \( q_1, \ldots, q_n \) satisfying
\[
|a_i - q_i| < \varepsilon/(2K) \quad (1 \leq i \leq n),
\]
where \( K \) is the constant
\[
K = 1 + ||c_1|| + \cdots + ||c_n||.
\]

Then
\[
z = q_1c_1 + \cdots + q_nc_n
\]
is an element of \( \langle C \rangle_\mathbb{Q} \). We estimate \( ||y - z|| \) as follows:
\[
||y - z|| = \left\| \sum_{i=1}^{n} (a_i - q_i)c_i \right\| \leq \sum_{i=1}^{n} |a_i - q_i| ||c_i||
\]
\[
< \left[ \varepsilon/(2K) \right] \sum_{i=1}^{n} ||c_i|| < \varepsilon/2.
\]

Then \( z \) can be shown to be within \( \varepsilon \) of \( x \) as follows, using a standard technique to generate inequalities:
\[
||x - z|| = ||x - y + y - z||
\]
\[
\leq ||x - y|| + ||y - z|| < (\varepsilon/2) + (\varepsilon/2).
\]

What we have pointed out is that a Schauder basis \( S \) is total in \( N \); but what we would like is the converse. Suppose that \( C \) is a countable total subset of \( N \); then, given \( x \in N \) and \( \varepsilon > 0 \), we know that there is some integer \( n \in \mathbb{N} \) such that
\[
\left\| x - \sum_{i=1}^{n} \lambda_i c_i \right\| < \varepsilon \tag{1}
\]
for suitable real numbers \( \lambda_i \) and elements \( c_i \in C \), whereas we need a series \( \sum_{i=1}^{\infty} a_i c_i \) converging to \( x \), if \( C \) is to be a Schauder basis. The snag is that there is no
guarantee that we can choose the same $\lambda_i$'s in Equation (1) for different values of $\epsilon$. However, the converse does hold in the case of most interest to us.

Before we prove this result, we must first prove a lemma and its corollary. You will see that these are very useful results and of independent interest.

**Lemma T.12.2.1(2)** Let $(v_n)$ be a sequence of vectors in a Euclidean space $E$ converging to a vector $v$,

$$\lim_{n \to \infty} \|v_n - v\| = 0.$$

If $y$ is any vector in $E$, then

$$\lim_{n \to \infty} |v_n - y - v \cdot y| = 0.$$

We usually abbreviate this as

$$v_n \to v \text{ implies } v_n \cdot y \to v \cdot y.$$

**Proof** Using Schwarz's inequality from Section 1,

$$|v_n \cdot y - v \cdot y| \leq |v_n - v \cdot y| \leq \|v_n - v\| \|y\|.$$

Now $\|y\|$ is some fixed positive number and $\|v_n - v\|$ can be made as small as desired, so the right-hand side can be made less than any pre-assigned $\epsilon$ by taking $n$ large enough. This proves the result.

The corollary enables us to interchange the order of an infinite sum and an inner product. We remind you that to say that an infinite series of vectors in $E$ converges,

$$\sum_{i=1}^{\infty} x_i = s \text{ say},$$

means that the partial sums

$$s_N = \sum_{i=1}^{N} x_i$$

converge in $E$ as $N \to \infty$. In other words, $s_N \to s$ as $N \to \infty$.

**Corollary T.12.2.1** Let $(x_i)$ be a countably infinite family of vectors in a Euclidean space $E$ such that the series $\sum_{i=1}^{\infty} x_i$ converges. Then if $y \in E$ is any vector,

$$y \cdot \left( \sum_{i=1}^{\infty} x_i \right) = \sum_{i=1}^{\infty} (y \cdot x_i).$$

In particular, the series on the right-hand side converges.

**Proof** From the previous lemma,

$$y \cdot s_N \to y \cdot s,$$

using the notation above. But

$$y \cdot s_N = \sum_{i=1}^{N} y \cdot x_i.$$

As this is the partial sum of an infinite series of real numbers, and we know it converges (to $y \cdot s$), we are therefore justified in writing

$$\lim_{N \to \infty} \sum_{i=1}^{N} y \cdot x_i = \sum_{i=1}^{\infty} y \cdot x_i.$$

As this limit equals $y \cdot s$ and $s = \sum_{i=1}^{\infty} x_i$, this completes the proof.

The results of this lemma and its corollary are useful, as we said, because it is always convenient to interchange the order of operations, here the inner product and limits. For those who have learnt that a function $f: \mathbb{R} \to \mathbb{R}$ is continuous if whenever $t_n \to t$ in $\mathbb{R}$, then $f(t_n) \to f(t)$, we make the following remark. This interchange of the order, of limit and function, is equivalent to the $\epsilon - \delta$ definition of continuity.
in the very general setting of functions $f : E \rightarrow F$ between metric spaces $E$, $F$. Granting this, take $E$ to be our Euclidean space, $F$ to be $\mathbb{R}$ and $f(x) = y \cdot x$. Then we have just shown in the lemma that $f$ is continuous.

This next theorem gives a complete justification for introducing the strange concepts of orthonormal, total and Schauder basis, and brings them all together.

**Theorem T.12.2.1(2)** Let $S = \{e_1, e_2, e_3, \ldots \}$ be a countably infinite and total orthonormal set of vectors in a Euclidean space $E$. Suppose that for every vector $x \in E$ the infinite series

$$s(x) = \sum_{i=1}^{\infty} (x \cdot e_i) e_i$$

converges. Then $S$ is a Schauder basis and $s(x) = x$; that is,

$$x = \sum_{i=1}^{\infty} (x \cdot e_i) e_i.$$

**Proof (Optional)** For any index $k$, we can use the previous lemma, the orthogonality of $S$ and the linear structure of the inner product, to see that

$$s(x) \cdot e_k = \left( \sum_{i=1}^{\infty} (x \cdot e_i) e_i \right) \cdot e_k = \sum_{i=1}^{\infty} ((x \cdot e_i) e_i) \cdot e_k$$

$$= \sum_{i=1}^{\infty} (x \cdot e_i)(e_i \cdot e_k) = x \cdot e_k,$$

and so $(s(x) - x) \cdot e_k = 0$. Hence $(s(x) - x) \cdot y = 0$ for all $y \in (S)$. This will enable us to show that $s(x)$ equals $x$, as follows. Assume that $s(x) - x \neq 0$. As $(S)$ is dense in $E$, we can find a $y \in (S)$ which approximates $s(x) - x$ as closely as desired. We make the choice $\varepsilon = \|s(x) - x\|$, and choose $y$ so that

$$\|s(x) - x - y\| < \|s(x) - x\|.$$

Thus

$$||s(x) - x||^2 + ||y||^2 = ||s(x) - x||^2 - 2(s(x) - x) \cdot y + ||y||^2$$

$$= ||s(x) - x||^2 < ||s(x) - x||^2,$$

so $||y||^2 < 0$. This contradiction implies that $s(x) - x = 0$, and so $s(x) = x$.

Suppose that $a_1, a_2, \ldots$ are real numbers such that $\sum_{i=1}^{\infty} a_i e_i$ converges to $x$. Then

$$a_k = \sum_{i=1}^{\infty} a_i (e_i \cdot e_k) = \sum_{i=1}^{\infty} (a_i e_i) \cdot e_k$$

$$= \left( \sum_{i=1}^{\infty} a_i e_i \right) \cdot e_k = x \cdot e_k$$

for any index $k$. Thus any $x \in E$ has a unique series expansion in terms of $S$, given by the expression for $s(x)$, so $S$ is indeed a Schauder basis, and the result follows.

It is clear from the proof of uniqueness given above that if $S$ is an orthonormal basis in the Euclidean space $E$, then the special convergence condition introduced in this theorem is automatically satisfied, since the unique expansion of $x \in E$ in terms of $S$ must be the one given by $s(x)$, which therefore converges. Thus this special convergence condition is not merely sufficient, but also necessary, to guarantee that a countably infinite total orthonormal set of vectors is a Schauder basis.

We shall see later on that if the Euclidean space $E$ has the property of metric completeness, this convergence condition is automatically satisfied; this will yield a simplified version of Theorem T.12.2.1 (2).
Exercise 7

Let
\[ e''_1 = \left( \frac{1}{\sqrt{2}}, 1/\sqrt{2}, 0, \ldots \right), \]
\[ e''_2 = \left( 1/\sqrt{2}, -1/\sqrt{2}, 0, \ldots \right), \]
and, in general,
\[ e''_{2i-1} = \left( 0, 0, \ldots, 0, 1/\sqrt{2}, 1/\sqrt{2}, 0, \ldots \right), \]
\[ e''_{2i} = \left( 0, 0, \ldots, 0, 1/\sqrt{2}, -1/\sqrt{2}, 0, \ldots \right), \]
with the non-zero entries in the \((2i-1)\)th and \(2i\)th places respectively.

Show that \( S'' = \{ e''_1, e''_2, \ldots \} \) is a Schauder basis for \( \ell^2 \).

Solution
All we have to do is to show that \( \{ e''_1, e''_2, \ldots \} \) is a total countable orthonormal set in \( \ell^2 \). This set is clearly countable, and it is easy to check that it is also orthonormal in \( \ell^2 \).

Checking back to Section 1, we see that \( \tilde{S} = \{ \tilde{e}_1, \tilde{e}_2, \ldots \} \) is a Schauder basis for \( \ell^2 \).

For we showed there that it was a total orthonormal set, and that \( \sum_{i=1}^{\infty} (x \cdot \tilde{e}_i) \tilde{e}_i \) converged (to \( x \)) for every vector. It is easy to check that:
\[ \tilde{e}_1 = \frac{1}{\sqrt{2}} \left( e''_1 + e''_2 \right), \]
\[ \tilde{e}_2 = \frac{1}{\sqrt{2}} \left( e''_1 - e''_2 \right). \]

Comparing the linear spans of these two sets of vectors, we see that
\[ \langle S'' \rangle \subset \langle \tilde{S} \rangle. \]

Similarly
\[ \tilde{e}_{2i-1} = \frac{1}{\sqrt{2}} \left( e''_{2i-1} + e''_{2i} \right), \]
\[ \tilde{e}_{2i} = \frac{1}{\sqrt{2}} \left( e''_{2i-1} - e''_{2i} \right), \]
so
\[ \langle \tilde{S} \rangle \subset \langle S'' \rangle. \]

Hence
\[ \langle S'' \rangle = \langle \tilde{S} \rangle, \]
and therefore \( S'' \) is a total set in \( \ell^2 \) since \( \tilde{S} \) is total.

We must also show that \( \sum_{i=1}^{\infty} (x \cdot e''_i) e''_i \) converges for each \( x \). Our theorem then tells us that \( S'' \) is a Schauder basis. A calculation gives
\[ \sum_{i=1}^{2N} (x \cdot e''_i) e''_i = \sum_{i=1}^{2N} (x \cdot \tilde{e}_i) \tilde{e}_i \]
for every \( N \) and every \( x \in \ell^2 \). Now the right-hand side defines partial sums, \( s_{2N} \), say, with \( s_{2N} \rightarrow x \) as \( N \rightarrow \infty \). Since \( x \in \ell^2 \), it follows that
\[ \lim_{N \rightarrow \infty} x \cdot e''_{2N-1} = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{2}} (x_{2N-1} + x_{2N}) = 0. \]

Hence
\[ \lim_{N \rightarrow \infty} \sum_{i=1}^{N} (x \cdot e''_i) e''_i = x, \]
and so the convergence condition is satisfied for \( S'' \) and so \( S'' \) is a Schauder basis for \( \ell^2 \).

Exercise 8

Show that the set of vectors \( S' = \{ e'_1, e'_2, \ldots \} \) defined in Exercise 6 of Section 1 is NOT a Schauder basis for \( \ell^2 \). (Hint: Show that there is no series of the form \( \sum_{i=1}^{\infty} a_i e'_i \) which converges in norm to the vector \( a = (1/\sqrt{2}, 0, 0, \ldots) \).)
Solution Consider \[ \| \sum_{i=1}^{N} a_i e'_i - a \|, \] where \( a = ((1/\sqrt{2}), 0, 0, \ldots) \) and \( a_1, \ldots, a_N \) are any real numbers. We have

\[
\sum_{i=1}^{N} a_i e'_i = (1/\sqrt{2}) (a_1, a_1, a_2, \ldots, a_N, 0, 0, \ldots),
\]

so

\[
\sum_{i=1}^{N} a_i e'_i - a = (1/\sqrt{2}) (a_1 - 1, a_1, a_2, \ldots, a_N, 0, 0, \ldots).
\]

Hence

\[
\left\| \sum_{i=1}^{N} a_i e'_i - a \right\|^2 = \frac{1}{2} \left[ (a_1 - 1)^2 + a_1^2 + 2 \left( a_2^2 + \cdots + a_N^2 \right) \right]
\]

\[
= a_1^2 - a_1 + \frac{1}{2} + \left( a_2^2 + \cdots + a_N^2 \right)
\]

\[
= \left( a_1 - \frac{1}{2} \right)^2 + \frac{1}{4} + \left( a_2^2 + \cdots + a_N^2 \right).
\]

Therefore

\[
\left\| \sum_{i=1}^{N} a_i e'_i - a \right\|^2 \geq \frac{1}{4}
\]

for any choice of \( a_1, \ldots, a_N \), so no choice of \( a_1, a_2, \ldots \) can be made such that

\[
\sum_{i=1}^{N} a_i e'_i
\]

converges in norm to \( a \). It follows that \( \{ e'_1, e'_2, \ldots \} \) is not a Schauder basis in \( l^2 \).

Remark It is perhaps worth mentioning that the notion of a Hamel basis applies to any vector space \( V \), whereas the definition of a Schauder basis applies to a normed infinite-dimensional space, so the two concepts are not equivalent. Not every infinite dimensional normed space has a Schauder basis, whereas every vector space has a Hamel basis. In Theorem T.12.2.3(1) we show that every separable complete infinite-dimensional Euclidean space does have a Schauder basis, but it is a deep result of a more advanced subject called Functional Analysis that any Hamel basis for such a space is uncountable. Thus Schauder bases, if they exist, are much easier to work with than Hamel bases (this is part of their attraction). Hamel bases will play no further part in this unit.

2.2 Fourier series

We have seen the importance of the numbers \( x \cdot e_i \) for any vector \( x \) and any orthonormal Schauder basis \( \{ e_i \} \) for a Euclidean space. This prompts us to make the following definition.

Definition T.12.2.2 Let \( E \) be a Euclidean space with an orthonormal Schauder basis \( \{ e_i \} \) and let \( x \in E \). Then the numbers \( x \cdot e_i \) are called the Fourier coefficients of \( x \) and

\[
\sum_{i=1}^{\infty} (x \cdot e_i) e_i
\]

is called the Fourier series of \( x \) with respect to this basis.

Theorem T.12.2.1(2) shows that the Fourier series for \( x \) converges in norm to \( x \).

Having established a framework giving the infinite-dimensional analogue of the first part of Theorem T.12.1.2, we can go on to prove the analogue of the second part. The results in this following theorem are extremely important.
Theorem T.12.2.2 Let $S = \{e_1, e_2, \ldots\}$ be a countable infinite orthonormal set of vectors in a Euclidean space $E$.

(a) Bessel’s Inequality For each $x \in E$, the series $\sum_{i=1}^{\infty} (x \cdot e_i)^2$ is convergent, and

$$\|x\|^2 \geq \sum_{i=1}^{\infty} (x \cdot e_i)^2.$$

(b) Parseval’s Equation The set $S$ is a Schauder basis for $E$ if and only if, for each $x \in E$,

$$\|x\|^2 = \sum_{i=1}^{\infty} (x \cdot e_i)^2.$$

(c) If $S$ is a Schauder basis for $E$, then for any $x, y \in E$,

$$x \cdot y = \sum_{i=1}^{\infty} (x \cdot e_i)(y \cdot e_i).$$

Proof As in Lemma T.12.1.2 (2), we see that

$$0 \leq \left\| x - \sum_{i=1}^{n} (x \cdot e_i)e_i \right\|^2 = \|x\|^2 - \sum_{i=1}^{n} (x \cdot e_i)^2$$

for any $n \geq 1$ and $x \in E$.

(a) If $x \in E$, the sequence of partial sums $\left( \sum_{i=1}^{n} (x \cdot e_i)e_i \right)$ is increasing and bounded above by $\|x\|^2$; the result follows.

(b) The above identity shows that $\|x\|^2 = \sum_{i=1}^{\infty} (x \cdot e_i)^2$ if and only if $s(x) = \sum_{i=1}^{\infty} (x \cdot e_i)e_i$ converges to $x$. We have already shown that $s(x) = x$ for all $x \in E$ if $S$ is a Schauder basis. Conversely, if $s(x) = x$ for all $x \in E$, then $S$ is a countably infinite total orthonormal set in $E$ which satisfies the convergence condition; hence $S$ is a Schauder basis.

(c) If $S$ is a Schauder basis for $E$, we have that

$$x \cdot y = s(x) \cdot y = \left( \sum_{i=1}^{\infty} (x \cdot e_i)e_i \right) \cdot y = \sum_{i=1}^{\infty} (x \cdot e_i)(y \cdot e_i)$$

for any $x, y \in E$.

Exercise 9 Verify that Bessel’s inequality holds and that Parseval’s equation does not hold for the element $x = (1, 0, 1, 0, 0, \ldots)$ in $\ell^2$ relative to the countable infinite orthonormal set of vectors $\{e'_1, e'_2, \ldots\}$ introduced in Exercise 6 of Section 1.

Solution We compute

$$\|x\|^2 = \sum_{i=1}^{\infty} x_i^2 = 1^2 + 1^2 = 2.$$

On the other hand, $(x \cdot e'_1) = 1/\sqrt{2}$, $(x \cdot e'_2) = 1/\sqrt{2}$, and $(x \cdot e'_i) = 0$ for $i > 2$. Therefore,

$$\sum_{i=1}^{\infty} (x \cdot e'_i)^2 = \frac{1}{2} + \frac{1}{2} = 1.$$

Now

$$2 = \|x\|^2 > \sum_{i=1}^{\infty} (x \cdot e'_i)^2 = 1,$$

and so Bessel’s inequality holds and Parseval’s equation does not.
Verify that Parseval's equation and statement (c) of Theorem T.12.2.2 hold for the vectors

\[ x = (1, 0, 1, 0, 0, \ldots), \]
\[ y = (1, \frac{1}{2}, \frac{1}{3}, \ldots, (1/n), \ldots) \]
in \(l^2\) relative to the Schauder basis \(\{e''_1, e''_2, \ldots\}\) introduced in Exercise 7 of Section 2.

Solution First,

\[ (x \cdot e''_1) = 1/\sqrt{2}, \quad (x \cdot e''_2) = 1/\sqrt{2}, \]
\[ (x \cdot e''_3) = 1/\sqrt{2}, \quad (x \cdot e''_4) = 1/\sqrt{2}, \]
and
\[ (x \cdot e''_i) = 0 \quad \text{for } i > 4. \]

Hence

\[ \sum_{i=1}^{\infty} (x \cdot e''_i)^2 = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 2 = ||x||^2, \]

so Parseval's equation holds for \(x\). Next,

\[ (y \cdot e''_{2i-1}) = \left(1/\sqrt{2}\right) \left(\frac{1}{2i-1} + \frac{1}{2i}\right), \]
\[ (y \cdot e''_{2i}) = \left(1/\sqrt{2}\right) \left(\frac{1}{2i-1} - \frac{1}{2i}\right), \]
for all \(i \geq 1\), whilst

\[ ||y||^2 = \sum_{i=1}^{\infty} \frac{1}{i^2}. \]

Now

\[ (y \cdot e''_{2i-1})^2 + (y \cdot e''_{2i})^2 = \frac{1}{2} \left(\frac{2}{(2i-1)^2} + \frac{2}{(2i)^2}\right) = \frac{1}{(2i-1)^2} + \frac{1}{(2i)^2}. \]

Hence

\[ \sum_{i=1}^{\infty} (y \cdot e''_i)^2 = \sum_{i=1}^{\infty} \frac{1}{i^2} = ||y||^2, \]

so Parseval's equation holds for \(y\).

Finally,

\[ x \cdot y = 1 + \frac{1}{3} = \frac{4}{3}, \]
and

\[ \sum_{i=1}^{\infty} (x \cdot e''_i)(y \cdot e''_i) = (1/\sqrt{2}) \left[(1/\sqrt{2}) \left(1 + \frac{1}{2}\right) + (1/\sqrt{2}) \left(1 - \frac{1}{2}\right)\right] \]
\[ + (1/\sqrt{2}) \left(\frac{3}{4} + \frac{1}{4}\right) + (1/\sqrt{2}) \left(\frac{1}{4} - \frac{1}{4}\right) \]
\[ = \frac{1}{2} \left[1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} - \frac{1}{4}\right] = \frac{4}{3}. \]

Hence

\[ x \cdot y = \sum_{i=1}^{\infty} (x \cdot e''_i)(y \cdot e''_i), \]
as required.
2.3 Hilbert space

In Section 2.2 we saw that, in a Euclidean space with an orthonormal Schauder basis, we can define Fourier series and Fourier coefficients and prove an infinite-dimensional analogue of all but part (c) of Theorem T.12.1.2. This part of the theorem tells us that any finite-dimensional Euclidean space with an orthonormal Hamel basis must be isomorphic to \( \mathbb{R}^k \) for some \( k \). We would like an infinite-dimensional analogue of this statement, and we have in the Euclidean space \( \ell^2 \) an obvious infinite-dimensional analogue of \( \mathbb{R}^k \). However, in general, it is not the case that a Euclidean space with an orthonormal Schauder basis is isomorphic to \( \ell^2 \), but it will be so if we impose one further condition on the Euclidean space. To see what this condition is, we compare the general situation with that of \( \ell^2 \).

Suppose that \( S = \{e_1, e_2, \ldots\} \) is an orthonormal Schauder basis of a Euclidean space \( E \). If \( a_1, a_2, \ldots \) are real numbers such that \( \sum_{i=1}^{\infty} a_i e_i \) converges in \( E \) to some \( x \in E \), then Theorem T.12.2.1 tells us that \( a_1 = x \cdot e_i \) for all \( i \); that is, the series is the same as the Fourier series for \( x \), and Bessel's inequality tells us that \( \sum_{i=1}^{\infty} a_i^2 \) converges.

Let us begin our search for a convergence criterion by considering the converse of the above statement. Is the following statement true?

If \( a_1, a_2, \ldots \) are real numbers and \( \sum_{i=1}^{\infty} a_i^2 \) converges, then \( \sum_{i=1}^{\infty} a_i e_i \) converges in \( E \).

So far, we have seen one particular case in which the answer is Yes. Theorem T.12.1.3 (Section 1) tells us that Statement (*) is true in \( \ell^2 \) for a particular choice of basis.

Example 2 Statement (*) is not always true. To see this, let \( V \) be the vector subspace of \( \ell^2 \) consisting of all sequences which are eventually zero, so that each vector has the form \( s = (s_1, s_2, \ldots, s_N, 0, 0, \ldots) \) for some \( N \), and where \( N \) can vary from vector to vector. Let \( S = \{\hat{e}_1, \hat{e}_2, \ldots\} \) be the orthonormal Schauder basis for \( \ell^2 \) defined in Section 1. Clearly \( \hat{e}_i \in V \) for each \( i \) and so \( S \) is also an orthonormal Schauder basis for the Euclidean space \( V \). Note that \( V = (S) \). If we define \( a_i = 1/i \) for each \( i \), then the series \( \sum_{i=1}^{\infty} a_i^2 \) is convergent. From Theorem T.12.1.3, we know that \( \sum_{i=1}^{\infty} a_i \hat{e}_i \) converges in \( \ell^2 \) to the vector \( a = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots) \).

However, this vector does not belong to \( V \), and so the series cannot converge to any vector in \( V \) because, if it did, the series would converge to two different limits in \( \ell^2 \), which is impossible.

We have seen that Statement (*) is false for the Euclidean space \( V \). What is lacking in \( V \) that is not lacking in \( \ell^2 \)? The key property is one introduced in Unit 11, where we identified certain sequences that we felt 'ought to converge', namely Cauchy sequences, and we defined a normed space in which they do converge to be complete. Here we have a similar situation.

We remind you that a sequence \( \{v_n \in V : n \geq 1\} \) of vectors in a normed space \( V \) is called a Cauchy sequence if, given any \( \varepsilon > 0 \), we can find an integer \( N(\varepsilon) \) such that for all \( n, m > N(\varepsilon) \), \( ||v_n - v_m|| < \varepsilon \).

Let \( E \) be a Euclidean space and let \( S = \{e_1, e_2, \ldots\} \) be a countably infinite orthonormal subset of \( E \). If \( a_1, a_2, \ldots \) are real numbers such that \( \sum_{i=1}^{\infty} a_i^2 \) converges, the orthonormality of the set \( S \) and the linear structure of the inner product together show that

\[
\sum_{i=1}^{n} a_i e_i - \sum_{i=1}^{m} a_i e_i = \left( \sum_{i=m+1}^{n} a_i e_i \right) - \left( \sum_{i=m+1}^{n} a_i e_i \right)
\]

for all integers \( n > m \geq 1 \). From the theory of infinite series, given \( \varepsilon > 0 \) we can find some integer \( N \geq 1 \) such that

\[
\left( \sum_{i=m+1}^{n} a_i^2 \right)^{\frac{1}{2}} < \varepsilon
\]
whenever \( n > m \geq N \). Thus the series \( \sum_{i=1}^{\infty} a_i e_i \) is a Cauchy series in \( E \). Thus, to obtain the truth of Statement (\( * \)), we need Cauchy sequences in \( E \) to converge. This leads us to the following definition. \( Q \)

**Definition T.12.2.3** A Hilbert space is a complete Euclidean space.

It is worth noting, before we proceed, that for any positive integer \( k \), \( \mathbb{R}^k \) is a Hilbert space, so that any finite-dimensional Euclidean space is a Hilbert space. Recall that we learnt that \( \mathbb{R} \) is a complete normed space in Unit II.

**Lemma T.12.2.3(1)** \( \mathbb{R}^k \) is a Hilbert space.

**Proof** If \( (x^{(n)}) \) is a Cauchy sequence in \( \mathbb{R}^k \), since

\[
\left[ x_i^{(m)} - x_i^{(n)} \right]^2 \leq \sum_{j=1}^{k} \left[ x_j^{(m)} - x_j^{(n)} \right]^2 = \left\| x^{(m)} - x^{(n)} \right\|^2
\]

for all \( m, n \geq 1 \) and \( 1 \leq i \leq k \), the sequence of real numbers \( (x_i^{(n)}) \) is a Cauchy sequence for all \( 1 \leq i \leq k \), and so converges to some limit \( x_i \). If we consider the vector \( x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k \), then

\[
\left\| x - x^{(n)} \right\|^2 = \sum_{i=1}^{k} \left[ x_i - x_i^{(n)} \right]^2
\]

for all \( n \geq 1 \), and so \( \left\| x - x^{(n)} \right\| \to 0 \). Thus the sequence \( (x^{(n)}) \) converges to \( x \) in \( \mathbb{R}^k \).

Notice that the rational span of a (finite) Hamel basis for a finite-dimensional Euclidean space is countable and dense, so that any finite-dimensional Euclidean space is a separable Hilbert space. We shall prove later on that infinite-dimensional separable Hilbert spaces exist.

It is time to take stock. We now know what a Euclidean space is, what a Schauder basis is and what a Hilbert space is. We shall soon see, in Theorem T.12.2.3(1), that if the Hilbert space is not 'too large', it necessarily has a Schauder basis. Two more steps are needed to obtain that result. In this next lemma we discover precisely which infinite series of orthonormal vectors converge. This result is very important. It is practical, since it may well be possible to check whether or not a sequence belongs to \( \ell^2 \). It is important for theory, as it sets the stage for proving that all separable Hilbert spaces are 'essentially the same' as \( \ell^2 \).

We are coming to the crux of the unit. As we noted in the introduction, this material will take a few readings to assimilate.

**Lemma T.12.2.3(2)** Let \( H \) be an infinite-dimensional Hilbert space, and let \( S = \{e_1, e_2, \ldots\} \) be a countably infinite orthonormal subset. If \( a = (a_1, a_2, \ldots) \) is a sequence of real numbers, then \( \sum_{i=1}^{\infty} a_i e_i \) converges in \( H \) if and only if \( a \in \ell^2 \).

Of course, any infinite-dimensional Hilbert space with a Schauder basis is separable. The crucial point to our definition is that any separable infinite-dimensional Hilbert space has an orthonormal Schauder basis. We present a proof now, with the proviso that it is non-assessed. However, the statement of the proposition is required knowledge.

**Proposition T.12.2.3(1)** Let \( H \) be an infinite-dimensional Hilbert space, and let \( S = \{e_1, e_2, \ldots\} \) be a countably infinite orthonormal subset. If \( S \) is total, then \( S \) is a Schauder basis for \( H \).

**Proof** (Optional) This is a direct corollary of Theorem T.12.2.1 (2), since the fact that \( H \) is a Hilbert space implies the convergence of \( s(x) = \sum_{i=1}^{\infty} (x \cdot e_i) e_i \) in \( H \) for all \( x \in H \).
Theorem T.12.2.3(1) An infinite-dimensional Hilbert space $H$ possesses a Schauder basis if and only if it is separable. Moreover, a separable infinite-dimensional Hilbert space $H$ possesses an orthonormal Schauder basis.

Proof (Optional) Certainly $H$ is separable if it possesses a Schauder basis (Theorem T.12.2.1). If $H$ is separable, let $D = \{d_1, d_2, \ldots\}$ be a countable dense subset of $H$. We can throw away any zero vector in $D$ without affecting the span $(D)$. Starting with $n = 2$, and using induction, throw away $d_n$ from $D$ if $d_n$ is linearly dependent on $d_1, \ldots, d_{n-1}$. Hence we obtain a linearly independent subset $C$ of $D$ such that $(C) = (D)$. Now apply the Gram–Schmidt orthogonalization procedure (see Weir, Theorem 7.5.2, page 191) to obtain an orthonormal set $S$, which is either finite or countably infinite, such that $(S) = (C) = (D)$. Since $D \subset (D) = (S)$ and $D$ is dense in $H$, it follows that $S$ is total.

If $S$ were finite, then $(S)$ would be a finite-dimensional subspace of $H$ and so itself a Hilbert space. If $x \in H$, for any $n \geq 1$ there exists an $x_n \in (S)$ with $||x - x_n|| < n^{-1}$. Then

\[
||x_m - x_n|| \leq ||x_m - x|| + ||x - x_n|| < m^{-1} + n^{-1}, \quad m, n \geq 1,
\]

so $(x_n)$ is a Cauchy sequence in $(S)$, which therefore converges to some $y \in (S)$. But the sequence $(x_n)$ clearly converges to $x$, and so $x = y \in (S)$. As $x$ was chosen arbitrarily, we conclude that $H = (S)$ is finite-dimensional. This contradiction implies that $S$ is a countably infinite total orthonormal subset of $H$, and hence is a Schauder basis for $H$.

Thus any separable Hilbert space possesses an orthonormal Schauder basis. It is worth noting something more. We may also deduce from this that orthonormal subsets of a separable Hilbert space are either finite or countably infinite, and so the theory of orthonormal subsets of separable Hilbert spaces is relatively easy to handle.

Proposition T.12.2.3(2) Any orthonormal subset of a separable Hilbert space $H$ is either finite or countably infinite.

Proof Let $S \subseteq H$ be infinite and orthonormal, and let $D$ be a countable dense subset of $H$. Write $D = \{d_1, d_2, \ldots\}$. For any $s \in S$, let $n(s) \geq 1$ be the least positive integer $n$ such that $||d_n - s|| < 1/\sqrt{2}$. Hence we have defined a function $n : S \rightarrow \mathbb{N}$. If $s, t \in S$ with $s \neq t$, then, using orthonormality,

\[
||s - t||^2 = ||s||^2 - 2s \cdot t + ||t||^2 = 2,
\]

and so

\[
\sqrt{2} = ||s - t|| \leq ||s - d_{n(s)}|| + ||d_{n(s)} - d_{n(t)}|| + ||d_{n(t)} - t|| < (1/\sqrt{2}) + ||d_{n(s)} - d_{n(t)}|| + (1/\sqrt{2})
\]

\[
= \sqrt{2} + ||d_{n(s)} - d_{n(t)}||,
\]

so that $||d_{n(s)} - d_{n(t)}|| > 0$. Then $d_{n(s)} - d_{n(t)} \neq 0$, so $d_{n(s)} \neq d_{n(t)}$, and so $n(s) \neq n(t)$. Hence the function $n : S \rightarrow \mathbb{N}$ is one-one (injective).

We can now write $S = \{s_1, s_2, \ldots\}$, where $s_1$ is the element $s \in S$ with the least value $n(s)$, $s_2$ is the element of $S$ with the second least value, and so on. Hence $S$ is indeed countably infinite.

It is also worth noting that nonseparable Hilbert spaces do exist, and that such Hilbert spaces do possess uncountable orthonormal subsets. For example, let $H$ denote the vector space of functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that

(a) the set $N_f = \{x : f(x) \neq 0\}$ is finite or countably infinite;
(b) the sum $\sum_{x \in \mathbb{R}} f(x)^2$ converges.

To say that a sum over $\mathbb{R}$ converges, that is,

\[
\sum_{x \in \mathbb{R}} a(x) < \infty,
\]
means that \( a(x) = 0 \) except at countably many points \( x_1, x_2, x_3, \ldots \) (at most); and then that
\[
\sum_{n=1}^{\infty} a(x_n)
\]
converges in the usual sense. As our functions have this property (by condition (a)), condition (b) is well-defined.

We define these vector operations as pointwise addition and pointwise scalar multiplication. It is possible to show that \( H \) is a Hilbert space with respect to the inner product
\[
f \cdot g = \sum_{x \in \mathbb{R}} f(x)g(x), \quad f, g \in H,
\]
and has a complete, but uncountable, orthonormal set \( \{f_x : x \in \mathbb{R}\} \), where
\[
f_x(y) = \begin{cases} 1, & y = x, \\ 0, & y \neq x. \end{cases}
\]
However, a detailed study of nonseparable Hilbert spaces does not form part of this unit!

In our continued effort to provide an infinite-dimensional analogue of Theorem T.12.1.2 we note that if \( H \) is a separable Hilbert space with Schauder basis \( S = \{e_1, e_2, \ldots\} \), then Theorem T.12.2.2 and Lemma T.12.2.3 (2) together provide a strong interrelationship between \( H \) and the Euclidean space \( \ell^2 \). We shall now set out this connection explicitly.

**Theorem T.12.2.3(2)** Let \( H \) be any separable infinite-dimensional Hilbert space. Then \( H \) is isomorphic to \( \ell^2 \). That is, there exists a linear transformation \( \varphi : H \rightarrow \ell^2 \) which is one-one and onto, such that
\[
\varphi(x) \cdot \varphi(y) = x \cdot y
\]
for \( x, y \in H \).

**Proof** Let \( S = \{e_1, e_2, \ldots\} \) be a Schauder basis for \( H \). Define \( \varphi : H \rightarrow \ell^2 \) by
\[
\varphi(x) = (x \cdot e_1, x \cdot e_2, \ldots), \quad x \in H,
\]
and define \( \psi : \ell^2 \rightarrow H \) by setting
\[
\psi(a) = \sum_{i=1}^{\infty} a_i e_i, \quad a \in \ell^2.
\]
It is clear that \( \varphi : H \rightarrow \ell^2 \) and \( \psi : \ell^2 \rightarrow H \) are linear, and that
\[
\varphi(x) \cdot \varphi(y) = \sum_{i=1}^{\infty} (x \cdot e_i)(y \cdot e_i) = x \cdot y
\]
for all \( x, y \in H \). Moreover \( \psi(a) \cdot e_i = a_i \) for all \( a \in \ell^2 \) and \( i \geq 1 \), so that \( \varphi(\psi(a)) = a \) for all \( a \in \ell^2 \). It is also clear that \( \psi(\varphi(x)) = s(x) = x \) for all \( x \in H \). Thus \( \varphi : H \rightarrow \ell^2 \) is one-one and onto, with inverse \( \psi : \ell^2 \rightarrow H \). Hence it also follows that
\[
\psi(a) \cdot \psi(b) = a \cdot b, \quad a, b \in \ell^2.
\]
Hence we can show that any two separable Hilbert spaces \( H_1, H_2 \) are isomorphic, for if we consider the maps \( \varphi_1 : H_1 \rightarrow \ell^2 \) and \( \psi_2 : \ell^2 \rightarrow H_2 \), we obtain a linear map \( \rho = \psi_2 \cdot \varphi_1 \) from \( H_1 \) to \( H_2 \) which is one-one and onto, and which satisfies
\[
\rho(x) \cdot \rho(y) = x \cdot y, \quad x, y \in H_1.
\]

Theorem T.12.2.3(2) tells us that there is ‘essentially’ only one separable infinite-dimensional Hilbert space. The jargon for ‘essentially’ is ‘up to isomorphism’. The isomorphisms in question are \( \varphi, \psi, \rho \) given above.
We can now consider Euclidean spaces from this point of view. You will see that we do not have to check completeness if we can find an isomorphism onto a Hilbert space.

**Theorem T.12.2.3(3)** Let $E$ be a Euclidean space, let $H$ be a Hilbert space, and let $\sigma : E \rightarrow H$ be a linear transformation which is one-one and onto, and which preserves inner products. Then $E$ is also a Hilbert space. In addition, if $H$ is separable, then so is $E$.

**Proof** Since $\sigma$ preserves inner products, it also preserves norms. Thus, if $(x_n)$ is a Cauchy sequence in $E$, then $(\sigma(x_n))$ is a Cauchy sequence in $H$ which converges to some limit $y \in H$. Let $x$ be the unique element of $H$ such that $y = \sigma(x)$. Then $\|x_n - x\| = \|\sigma(x_n) - y\| = \|\sigma(x_n) - y\|$ converges to 0 as $n$ tends to infinity, and so the sequence $(x_n)$ converges to $x$ in $E$. Hence $E$ is a Hilbert space.

If $H$ is separable, let $D$ be a countable dense subset of $H$. Let $C = \sigma^{-1}(D)$. Since $\sigma$ provides a bijection between $C$ and $D$, it is clear that $C$ is a countable subset of $E$. If $x \in E$ and $\varepsilon > 0$, we can find $d \in D$ such that $\|\sigma(x) - d\| < \varepsilon$. Then $c = \sigma^{-1}(d) \in C$ and $\|x - c\| = \|\sigma(x) - c\| = \|\sigma(x) - d\| < \varepsilon$. Hence $C$ is dense in $E$, so that $E$ is separable.

Thus, if $H$ is any separable Hilbert space, the map $\psi : \ell^2 \rightarrow H$, defined in the proof of Theorem T.12.2.3(2), can now be used to prove that $\ell^2$ is also a separable Hilbert space. We have already seen that $\ell^2$ possesses a Schauder basis. We could prove the existence of separable infinite-dimensional Hilbert spaces by proving directly that $\ell^2$ is a Hilbert space, which has a Schauder basis, and so is separable (SAQ 5, in Section ?). However, in view of the work of previous units, we have chosen not to do things this way, but rather to prove that $L^2(\mathbb{R})$, the space of square-integrable functions on $\mathbb{R}$, is a separable Hilbert space. We shall do this in Section 3.

The fact that any two separable infinite-dimensional Hilbert spaces are isomorphic does not imply that all separable Hilbert spaces are in all senses indistinguishable. They are indistinguishable, if you are only interested in properties of the linear structure and the inner product. However, $\ell^2$, $L^2[-\pi, \pi]$ and $L^2(\mathbb{R})$, which are all separable infinite-dimensional Hilbert spaces, are clearly different spaces for other purposes!

**Exercise 11**

Assuming the existence of a separable Hilbert space, show that the subspace $W$ of $\ell^2$ consisting of those elements $a \in \ell^2$ for which $a_{2i} = 0$ for all $i \geq 1$ is a separable Hilbert space.

**Solution** We have just noted that if one separable Hilbert space exists, then $\ell^2$ is another. We may therefore assume that $\ell^2$ is a Hilbert space. It is clear that the map $\sigma : W \rightarrow \ell^2$ defined by

$$\sigma(a) = (a_1, a_3, a_5, \ldots) \quad a \in W,$$

is a linear map such that

$$\sigma(a) \cdot \sigma(b) = a \cdot b \quad a, b \in W,$$

so if $a \in W$ with $\sigma(a) = 0$, then $\|a\|^2 = a \cdot a = \sigma(a) \cdot \sigma(a) = 0$, and hence $a = 0$.

Thus $\sigma : W \rightarrow \ell^2$ is one-one.

If $b \in \ell^2$, consider $c \in W$ given by

$$c_{2i} = 0, \quad c_{2i-1} = b_i, \quad i \geq 1.$$

It is clear that $\sigma(c) = b$; hence $\sigma : W \rightarrow \ell^2$ is also onto. Thus $W$ is a separable Hilbert space by Theorem T.12.2.3(3).

If $\hat{S} = \{\hat{e}_1, \hat{e}_2, \ldots\}$ is the usual Schauder basis for $\ell^2$, it is clear that

$$\{\hat{e}_1, \hat{e}_3, \hat{e}_5, \ldots\} = \sigma^{-1}((\hat{S}))$$

is a Schauder basis for $W$.  


3 The Hilbert space $L^2(\mathbb{R})$

This leaves us to tie up the last loose end. For it remains to be shown that at least one separable infinite-dimensional Hilbert space exists — otherwise the theory is vacuous. This brings us full circle, and we return to $L^2(\mathbb{R})$, the space of all square-integrable functions. Our aim is to show that $L^2(\mathbb{R})$ is a separable Hilbert space. Hence it will follow that $l^2$ is also a separable Hilbert space, and that $L^2(\mathbb{R})$ and $l^2$ are isomorphic.

We have already seen that $L^2(\mathbb{R})$ is a Euclidean space with respect to the inner product defined by

$$ f \cdot g = \int fg, \quad f, g \in L^2(\mathbb{R}), $$

and, as we saw in Unit 11, $L^2(\mathbb{R})$ is complete with respect to the norm induced by this inner product. Hence $L^2(\mathbb{R})$ is a Hilbert space. It remains to show that $L^2(\mathbb{R})$ is separable. To do this we show that $L^2(\mathbb{R})$ possesses a countable total subset.

**Theorem T.12.3.1** $L^2(\mathbb{R})$ is separable.

**Proof** Let $C = \{ \chi_{[a,b]} : a, b \in \mathbb{Q}, a \leq b \}$ be the set of characteristic functions of closed subintervals of $\mathbb{R}$ with rational endpoints. Clearly, the set of ordered pairs $\{(a, b) : a, b \in \mathbb{Q}, a \leq b\}$ is a subset of the countable set $\mathbb{Q} \times \mathbb{Q}$, and so is itself countable. Hence $C$ is a countable subset of $L^2(\mathbb{R})$. It remains to show that $C$ is total.

The details of the proof that $C$ is total are rather tedious, and involve many steps, each of which is not difficult. As long as you understand the statements of Stages 1 and 2 and the whole of Stage 3 of the proof, you may omit the details.

Let $f \in L^2(\mathbb{R})$, and choose $\varepsilon > 0$.

(a) **Stage 1** To find a step function $\varphi$ such that $\|f - \varphi\|_2 < \frac{1}{2}\varepsilon$.

For any $n \geq 1$, we define $f_n : \mathbb{R} \rightarrow \mathbb{R}$ by setting

$$ f_n(x) = \begin{cases} f(x), & \text{if } |x| \leq n \text{ and } |f(x)| \leq n, \\ 0, & \text{otherwise}. \end{cases} $$

Note that

$$ f_n = \max[\min(f, n), -n] \cdot \chi_{[-n,n]}, $$

and so $f_n$ is measurable (see Weir, Theorem 6.1.1, page 121) and so $f_n^2$ is also measurable. Note that

$$ |f_n| \leq n \cdot \chi_{[-n,n]}, \quad 0 \leq f_n^2 \leq n^2 \cdot \chi_{[-n,n]}, $$

so that $f_n, f_n^2 \in L^1(\mathbb{R})$ (see Weir, Proposition 6.1.1, page 122), and hence $f_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

Now $f_n \rightarrow f$ pointwise, and $f_n - f \in L^2(\mathbb{R})$, so that $((f_n - f)^2)$ is a sequence of functions in $L^1(\mathbb{R})$ which converge pointwise to 0 such that

$$ 0 \leq (f_n - f)^2 \leq f^2, $$

where $f^2 \in L^1(\mathbb{R})$. Hence the Dominated Convergence Theorem (see Weir, Theorem 5.2.1, page 109) implies that $\int (f_n - f)^2$ converges to 0 as $n$ tends to infinity, so that $\|f_n - f\|_2$ tends to 0. Hence we can find $N \geq 1$ such that $\|f - f_N\|_2 < \frac{1}{4}\varepsilon$.

By Weir, Exercise 5, page 42 (solution on page 243), we can find a sequence $(\varphi_m)$ of step functions such that $\int |f_N - \varphi_m|$ tends to 0 as $m$ tends to infinity. Since $|f_N| \leq N$ we may assume that $|\varphi_m| \leq N$ for all $m$. Choose a step function $\varphi = \varphi_m$ such that

$$ \int |f_N - \varphi| < \varepsilon^2/32N. $$

Since

$$ |f_N - \varphi| \leq 2N $$
we see that
\[ \|f_N - \varphi\|^2 \leq \int |f_N - \varphi| |f_N - \varphi| \leq 2N \int |f_N - \varphi| < \varepsilon^2/16, \]
so that
\[ \|f_N - \varphi\|_2 < \varepsilon/4, \]
and hence
\[ \|f - \varphi\|_2 \leq \|f - f_N\|_2 + \|f_N - \varphi\|_2 < \frac{1}{2} \varepsilon. \]

(b) Stage 2 To find \( \psi \in (C) \) such that \( \|\varphi - \psi\|_2 < \frac{1}{2} \varepsilon. \)

As \( \varphi \) is a step function, we can write it as a finite linear combination of characteristic functions of intervals with real endpoints, say
\[ \varphi = \sum_{j=1}^{L} c_j \chi_{[a(j),b(j)]}. \]

Let
\[ K = 1 + 2 \sum_{j=1}^{L} |c_j|. \]

For \( 1 \leq j \leq L \), we can find rational numbers \( A(j), B(j) \in \mathbb{Q} \) such that
\[ a(j) - \varepsilon/2K^2 < A(j) \leq a(j) < B(j) < b(j) < b(j) + \varepsilon^2/2K^2. \]

Now consider the element \( \psi \in (C) \) given by
\[ \psi = \sum_{j=1}^{L} c_j \chi_{[A(j),B(j)]} \in (C). \]

Then \( \psi \) is the required element of \((C). \) For
\[ \|\varphi - \psi\|_2 = \left\| \sum_{j=1}^{L} c_j (\chi_{[a(j),b(j)]} - \chi_{[A(j),B(j)]}) \right\|_2 \]
\[ \leq \sum_{j=1}^{L} |c_j| \left\| \chi_{[a(j),b(j)]} - \chi_{[A(j),B(j)]} \right\|_2 \]
\[ = \sum_{j=1}^{L} |c_j|[A(j) - a(j)] + [B(j) - b(j)]^{1/2} < \frac{\varepsilon}{K} \sum_{j=1}^{L} |c_j| < \frac{1}{2} \varepsilon, \]
as required.

(c) Stage 3 Hence
\[ \|f - \psi\|_2 \leq \|f - \varphi\|_2 + \|\varphi - \psi\|_2 < \frac{1}{2} \varepsilon + \frac{1}{2} \varepsilon = \varepsilon. \]

Thus \((C)\) is dense in \( L^2(\mathbb{R}) \) and so \( C \) is total. Hence \( L^2(\mathbb{R}) \) is separable.
And so we have come to the end of this long and difficult unit. What have we accomplished? Rather a lot, in fact. Checking with the glossary below, you should now know what a Hilbert space is; the structure of all separable Hilbert spaces; that there are such spaces, $\ell^2$ and $L^2(\mathbb{R})$ being examples. You know what a Schauder basis is and that every separable Hilbert space has one. You can even supply an example of one for the space $\ell^2$. As we stated in the introduction, we have discussed existence and uniqueness and completely settled these questions for separable Hilbert spaces.

What next? There is a theory of linear transformations on Hilbert space, very deep and very important. The entire theory of Hilbert spaces has an extremely important application: the mathematical structure of quantum theory. Apparently Mother Nature also knows about Hilbert spaces, and she uses them quite extensively in her work at the atomic level.

We have taken the liberty of suggesting some advanced books which take the theory of Hilbert spaces further, some books on functional analysis, and some on applications to quantum theory. We are not suggesting that you immediately purchase any of these books. Rather, if you wish to carry your work further at some future time, these are good books to look at.

We began by studying Euclidean spaces, that is, normed spaces in which the norm is derived from an inner product, and we derived a criterion, the Parallelogram Law, which enables us to decide which norms are derived from an inner product. Using inner products, we also introduced the ideas of linearly independent, orthogonal, and orthonormal sets of vectors. We studied examples of such sets in the space of sequences $\ell^2$, which is an infinite-dimensional analogue of $\mathbb{R}^d$.

The next stage in our search for the right setting in which to discuss Fourier series led us to the idea of a Schauder basis in a normed space; in a Euclidean space, this is equivalent to the concept of a countable total orthonormal set of vectors. We then
defined *Fourier coefficients* and the *Fourier series* for an element $x$ in a Euclidean space with an orthonormal Schauder basis, and we saw that the Fourier series of $x$ must always converge strongly to $x$; we proved that the sum of squares of the Fourier coefficients also converges. In attempting to demonstrate the converse, we found that we had to impose the condition of completeness, which led us to the definition of a Hilbert space as a Euclidean space which is complete.

We found also that any separable Hilbert space must be isomorphic to $\ell^2$, and, finally, we showed that $L^2$, the space of square integrable functions, is a separable Hilbert space.

*In the following subsections references to the set book or this text are preceded by $W$ or $T$ respectively. For example, a reference to page 9 of the set book is indicated by $[W9]$ and to Section 3 of this text by $[T3]$."

### 4.1 Notation

- $PC[-\pi,\pi]$ [T1]
- $\ell^2$ [T1]
- $L^2$ [T1]
- $x \cdot y$ [T1]
- $L^1$, $L^3$ [T1]
- $\{e_1, e_2, \ldots\}$ [T1]
- $\ell^\infty$ [T1]
- $L^2(-\pi, \pi)$ [T2]
- $H$ [T2]

### 4.2 Glossary

- Bessel's inequality [T2]
- Cauchy-Schwarz inequality [T1]
- complete space [T2]
- dense set [T2]
- Euclidean space [T1]
- Fourier coefficients [T2]
- Fourier series [T2]
- Hamel basis [T2]
- Hilbert space [T2]
- inner product [T1]
- linearly independent [T1]
- non-separable Hilbert space [T2]
- orthogonal [T1]
- orthonormal [T1]
- Parallelogram Law [T1]
- Parseval's equation [T2]
- piecewise continuous [T1]
- Schauder basis [T2]
- Schwarz's inequality [T1]
- separable Hilbert space [T2]
- separable space [T2]
- total set [T2]

### 4.3 Results

**Lemma T.12.1.1** [T1] If $V$ is a Euclidean space, then $V$ is a normed space with respect to the norm defined by

$$||v|| = \sqrt{v \cdot v} \quad (v \in V).$$

**Theorem T.12.1.1** [T1] If $V$ is a vector space with a norm derived from an inner product, then the Parallelogram Law holds; that is, for any $v$, $w \in V$,

$$||v + w||^2 + ||v - w||^2 = 2(||v||^2 + ||w||^2).$$

**Lemma T.12.1.2(1)** [T1] An orthogonal set of vectors in a Euclidean space $V$ is linearly independent.

**Theorem T.12.1.2** [T1] Let $V$ be a real finite-dimensional Euclidean space with an orthonormal Hamel basis $\{e_1, e_2, \ldots, e_k\}$.

(a) Each $v \in V$ can be expressed in the form

$$v = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_k e_k,$$

where the unique coefficients are given by

$$\lambda_i = v \cdot e_i, \quad i = 1, \ldots, k.$$
(b) For each \(v, w \in V\),
\[
v \cdot w = \sum_{i=1}^{k} (v \cdot e_i)(w \cdot e_i);
\]
in particular,
\[
\|v\|^2 = \sum_{i=1}^{k} (v \cdot e_i)^2.
\]
(c) If \(\mathbb{R}^k\) is considered as a Euclidean space with respect to the standard inner product
\[
(a_1, a_2, \ldots, a_k) \cdot (b_1, b_2, \ldots, b_k) = \sum_{i=1}^{k} a_i b_i,
\]
then there is an isomorphism \(\theta : V \rightarrow \mathbb{R}^k\) given by
\[
\theta(v) = ((v \cdot e_1), (v \cdot e_2), \ldots, (v \cdot e_k)) \quad (v \in V),
\]
which preserves inner products, so that
\[
v \cdot w = \theta(v) \cdot \theta(w).
\]

**Lemma T.12.1.2** [T1] Let \(\{e_1, e_2, \ldots, e_n\}\) be an orthonormal set of vectors in a Euclidean space \(V\), and let \(v \in V\). Then
\[
\left\| v - \sum_{i=1}^{n} \lambda_i e_i \right\| \geq \left\| v - \sum_{i=1}^{n} (v \cdot e_i)e_i \right\|
\]
whenever \(N \leq n\), and \(\lambda_1, \ldots, \lambda_N\) are any real numbers.

**Lemma T.12.1.3** [T1] If the series \(\sum_{i=1}^{\infty} a_i^2, \sum_{i=1}^{\infty} b_i^2\) are convergent, then the series \(\sum_{i=1}^{\infty} a_i b_i\) is absolutely convergent.

**Proposition T.12.1.3** [T1] \(\ell^2\) is a Euclidean space, with inner product given by
\[
a \cdot b = \sum_{i=1}^{\infty} a_i b_i
\]
for \(a, b \in \ell^2\).

**Corollary T.12.1.3** [T1] If \(a, b \in \ell^2\) then
\[
\left\| \sum_{i=1}^{\infty} a_i b_i \right\| \leq \left( \sum_{i=1}^{\infty} a_i^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{\infty} b_i^2 \right)^{\frac{1}{2}}.
\]

**Theorem T.12.1.3** [T1] If the vectors \(\hat{e}_i \in \ell^2\) are defined as above, and \(a\) is any vector in \(\ell^2\), then the series
\[
\sum_{i=1}^{\infty} (a \cdot \hat{e}_i) \hat{e}_i
\]
converges to \(a\), in the norm derived from the inner product on \(\ell^2\).

**Lemma T.12.2.1(1)** [T2] If \(S = \{s_1, s_2, \ldots\}\) is a Schauder basis in the normed space \(N\), then \(S\) is linearly independent.

**Theorem T.12.2.1(1)** [T2] Let \(N\) be a normed space containing a countably infinite total subset \(C\) (for example a Schauder basis). Then \(N\) is separable.
Lemma T.12.2.1(2) [T2] Let \((v_n)\) be a sequence of vectors in a Euclidean space \(E\) converging to a vector \(v\),
\[
\lim_{n \to \infty} \|v_n - v\| = 0.
\]
If \(y\) is any vector in \(E\), then
\[
\lim_{n \to \infty} |v_n \cdot y - v \cdot y| = 0.
\]
We usually abbreviate this as:
\[
v_n \rightarrow v \text{ implies } v_n \cdot y \rightarrow v \cdot y.
\]

Corollary T.12.2.1 [T2] Let \((x_i)\) be a countably infinite family of vectors in a Euclidean space \(E\) such that the series \(\sum_{i=1}^{\infty} x_i\) converges. Then if \(y \in E\) is any vector,
\[
y \cdot \sum_{i=1}^{\infty} x_i = \sum_{i=1}^{\infty} y \cdot x_i.
\]
In particular, the series on the right-hand side converges.

Theorem T.12.2.1(2) [T2] Let \(S = \{e_1, e_2, e_3, \ldots\}\) be a countably infinite and total orthonormal set of vectors in the Euclidean space \(E\). Suppose that for every vector \(x \in E\) the infinite series
\[
s(x) = \sum_{i=1}^{\infty} (x \cdot e_i)e_i
\]
converges. Then \(S\) is a Schauder basis and \(s(x) = x\); that is,
\[
x = \sum_{i=1}^{\infty} (x \cdot e_i)e_i.
\]

Theorem T.12.2.2 [T2] Let \(S = \{e_1, e_2, \ldots\}\) be a countable infinite orthonormal set of vectors in a Euclidean space \(E\).

(a) Bessel's Inequality For each \(x \in E\), the series \(\sum_{i=1}^{\infty} (x \cdot e_i)^2\) is convergent, and
\[
||x||^2 \geq \sum_{i=1}^{\infty} (x \cdot e_i)^2.
\]

(b) Parseval's Equation The set \(S\) is a Schauder basis for \(E\) if and only if, for each \(x \in E\),
\[
||x||^2 = \sum_{i=1}^{\infty} (x \cdot e_i)^2.
\]

(c) If \(S\) is a Schauder basis for \(E\), then for any \(x, y \in E\),
\[
x \cdot y = \sum_{i=1}^{\infty} (x \cdot e_i)(y \cdot e_i).
\]

Lemma T.12.2.3(1) [T2] \(\mathbb{R}^k\) is a Hilbert space.

Lemma T.12.2.3(2) [T2] Let \(H\) be an infinite-dimensional Hilbert space, and let \(S = \{e_1, e_2, \ldots\}\) be a countably infinite orthonormal subset. If \(a = (a_1, a_2, \ldots)\) is a sequence of real numbers, then \(\sum_{i=1}^{\infty} a_i e_i\) converges in \(H\) if and only if \(a \in \ell^2\).

Proposition T.12.2.3(1) [T2] Let \(H\) be an infinite-dimensional Hilbert space, and let \(S = \{e_1, e_2, \ldots\}\) be a countably infinite orthonormal subset. If \(S\) is total, then \(S\) is a Schauder basis for \(H\).

Theorem T.12.2.3(1) [T2] An infinite-dimensional Hilbert space \(H\) possesses a Schauder basis if and only if it is separable. Moreover, a separable infinite-dimensional Hilbert space \(H\) possesses an orthonormal Schauder basis.
**Proposition T.12.2.3(2)** [T2] Any orthonormal subset of a separable Hilbert space \( H \) is either finite or countably infinite.

**Theorem T.12.2.3(2)** [T2] Let \( H \) be any separable infinite-dimensional Hilbert space. Then \( H \) is isomorphic to \( \ell^2 \) in that there exists a linear transformation \( \varphi : H \rightarrow \ell^2 \) which is one-one and onto, such that

\[
\varphi(x) \cdot \varphi(y) = x \cdot y
\]

for \( x, y \in H \).

**Theorem T.12.2.3(3)** [T2] Let \( E \) be a Euclidean space, let \( H \) be a Hilbert space, and let \( \sigma : E \rightarrow H \) be a linear transformation which is one-one and onto, and which preserves inner products. Then \( E \) is also a Hilbert space. If \( H \) is separable, then so is \( E \).

**Theorem T.12.3** [T3] \( L^2(\mathbb{R}) \) is separable.

**Theorem 7.5.2** (Gram–Schmidt) [W191] Let \( f_1, \ldots, f_m \in L^2 \). Then there exist orthogonal functions \( h_1, \ldots, h_m \) in \( L^2 \) such that for \( r = 1, \ldots, m \), the sets \( \{f_1, \ldots, f_r\} \) and \( \{h_1, \ldots, h_r\} \) span the same linear subspace of \( L^2 \).

### 4.4 Further reading

Most texts on Functional Analysis treat the topic of Hilbert space in more detail than we have been able to do in this unit. We recommend the following books.


The following more advanced books deal with the theory and application of Hilbert space theory to quantum mechanics.


5 Self-Assessment Questions

5.1 Euclidean spaces

SAQ 1
Let $p$ be a positive integer and $L^p$ be the space of measurable functions $f$ for which $|f|^p$ is integrable; then we know that $L^p$ is a normed space with a norm defined by

$$
\|f\|_p = \left( \int |f|^p \right)^{1/p}.
$$

Find those values of $p$ for which this norm derives from an inner product.

SAQ 2
Let $p$ be a positive integer, and let $\ell^p$ be the space of sequences of real numbers,

$$
a = (a_1, a_2, \ldots),
$$

for which the series $\sum_{i=1}^{\infty} |a_i|^p$ converges.

Given that $\ell^p$ is a normed space with norm defined by

$$
\|a\|_p = \left( \sum_{i=1}^{\infty} |a_i|^p \right)^{1/p},
$$

find those values of $p$ for which this norm derives from an inner product.

SAQ 3
Show that there is a countable total set in each of the normed spaces $\ell^p$, introduced in SAQ 2.

5.2 Hilbert space

SAQ 4
Prove Lemma T.12.1.1.

SAQ 5
Let $E$ be a Euclidean space and $H$ be a subspace of $E$ which is a separable Hilbert space with $E \neq H$. Show that every $x \in E \setminus H$ can be written uniquely in the form $x = h + g$, where $h \in H$ and $g$ is orthogonal to each element in $H$. (*Hint: Suppose $\{e_1, e_2, \ldots\}$ is an orthonormal Schauder basis of $H$; for $x \in E$, write $h = \sum_{i=1}^{\infty} (x \cdot e_i)e_i$. You will need to use Corollary T.12.2.1.)*

SAQ 6
Prove that $\ell^2$ is complete. As $\ell^2$ has a Schauder basis, this affords a direct proof that it is a separable Hilbert space.
Solutions to Self-Assessment Questions

Solution 1
We have to find those values of \( p \) for which the Parallelogram Law holds. As in the example below Theorem T.12.1.1, let \( f = \chi_{[0,1]} \) and \( g = \chi_{[2,3]} \). Then \( f, g \in L^p \) for all \( p \), and
\[
\begin{align*}
&f = |f| = |f|^p, \\
g = |g| = |g|^p, \\
f + g = |f + g| = |f + g|^p, \\
|f - g| = |f + g|.
\end{align*}
\]
Now \( \int f = \int g = 1 \), and \( \int (f + g) = 2 \), so we have
\[
\begin{align*}
\|f\|_p &= \|g\|_p = 1, \\
\|f + g\|_p &= \|f - g\|_p = 2^{1/p}.
\end{align*}
\]
If the Parallelogram Law holds, we must have
\[
\|f + g\|^p_2 + \|f - g\|^p_2 = 2 \left( \|f\|^p_2 + \|g\|^p_2 \right),
\]
that is,
\[
2(4)^{1/p} = 4 \quad \text{or} \quad 4 = 2^p.
\]
Therefore, if \( p \neq 2 \), \( L^p \) is not a Euclidean space.

Finally, we have seen that \( L^2 \) is a Euclidean space, so \( p = 2 \) is the only value for which the norm in \( L \) is derived from an inner product.

Solution 2
A similar argument to that in SAQ 1 can be used to show that the only value for which the given norm derives from an inner product is \( p = 2 \). If we take \( a = (1,0,0,\ldots) \) and \( b = (0,1,0,\ldots) \),
then
\[
\|a\| = \|b\| = 1 \quad \text{and} \quad \|a - b\| = \|a + b\| = 2^{1/p}.
\]
A computation similar to that in SAQ 1 shows that the only possibility is \( p = 2 \).

Finally, we have already seen that \( L^2 \) is a Euclidean space, so \( p = 2 \) is the only value for which the norm in \( L \) is derived from an inner product.

Solution 3
The obvious set to try is the set that worked in \( \ell^2 \). For each positive integer \( i \), let \( \hat{e}_i = (0,\ldots,0,1,0,\ldots) \), where the non-zero entry is in the \( i \)-th place; then \( \hat{e}_i \in \ell^p \) for all \( i \) and all \( p \), and clearly \( \{\hat{e}_1, \hat{e}_2, \ldots\} \) is a countable set.

To show that \( \{\hat{e}_1, \hat{e}_2, \ldots\} \) is total in \( \ell^p \), we have to show that, given any \( a \in \ell^p \), say \( a = (a_1, a_2, \ldots) \), and any \( \varepsilon > 0 \), there is a finite linear combination \( f \) of the \( \hat{e}_i \) such that
\[
\|a - f\| < \varepsilon.
\]
Now, if \( a \in \ell^p \), then \( \sum_{i=1}^{\infty} |a_i|^p \) converges, so there exists an integer \( N \) such that
\[
\sum_{i=N+1}^{\infty} |a_i|^p < \varepsilon^p.
\]
Now put
\[
f = \sum_{i=1}^{N} a_i \hat{e}_i = (a_1, a_2, \ldots, a_N, 0, 0, \ldots);
\]
then
\[ \|a - f\| = \|(0, 0, \ldots, 0, a_{N+1}, a_{N+2}, \ldots)\| = \left(\sum_{i=N+1}^{\infty} |a_i|^p\right)^{1/p} < [\varepsilon^p]^{1/p} = \varepsilon. \]

Therefore \( \{\hat{e}_1, \hat{e}_2, \ldots\} \) is total in \( \mathcal{G} \), as required.

**Solution 4**

Most of the proof is fairly direct. If \( v = 0 \) is the zero vector, then \( v \cdot v = 0 \), so
\( (||v||)^2 = 0 \); hence \( ||0|| = 0 \). On the other hand, if \( ||v|| = 0 \), its square vanishes:
\( v \cdot v = 0 \). But from condition (iv) for inner products, \( v = 0 \). The scaling law is just as easy. If \( v \) is any vector and \( a \in \mathbb{R} \), then
\( (av) \cdot (av) = a^2(v \cdot v). \)

Taking (positive) square roots gives
\[ ||av|| = |a| ||v||. \]

It is the subadditivity of the norm which poses the only difficulty. The solution is an exercise in using inequalities, and should be studied carefully as a good model for such problems. Now because \( v \cdot v \geq 0 \) for all vectors \( v \), we have
\[ 0 \leq (v + w) \cdot (v + w) = ||v||^2 + ||w||^2 + 2(v \cdot w) \]
\[ = ||v||^2 + ||w||^2 + 2||v|| ||v|| \leq (||v|| + ||w||)^2. \]

This is the first crucial step. Even though \( v \cdot w \) can be negative, it appears in a combination known to be positive, so we can put an overall absolute value around everything. Then we apply the triangle inequality to obtain an expression involving \( ||v \cdot w|| \). This now enables us to apply the Schwarz inequality. We find, upon rewriting,
\[ 0 \leq (v + w) \cdot (v + w) \leq (||v|| + ||w||)^2. \]

That is,
\[ ||v + w||^2 \leq (||v|| + ||w||)^2; \]

taking square roots now completes the proof.

**Solution 5**

Since \( H \) is a separable Hilbert space, it has an orthonormal Schauder basis, \( \{e_1, e_2, \ldots\} \). Suppose \( x \in E \setminus H \). By Theorem T.12.2.2(a), Bessel’s inequality, the series
\[ \sum_{i=1}^{\infty} (x \cdot e_i)^2 \]
converges. Hence, because \( H \) is a Hilbert space, the series
\[ \sum_{i=1}^{\infty} (x \cdot e_i)e_i \]
converges in \( H \), by Theorem T.12.2.3(2). Let this sum be \( h \), and put \( g = x - h \). We have
\[ x = g + h, \]
and wish to show that \( g \) is orthogonal to every vector in \( H \). Clearly \( g \neq 0 \).

If \( h' \in H \), we have
\[ h' \cdot g = h' \cdot (x - h) = h' \cdot x - h' \cdot h. \]
Now, by Theorem T.12.2.2(c),
\[
h' \cdot h = \sum_{i=1}^{\infty} (h' \cdot e_i)(h \cdot e_i) = \sum_{i=1}^{\infty} (h' \cdot e_i)(x \cdot e_i),
\]
as
\[
h = \sum_{i=1}^{\infty} (x \cdot e_i)e_i.
\]
Now as $h' \in H$ and $\{e_1, e_2, \ldots\}$ is a Schauder basis of $H$, there is a series $\sum_{i=1}^{\infty} a_i e_i$ that converges to $h'$. Hence we may apply the result of Corollary T.12.2.1 to obtain
\[
h' \cdot x = \sum_{i=1}^{\infty} (h' \cdot e_i)(x \cdot e_i).
\]
Thus $h' \cdot g = 0$, as required.

Our last task is to prove uniqueness. So suppose
\[
x = g + h,
\]
\[
x = g' + h'
\]
are two such expressions for $x$. Then
\[
g - g' = h' - h.
\]
But
\[
g \cdot (h' - h) = 0 = g' \cdot (h' - h),
\]
and so
\[
(g - g') \cdot (h' - h) = 0.
\]
Thus
\[
\|h' - h\|^2 = 0,
\]
so
\[
h' - h = 0,
\]
We conclude, then, that $h = h'$, and so $g = g'$. This proves that our expression for $x$ is unique, and our solution is complete.

**Solution 6**

Let $(x^{(k)})$ be a Cauchy sequence in $\ell^2$, and write $x^{(k)} = (x_n^{(k)})$, $k \in \mathbb{N}$. Then given any $\varepsilon > 0$ we can find an integer $N(\varepsilon)$ such that for all $k, p > N(\varepsilon)$ we have
\[
\left\|x^{(k)} - x^{(p)}\right\|^2 = \sum_{n=1}^{\infty} \left| x_n^{(k)} - x_n^{(p)} \right|^2 < \varepsilon^2.
\]
Then for each $n$
\[
\left| x_n^{(k)} - x_n^{(p)} \right| < \varepsilon, \quad \text{for} \ k, p > N(\varepsilon).
\]
Consequently the sequence $(x_n^{(k)})$ converges as $k \to \infty$, with $n$ fixed. That is, there is a number $x_n$ such that
\[
\lim_{k \to \infty} x_n^{(k)} = x_n.
\]
From the estimate for $\left\|x^{(k)} - x^{(p)}\right\|^2$ above, it follows that for arbitrary $M$ we have
\[
\sum_{n=1}^{M} \left| x_n^{(k)} - x_n^{(p)} \right|^2 < \varepsilon^2, \quad \text{for} \ k, p > N(\varepsilon).
We note that
\[
\lim_{k \to \infty} \left| x_n^{(k)} - x_n^{(p)} \right|^2 = \left| x_n - x_n^{(p)} \right|^2,
\]
and so taking the limit \( k \to \infty \) in the previous sum gives
\[
\sum_{n=1}^{N(e)} \left| x_n - x_n^{(p)} \right|^2 \leq \epsilon^2, \quad \text{for } p > N(e).
\]
As \( N(e) \) does not depend on \( M \) we can now let \( M \to \infty \), which yields
\[
\sum_{n=1}^{\infty} \left| x_n - x_n^{(p)} \right|^2 \leq \epsilon^2, \quad \text{for } p > N(e).
\]
This we interpret as meaning that for any \( p > N(e) \), the sequence
\[
(x_n - x^{(p)}) = (x_1 - x_1^{(p)}, x_2 - x_2^{(p)}, \ldots) \in l^2.
\]
As \( (x_n^{(p)}) \in l^2 \) it follows that
\[
(x_n) \in l^2, \quad \text{and more: } (x_n) \text{ is the limit of the } x^{(p)} \text{ as } p \to \infty.
\]
Hence any Cauchy sequence in \( l^2 \) has a limit in \( l^2 \), so \( l^2 \) is complete.

6 Appendix

1. **Definition** A sequence \((x_n)\) of real numbers converges to a real number \(x\) if for any \(\epsilon > 0\) there exists an integer \(N \geq 1\) such that \(|x_n - x| < \epsilon\) whenever \(n \geq N\).

2. **Definition** A sequence \((x_n)\) of real numbers is a Cauchy sequence if for any \(\epsilon > 0\) there exists an integer \(N \geq 1\) such that \(|x_m - x_n| < \epsilon\) whenever \(m, n \geq N\).

3. **Theorem** A sequence \((x_n)\) is convergent if and only if it is a Cauchy sequence.

4. **Theorem** A monotonic increasing bounded sequence \((x_n)\) is convergent, and its limit is \(\sup x_n\). In particular, if \(x_1 \leq x_2 \leq \ldots \leq x_n \leq \ldots \leq K\), then \(\lim_{n \to \infty} x_n \leq K\).

5. **Theorem** If \((x_n), (y_n)\) are sequences of real numbers converging to \(x, y\) respectively, and if \(\lambda\) is a real number, then \((\lambda x_n + y_n)\) is a sequence converging to \(\lambda x + y\).

6. **Definition** If \(a_1, a_2, \ldots\) are real numbers, the series \(\sum_{i=1}^{\infty} a_i\) is said to converge if the sequence \((s_n)\) of partial sums
\[
s_n = \sum_{i=1}^{n} a_i
\]
converges to a limit \(s\). We write \(s = \sum_{i=1}^{\infty} a_i\). Thus, if \(\sum_{i=1}^{\infty} a_i\) is convergent, then
\[
\sum_{i=n+1}^{\infty} a_i = s - s_n
\]
converges to 0 as \(n \to \infty\).

7. **Theorem** If \(a_1, a_2, \ldots\) are real numbers, the series \(\sum_{i=1}^{\infty} a_i\) is convergent if and only if for any \(\epsilon > 0\) there exists an integer \(N \geq 1\) such that
\[
\left| \sum_{i=m+1}^{n} a_i \right| < \epsilon,
\]
whenever \(n > m \geq N\).
8. **Theorem** If \(a_i \geq 0\) for all \(i\), and if there exists \(K > 0\) such that \(\sum_{i=1}^{n} a_i \leq K\) for all \(n \geq 1\), then the series \(\sum_{i=1}^{\infty} a_i\) is convergent, and its limit \(s\) satisfies the inequality \(s \leq K\).

9. **Theorem** If \(\sum_{i=1}^{\infty} a_i\), \(\sum_{i=1}^{\infty} b_i\) are convergent series, and \(\lambda\) is real, then \(\sum_{i=1}^{\infty} (\lambda a_i + b_i)\) is convergent, and \(\sum_{i=1}^{\infty} (\lambda a_i + b_i) = \lambda (\sum_{i=1}^{\infty} a_i) + \sum_{i=1}^{\infty} b_i\).

10. **Definition** A series \(\sum_{i=1}^{\infty} a_i\) is absolutely convergent if the series \(\sum_{i=1}^{\infty} |a_i|\) is convergent.

11. **Theorem** If a series \(\sum_{i=1}^{\infty} a_i\) is absolutely convergent, then it is also convergent, and
\[
\left| \sum_{i=1}^{\infty} a_i \right| \leq \sum_{i=1}^{\infty} |a_i|.
\]

12. **Theorem** If \(\sum_{i=1}^{\infty} a_i\) is convergent, then \(\lim_{i \to \infty} a_i = 0\).

13. **The Schröder-Bernstein Theorem** Recall that a set \(T\) is countably infinite if there exists a one-one and onto map from \(T\) to \(\mathbb{N}\), the set of natural numbers, and that \(T\) is countable if it is either finite or countably infinite. A one-one map is also known as an injective map or an injection; an onto map is also known as a surjective map or a surjection; a map which is injective and surjective is also known as a bijection or a bijection.

The following two conditions on a set \(T\) are equivalent, and a set \(T\) is countable if and only if either (and so both) of these conditions holds.

(a) There exists an injection from \(T\) to \(\mathbb{N}\).

(b) There exists a surjection from \(\mathbb{N}\) to \(T\).

This result is the version for \(\mathbb{N}\) of a much more general result.

14. **Corollary** By the (Cartesian) product \(S \times T\) of two sets we mean the set of ordered pairs:
\[
S \times T = \{(s, t) : s \in S, t \in T\}.
\]

Writing
\[
S \times T = \bigcup_{s \in S} \{(s, t) : t \in T\},
\]
it follows that the product of two countable sets is countable. By induction, the product \(S_1 \times \ldots \times S_n\) for \(n\) countable sets is countable, \(n \in \mathbb{N}\).

15. If \(f : X \to Y\) is a bijection between the sets \(X, Y\), and \(g : Y \to Z\) is a surjective map between the sets \(Y, Z\), then \(g \circ f : X \to Z\) is a surjective map between \(X\) and \(Z\).

16. For any \(n \in \mathbb{N}\), the \(n\)-fold product, \(\mathbb{Q}^n\), of the rationals with itself is countable; write \(f : \mathbb{N} \to \mathbb{Q}^n\) for some bijection between the indicated sets. Let \(A\) be a set such that there exists a surjection \(g : \mathbb{Q}^n \to A\). Then \(g \circ f : \mathbb{N} \to A\) is surjective, and so \(A\) is countable.