February 11, 2015

AN INTRODUCTION TO HILBERT SPACES

BASED ON NOTES OF RODICA D. COSTIN

Contents

1. Going from finite to infinite dimension 3
1.1. Recall some basic facts about vector spaces 3
1.2. Inner product 5
1.3. Norm 6
1.4. The space ℓ^2 7
1.5. Metric Spaces 8
2. Completeness 9
2.1. Complete spaces 9
2.2. Example 1 10
2.3. Further examples 12
2.4. Completion and closure 13
2.5. The Hilbert space $L^2[a,b]$ 13
2.6. The Banach space $C[a,b]$ 14
2.7. The Banach spaces L^p 15
2.8. Closed sets, dense sets 15
2.9. Sets dense in the Hilbert space L^2 16
2.10. Polynomials are dense in the Banach space $C[a,b]$ 16
3. Hilbert Spaces 16
3.1. When does a norm come from an inner product? 17
3.2. The inner product is continuous 17
3.3. Orthonormal bases 17
3.4. Generalized Fourier series in L^2 20
3.5. Example of orthonormal basis in $L^2[a,b]$: Fourier series 20
3.6. Other bases for L^2: orthogonal polynomials 22
3.7. Orthogonal complements, The Projection Theorem 22
3.8. Least squares approximation via subspaces 23
4. Linear operators in Hilbert spaces 24
4.1. Shift operators on ℓ^2 25
4.2. Unitary operators. Isomorphic Hilbert spaces 25
4.3. Integral operators 26
4.4. Differential operators in $L^2[a,b]$ 28
4.5. A second order boundary value problem 29

1. Going from finite to infinite dimension

1.1. **Recall some basic facts about vector spaces.** Vector spaces are modeled after the familiar vectors in the line, plane, space etc. abstractly written as $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3, \ldots, \mathbb{R}^n, \ldots$; these are vector spaces over the scalars $\mathbb{R}$. It turned out that there is a great advantage to allow for complex coordinates, and then we may also look at $\mathbb{C}$, $\mathbb{C}^2$, $\mathbb{C}^3, \ldots, \mathbb{C}^n, \ldots$; these are vector spaces over the scalars $\mathbb{C}$.

In general:

Definition 1. *A* vector space *over the scalar field F is a set V endowed with two operations, one between vectors: if $x, y \in V$ then $x + y \in V$, and one between scalars and vectors: if $c \in F$ and $x \in V$ then $cx \in V$ having the following properties:*
- commutativity of addition: $x + y = y + x$
- associativity of addition: $x + (y + z) = (x + y) + z$
- existence of zero: there is an element $0 \in V$ such that $x + 0 = x$ for all $x \in V$
- existence of the opposite: for any $x \in V$ there is an opposite, denoted by $-x$, such that $x + (-x) = 0$
- distributivity of scalar multiplication with respect to vector addition: $c(x + y) = cx + cy$
- distributivity of scalar multiplication with respect to field addition: $(c + d)x = cx + dx$
- compatibility of scalar multiplication with field multiplication: $c(dx) = (cd)x$
- identity element of scalar multiplication: $1x = x$.

Some familiar definitions are reformulated below in a way that allows us to tackle infinite dimensions too.

Definition 2. *A set of vectors $S \subset V$ is called* linealy independent *if whenever, for some $x_1, x_2, \ldots, x_n \in S$ (for some n) there are scalars $c_1, c_2, \ldots, c_n \in F$ such that $c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0$ then this necessarily implies that all the scalars are zero: $c_1 = c_2 = \ldots = c_n = 0$.*

Note that the zero vector can never belong to in a linearly independent set.

Definition 3. *Given any set of vectors $S \subset V$* the span of *S is the set*

$$Sp(S) = \{c_1x_1 + c_2x_2 + \ldots + c_nx_n \,|\, x_j \in S,\, c_j \in F,\, n = 1, 2, 3 \ldots\}$$

Note that $Sp(S)$ forms a vector space, included in V; it is a *subspace* of V.

Definition 4. *A set of vectors $S \subset V$* is a basis of *V if it is linearly independent and its span equals V.*

Theorem 5. *Any vector space has a basis. Moreover, all the basis of V have the same cardinality, which is called the dimension of V.*

By the cardinality of a set we usually mean "the number of elements". However, if we allow for infinite dimensions, the notion of "cardinality" helps us distinguish between different types of "infinities". For example, the positive integers form an infinite set, and so do the real numbers; we somehow feel that we should say there are "fewer" positive integers than reals. We call the cardinality of the positive integers *countable*.

Please note that the integers $\mathbb{Z}$ are also countable (we can "count" them: $0, 1, -1, 2, -2, 3, -3, \ldots$), pairs of positive integers $\mathbb{Z}_+^2$ are also countable (count: $(1,1),(2,1),(1,2),(3,1),(2,2),(1,3),\ldots$), and so are $\mathbb{Z}^2$, and the set of rational numbers $\mathbb{Q}$ (rational numbers are ratios of integers m/n). It can be proved that the reals are not countable.

Notations The positive integers are denoted $\mathbb{Z}_+ = \{1, 2, 3, \ldots\}$ and the natural numbers are denoted $\mathbb{N} = \{0, 1, 2, 3 \ldots\}$. However: some authors do not include 0 in the natural numbers, so when you use a book make sure you know the convention used there.

Examples

1. $\mathbb{R}^n = \{(x_1, \ldots x_n) \,|\, x_j \in \mathbb{R}\}$ is a vector space over $\mathbb{R}$ of dimension n, with a basis consisting of the vectors
$e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0 \ldots, 0), \ldots, e_n = (0, 0, \ldots 0, 1)$
Remark: from now on we will prefer to list horizontally the components of vectors.

Any $x = (x_1, \ldots x_n)$ can be written as a linear combination of them

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

and there is an inner product:

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$$

and any vector has a norm

$$\|x\|^2 = x_1^1 + \ldots + x_n^2$$

2. $\mathbb{R}^{\mathbb{Z}_+} = \{x = (x_1, x_2, x_3, \ldots) \,|\, x_j \in \mathbb{R}\}$ is a vector space over $\mathbb{R}$. By analogy with $\mathbb{R}^n$ we can formulate the following *wish list*.

We would like to say that a norm is defined by

$$\|x\|^2 = x_1^1 + \ldots + x_n^2 + \ldots$$

that an inner product of two sequences is

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n + \ldots$$

and that a basis consists of the vectors

$$e_1 = (1, 0, 0, 0, \ldots),\ e_2 = (0, 1, 0, 0, 0 \ldots),\ e_3 = (0, 0, 1, 0, \ldots), \ldots \tag{1}$$

since any $x = (x_1, x_2, x_3, \ldots)$ can be written as an (infinite) linear combination

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3 + \ldots = \sum_{n=1}^{\infty} x_n \, e_n \tag{2}$$

But all these are not quite correct, because we have series rather than finite sums. It looks like we should accept series as expansions, and to restrict the sequences we work with in order to get a nice extension to infinite dimensions!

3. Similarly, $\mathbb{C}^n = \{(z_1, \ldots z_n) \,|\, z_j \in \mathbb{C}\}$ is a vector space over $\mathbb{C}$ of dimension n, with a basis consisting of the vectors
$(1, 0, \ldots, 0)$, $(0, 1, 0 \ldots, 0), \ldots, (0, 0, \ldots 0, 1)$.

4. $\mathbb{C}^{\mathbb{Z}_+} = \{z = (z_1, z_2, \ldots z_n, \ldots) \,|\, z_j \in \mathbb{C}\}$ is a vector space over $\mathbb{C}$ and any z could be thought as the infinite sum $z_1 e_1 + z_2 e_2 + z_3 e_3 + \ldots$ where e_n are given by (1) - but again, this is not a basis in the sense of the definition for vector spaces).

5. The set of polynomials of degree at most n, with coefficients in F (which is $\mathbb{R}$ or $\mathbb{C}$)

$$\mathcal{P}_n = \{p(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \,|\, a_j \in F\}$$

is a vector space over the field of scalars F, of dimension $n + 1$, and a basis is $1,\ t\,, t^2, \ldots\ t^n$.

6. The set of all polynomials with coefficients in F (which is $\mathbb{R}$ or $\mathbb{C}$)

$$\mathcal{P} = \{p(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \,|\, a_j \in F, n \in \mathbb{N}\}$$

is a vector space over the field of scalars F, and has the countable basis $1,\ t,\ t^2,\ t^3 \ldots$.

7. The set of all functions continuous on a closed interval (it could also be open, or extending to ∞):

$$C[a, b] = \{f : [a, b] \to F \,|\, f \text{continuous}\}$$

8. The set of all functions f with absolute value $|f|$ integrable on $[a, b]$:

$$L^1[a, b] = \{f : [a, b] \to F \,|\, |f| \text{integrable}\}$$

Warning: I did not specify what "integrable" means. $L^1[a, b]$ is a bit more complicated, but for practical purposes this is good enough for now. (We will comment more later.)

Note that $\mathcal{P} \subset C[a, b] \subset L^1[a, b]$.

1.2. **Inner product.** Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$ have an interesting operation:

Definition 6. *An* **inner product** *on vector space V over F (=$\mathbb{R}$ or $\mathbb{C}$) is an operation which associate to two vectors $x, y \in V$ a scalar $\langle x, y\rangle \in F$ that satisfies the following properties:*
- it is conjugate symmetric: $\langle x, y\rangle = \overline{\langle y, x\rangle}$

- *it is linear in the second argument:* $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$ *and* $\langle x, cy\rangle = c\langle x, y\rangle$
- *its is positive definite:* $\langle x, x\rangle \geq 0$ *with equality only for* $x = 0$*.*

Note that conjugate symmetry combined with linearity implies that $\langle .,.\rangle$ is conjugate linear in the first variable.

Note that for $F = \mathbb{R}$ an inner product is symmetric and linear in the first argument too.

Please keep in mind that most mathematical books use inner product linear in the first variable, and conjugate linear in the second one. You should make sure you know the convention used by each author.

Definition 7. *A vector space* V *equipped with an inner product* $(V, \langle .,.\rangle)$ *is called an* **inner product space***.*

Examples

On $\mathbb{C}^n$ the most used inner product is $\langle x, y\rangle = \sum_{j=1}^{n} \overline{x_j} y_j$.

We may wish to introduce a similar inner product on $\mathbb{C}^{\mathbb{Z}_+}$: $\langle x, y\rangle = \sum_{j=1}^{\infty} \overline{x_j} y_j$. The problem is that the series may not converge!

On the space of polynomials $\mathcal{P}$ or on continuous functions $C[a, b]$ we can introduce the inner product

$$\langle f, g\rangle = \int_a^b \overline{f(x)}\, g(x)\, dx$$

or more generally, using a weight (which is a positive function $w(x)$),

$$\langle f, g\rangle_w = \int_a^b \overline{f(x)}\, g(x)\, w(x) dx$$

We may wish to introduce a similar inner product on $L^1[a, b]$, only the integral may not converge. For example, $f(x) = 1/\sqrt{x}$ is integrable on $[0, 1]$, but $f(x)^2$ is not.

1.3. **Norm.** The inner product defines a length by $\|x\| = \sqrt{\langle x, x\rangle}$. This is a norm, in the following sense:

Definition 8. *Given a vector space* V*, a* **norm** *is a function on* V *such that:*
- *it is positive definite:* $\|x\| \geq 0$ *and* $\|x\| = 0$ *only for* $x = 0$
- *it is positive homogeneous:* $\|cx\| = |c|\, \|x\|$ *for all* $c \in F$ *and* $x \in V$
- *satisfies the* **the triangle inequality** *(i.e. it is subadditive):*

$$\|x + y\| \leq \|x\| + \|y\|$$

Definition 9. *A vector space* V *equipped with a norm* $(V, \|.\|)$ *is called a* **normed space***.*

An inner product space is, in particular a normed space (the first two properties of the norm are immediate, the triangle inequality is a geometric property in finite dimensions and requires a proof in infinite dimensions).

There are some very useful normed spaces (of functions), which are not inner product spaces (more about that later).

1.4. **The space ℓ^2.** Let us consider again $\mathbb{C}^{\mathbb{Z}_+}$, which appears to be the most straightforward way to go from finite dimension to an infinite one. To do linear algebra we need an inner product, or at least a norm. We must then restrict to the "vectors" which do have a norm, in the sense that the series $\|z\|^2 = \sum_{n=1}^{\infty} |z_n|^2$ converges.

Define the vector space ℓ^2 by

$$\ell^2 = \{z = (z_1, z_2, z_3, \ldots) \in \mathbb{C}^{\mathbb{Z}_+} \mid \sum_{n=1}^{\infty} |z_n|^2 < \infty\} \tag{3}$$

On ℓ^2 we therefore have a norm: $\|z\| = \left(\sum_{n=1}^{\infty} |z_n|^2\right)^{1/2}$.

Examples.

1. The constant sequence with $z_n = c$ for all n is not in ℓ^2 unless $c = 0$.

2. The sequences with $z_n = \frac{1}{n^a}$ with $a \in \mathbb{R}$ is in ℓ^2 only for $a > 1/2$ (why?).

3. For complex powers of n, recall that they are defined as

$$n^{a+ib} = e^{(a+ib)\ln n} = e^{a\ln n} e^{ib\ln n} = n^a\, e^{ib\ln n}$$

therefore if $z_n = \frac{1}{n^{a+ib}}$ the sequence is in ℓ^2 only for $a > 1/2$.

Actually, on ℓ^2 the inner product converges as well, due to the following inequality, which is one of the most important and powerful tools in infinite dimensions (the triangle inequality is also fundamental):

Theorem 10. The Cauchy-Schwartz inequality

In an inner product space we have

$$\left| \langle x, y\rangle \right| \leq \|x\| \, \|y\|$$

Therefore, if $\|x\|$ and $\|y\|$ converge, then $\langle x, y\rangle$ converges and moreover, equality holds if and only if x, y are linearly dependent (which means $x = 0$ or $y = 0$ or $x = cy$)

Proof.

Intuitively: if $\|x\|$ and $\|y\|$ converge, then $Sp(x, y)$ is an inner product space which is two-dimensional at most, therefore the Cauchy-Schwartz inequality follows from the one in finite dimensions.

Here are detailed rigorous arguments for the case of ℓ^2, with a review of the main concepts on convergent series.

Recall: a series $\sum_{n=1}^{\infty} a_n$ is said to converge to A, and we write $\sum_{n=1}^{\infty} a_n = A$ if the sequence formed by its partial sums $S_N = \sum_{n=1}^{N} a_n$ converges to A as $N \to \infty$.

Recall: a series $\sum_{n=1}^{\infty} a_n$ is said to converge absolutely if $\sum_{n=1}^{\infty} |a_n|$ converges.

Recall: absolute convergence implies convergence.

Recall: a series with positive terms either converges or has the limit $+\infty$. For such a series, say $\sum_{n=1}^{\infty} |a_n|$, it is customary to write $\sum_{n=1}^{\infty} |a_n| < \infty$ to express that it converges.

Take the partial sums of $\langle x, y\rangle$, use the triangle inequality in dimension N, then Cauchy-Schwartz in dimension N (which follows from $u \cdot v = \|u\| \, \|v\| \cos\theta$):

(4)

$$\Big|\sum_{n=1}^{N} \overline{x_n}\, y_n\Big| \le \sum_{n=1}^{N} \left|\overline{x_n}\, y_n\right| = \sum_{n=1}^{N} |x_n|\,|y_n| \le \left(\sum_{n=1}^{N} |x_n|^2\right)^{1/2} \left(\sum_{n=1}^{N} |y_n|^2\right)^{1/2}$$

Taking the limit $N \to \infty$ the convergence of the $\langle x, y\rangle$ series follows if $\|x\|$ and $\|y\|$ converge.

The argument showing when equality holds is not given here, as it is in accordance with what happens in finite dimensions. $\square$

We obtained that ℓ^2 is an inner product space, by the Cauchy-Schwartz inequality.

The vectors e_n of (1) do belong to ℓ^2, and they do form an *orthonormal set*: $e_n \perp e_k$ for $n \neq k$ (since $\langle e_n, e_k\rangle = 0$) and $\|e_n\| = 1$).

1.5. **Metric Spaces.** Now we would like to make sense of the expansion (2). For this, we need to state what we mean by convergence in ℓ^2.

We can do that using the usual definition of convergence (in $\mathbb{R}$) by replacing the distance between two vectors x, y by $d(x, y) = \|x - y\|$. This distance is a metric, in the following sense:

Definition 11. *A distance (or a metric) on a set M is a function $d(x, y)$ for $x, y \in M$ with the following properties:*
it is nonnegative: $d(x, y) \ge 0$
it separates the points: $d(x, y) = 0$ if and only if $x = y$
it is symmetric: $d(x, y) = d(y, x)$
it satisfies the triangle inequality: $d(x, z) \le d(x, y) + d(y, z)$.

(Two of conditions follow from the other, but it is better to have them all listed.)

Definition 12. *A metric space (M, d) is a set M equipped with a distance d.*

Note that any normed space is a metric space by defining the distance d:

$$d(x, y) = \|x - y\| \tag{5}$$

But there are many interesting metric spaces which are not normed (they may not even be vector spaces!). For example, we can define distances on a sphere (or on the surface of the Earth!) by measuring the (shortest) distance between two points on a large circle joining them.

Once we have a distance, we can define convergence of sequences:

Definition 13. *Consider a metric space (M, d).*
We say that a sequence $s_1, s_2, s_3 \ldots \in M$ converges to $L \in M$ if for any $\varepsilon > 0$ there is an N (N depends on ε) such that $d(s_n, L) < \varepsilon$ for all $n \geq N$.

Note that we can use this definition for convergence in normed spaces using the distance (5).

The series expansion (2) of any $x \in \ell^2$ converges (why?). We now have a satisfactory theory of ℓ^2.

2. Completeness

There is one more property that is essential for calculus: there need to exist enough limits.

The spaces $\mathbb{R}^n$, $\mathbb{C}^n$, ℓ^2 all have this property (inherited from $\mathbb{R}$). However, the space of polynomials $\mathcal{P}$ and $C[a, b]$ as inner product spaces with $\langle f, g\rangle = \int \overline{f} g$ do not have this property.

2.1. Complete spaces.

In the case of $\mathbb{R}$ this special property can be intuitively formulated as: if a sequence of real numbers a_n tends to "pile up" then it is convergent. Here is the rigorous formulation:

Definition 14. *The sequence $\{a_n\}_{n\in\mathbb{N}}$ is called a* Cauchy sequence *if for any $\varepsilon > 0$ there is an $N \in \mathbb{N}$ (N depending on ε) such that*

$$|a_n - a_m| < \varepsilon \quad \textit{for all } n, m > N$$

Theorem 15. *Any Cauchy sequence of real numbers is convergent.*

(This is a fundamental property of real numbers, at the essence of what they are.)

We can define Cauchy sequences in a metric space very similarly:

Definition 16. *Let (M, d) be a metric space.*

The sequence $\{a_n\}_{n\in\mathbb{N}} \subset M$ is called a Cauchy sequence *if for any $\varepsilon > 0$ there is an $N \in \mathbb{N}$ (N depending on ε) such that*

$$d(a_n, a_m) < \varepsilon \quad \textit{for all } n, m > N$$

Theorem 17. If a sequence in a metric space is convergent, then it is Cauchy

Proof. Let (x_n) be the sequence and l the limit. By the definition of convergence, for any ε there is an n_0 s.t. for all $n \geq n_0$ we have $d(x_n, l) < \varepsilon/2$. Let $n_1, n_2 > n_0$. Then $d(x_{n_1}, x_{n_2}) \leq d(x_{n_1}, l) + d(l, x_{n_2}) < \varepsilon$. □

But the converse is not always true: it is not always the case that in a metric space Cauchy sequences automatically converge. Fortunately if this is not the case, we can enlarge them to enforce this convergence condition. More about that later. For now we just define the spaces we will most often work with: complete ones.

Definition 18. *A metric space* (M, d) *is called* complete *if every Cauchy sequence is convergent in* M*.*

(Note that the limit of the Cauchy sequence *must belong to* M.)

In the particular case of normed spaces (when the distance is given by the norm):

Definition 19. *A normed space* $(V, \| \,.\, \|)$ *which is* complete *is called a Banach*[1] *space.*

In the even more special case of inner product spaces (when the norm is given by an inner product):

Definition 20. *An inner product space* $(V, \langle \,.\, \rangle)$ *which is* complete *is called a Hilbert*[2] *space.*

2.2. **Example 1.**

Theorem 21. *The space* $\ell^2(\mathbb{N})$ *in Definition 19 with the scalar product* $\langle x, y \rangle$ *is a Hilbert space.*

Proof. We have already shown that $\langle x, y \rangle$ is well defined, see (4) and the paragraph following it.

Assume $x^{[n]}$ is a Cauchy sequence of elements of ℓ^2. First we show that the sequence is bounded, that is there is an M such that

$$\|x^{[n]}\| \leq M \quad \forall n \in \mathbb{N} \tag{6}$$

Indeed, by the definition of a Cauchy sequence, for any $\varepsilon > 0$ there is an n_0 s.t.

$$\|x^{[n_2]} - x^{[n_1]}\| < \varepsilon \;\forall n_1, n_2 > n_0 \tag{7}$$

[1]Stefan Banach (1892-1945) was a Polish mathematician, founder of modern functional analysis - a domain of mathematics these lectures belong to.

[2]David Hilbert (1862-1943) was a German mathematician, recognized as one of the most influential and universal mathematicians of the 19th and early 20th centuries. He discovered and developed a broad range of fundamental ideas in many areas, including the theory of Hilbert spaces, one of the foundations of functional analysis.

In particular,

$$\|x^{[n_2]} - x^{[n_0]}\| < \varepsilon \ \forall n > n_0 \tag{8}$$

Let $M_1 = \|x^{[n_0]}\|$ and $M = M_1 + 2\varepsilon$. Then,

$$\|x^{[n]}\| = \|x^{[n]} - x^{[n_0]} + x^{[n_0]}\| \le \|x^{[n]} - x^{[n_0]}\| + \|x^{[n_0]}\| \le \varepsilon + M_1 < M \tag{9}$$

- Next we show that for each j

$$x_j^{[n]} \to L_j \quad \text{for some } L \in \mathbb{C} \tag{10}$$

Indeed,

$$|x_j^{[n_2]} - x_j^{[n_1]}|^2 \le \sum_{j=1}^{\infty} |x_j^{[n_2]} - x_j^{[n_1]}|^2 \tag{11}$$

and thus $\{x_j^{[n]}\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{C}$, and it must converge to some (of course finite) L_j.

- Next we show that $(L_1, L_2, ...) = L \in \ell^2$. Indeed, for any n and N,

$$\sum_{j=1}^{N} |L_j|^2 \le \sum_{j=1}^{N} |L_j - x_j^{[n]}|^2 + \sum_{j=1}^{N} |x_j^{[n]}|^2 \le \sum_{j=1}^{N} |L_j - x_j^{[n]}|^2 + \sum_{j=1}^{\infty} |x_j^{[n]}|^2 \le \sum_{j=1}^{N} |L_j - x_j^{[n]}|^2 + M \tag{12}$$

and since $|L_j - x_j^{[n]}| \to 0$ for any fixed N, we have,

$$\sum_{j=1}^{N} |L_j|^2 \le M \tag{13}$$

- Finally, we show that

$$\|x^{[n]} - L\| \to 0 \quad \text{as } n \to \infty \tag{14}$$

Indeed, for any given ε, there is an n_0 s.t., for all $n > n_0$ we have $\|x^{[n]} - x^{[n_0]}\| < \varepsilon/2$. Now, for the same ε there is an N s.t.

$$\sum_{N+1}^{\infty} \|x_j^{[n_0]}\|^2 < \varepsilon/2 \tag{15}$$

(why is that the case?). We then have, for $n > n_0$

$$\sum_{N+1}^{\infty} \|x_j^{[n]}\|^2 = \sum_{N+1}^{\infty} \|x_j^{[n]} - x_j^{[n_0]} + x_j^{[n]}\|^2 \le \sum_{1}^{\infty} \|x_j^{[n]} - x_j^{[n_0]}\| + \sum_{N+1}^{\infty} \|x_j^{[n_0]}\|^2 < \varepsilon \tag{16}$$

Let $N_2 \geq N$ be s.t.

$$\sum_{N_1+1}^{\infty} \|L_j\|^2 < \varepsilon/2 \tag{17}$$

Now,

Then, for any $n > n_0$ we have

(18)

$$\sum_{j=1}^{\infty} |L_j - x^{[n]}|^2 \leq \sum_{j=1}^{N} |L_j - x^{[n]}|^2 + \sum_{N_1+1}^{\infty} |L_j|^2 + \sum_{N_1+1}^{\infty} |x_j^{[n]}|^2 \leq \sum_{j=1}^{N} |L_j - x^{[n]}|^2 + \varepsilon \leq \varepsilon$$

since the first sum goes to zero. □

2.3. **Further examples.** These are very important examples (some proofs are needed but not given here).

1. The usual $\mathbb{R}^n$ and $\mathbb{C}^n$ are Hilbert spaces.

2. The space ℓ^2 of sequences:

$$\ell^2 \equiv \ell^2(\mathbb{Z}_+) = \{x = (x_1, x_2, x_3, \ldots) \,|\, x_n \in \mathbb{C},\ \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$$

endowed with the $\boldsymbol{\ell^2}$ **inner product**

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$$

is a Hilbert space.

For example, the sequence $x = (x_1, x_2, x_3, \ldots)$ with $x_n = 1/n$ belongs to ℓ^2 and so do sequences with $x_n = ca^n$ if $|a| < 1$.

A variation of this space is that of bilateral sequences:

$$\ell^2(\mathbb{Z}) = \{x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2 \ldots) \,|\, x_n \in \mathbb{C},\ \sum_{n=-\infty}^{\infty} |x_n|^2 < \infty\}$$

with *the* ℓ^2 *inner product* $\langle x, y\rangle = \sum_{n=-\infty}^{\infty} \overline{x_n} y_n$ is a Hilbert space.

Recall that a bilateral series $\sum_{n=-\infty}^{\infty} a_n$ is called convergent if both series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=-\infty}^{0} a_n$ are convergent.

3. The space $C[a, b]$ of continuous function on the interval $[a, b]$, endowed with the $\mathbf{L^2}$ **inner product**

$$\langle f, g\rangle = \int_a^b \overline{f(t)}\, g(t)\, dt$$

is *not* a Hilbert space, since *it is not complete.*

For example, take the following approximations of step functions

$$f_n(t) = \begin{cases} 0 & \text{if } t \in [0,1] \\ n\,(t-1) & \text{if } 1 < t < 1 + \frac{1}{n} \\ 1 & \text{if } t \in [1+\frac{1}{n}, 2] \end{cases}$$

The f_n are continuous on $[0, 2]$ (the middle line in the definition of f_n represents a segment joining the two edges of the step). The sequence $\{f_n\}_n$ is Cauchy in the L^2 norm: (say $n < m$)

$$\|f_n - f_m\|^2 = \int_0^2 |f_n(t) - f_m(t)|^2\,dt = \int_1^{1+\frac{1}{n}} |f_n(t) - f_m(t)|^2\,dt$$

$$= \int_1^{1+\frac{1}{m}} [n\,(t-1) - m\,(t-1)]^2\,dt + \int_{1+\frac{1}{m}}^{1+\frac{1}{n}} [n\,(t-1) - 1]^2\,dt =$$

$$= 1/3\,\frac{(n-m)^2}{m^3} - 1/3\,\frac{(n-m)^3}{m^3 n} = 1/3\,\frac{(n-m)^2}{m^2 n} < \frac{1}{3n} \to 0$$

However, f_n is not L^2 convergent in $C[0, 2]$. Indeed, in fact f_n converges to the step function

$$f(t) = \begin{cases} 0 & \text{if } t \in [0,1) \\ 1 & \text{if } t \in [1,2] \end{cases}$$

because

$$\|f_n - f\|^2 = \int_0^2 |f_n(t) - f(t)|^2\,dt = \int_1^{1+\frac{1}{n}} [n\,(t-1) - 1]^2\,dt = \frac{1}{3n} \to 0$$

therefore the L^2 limit of f_n is f, a function that does not belong to $C[0, 2]$.

2.4. **Completion and closure.** The last example suggests that if a metric space of interest is not closed, then one can add to that space all the possible limits of the Cauchy sequences, and then we obtain a closed space.

This procedure is called "closure", and the closure of a metric space M is denoted by $\overline{M}$.

2.5. **The Hilbert space $L^2[a, b]$.** The closure of $C[a, b]$ *in the L^2 norm* (the closure depends on the metric!) is a space denoted by $L^2[a, b]$, *the space of square integrable*[3] *functions*:

$$L^2[a,b] = \{f : [a,b] \to \mathbb{C}(\text{or } \mathbb{R}) \mid |f|^2 \text{ is integrable on } [a,b]\}$$

or, for short,

$$L^2[a,b] = \{f : [a,b] \to \mathbb{C} \mid \int_a^b |f(t)|^2\,dt < \infty\} \tag{19}$$

[3]These are Lebesgue integrable functions. They are quite close to the familiar Riemann integrable functions, but the advantage is that the Lebesue integrability behaves better when taking limits.

This notation is very suggestive: for practical purposes, if you have an f such that you can integrate $|f|^2$ (as a proper or improper integral), and the result is a finite number, then $f \in L^2$.

Examples of functions in L^2: continuous functions, functions with jump discontinuities (a finite number of jumps, or even countably many!), some functions which go to infinity: for example $x^{-1/4} \in L^2[0,1]$ because $\int_0^1 (x^{-1/4})^2 dx = \int_0^1 x^{-1/2} dx = 2x^{1/2}|_0^1 = 2 < \infty$.

A wrinkle: in $L^2[a,b]$, if two functions differ by their values at only a finite number of points (or even on a countable set) they are considered equal. For example, the functions f_1 and f_2 below are equal:

$$f_1(t) = \begin{cases} 0 & \text{if } t \in [0,1] \\ 1 & \text{if } t \in (1,2] \end{cases} \qquad f_2(t) = \begin{cases} 0 & \text{if } t \in [0,1) \\ 1 & \text{if } t \in [1,2] \end{cases} \qquad f_1 = f_2 \text{ in } L^2[0,2]$$

as indeed the L^2-norm $\|f_1 - f_2\| = 0$.

Therefore $L^2[a,b] = L^2(a,b)$.

Note the similarity of the definition (19) of L^2 with the definition of ℓ^2.

Note that by the Cauchy-Schwartz inequality, the L^2 inner product of two function in $L^2[a,b]$ is finite. Therefore the space $L^2[a,b]$ is a Hilbert space.

The space L^2 is also used on infinite intervals, like $L^2[a,\infty)$, and $L^2(\mathbb{R})$.

2.5.1. *An important variation of the L^2 space.* Instead of using the usual element of length dt we can use an element of "weighted" length $w(t)dt$. Here, $w > 0$. One physical interpretation is that if $[a,b]$ represents a wire of variable density $w(t)$ then the element of mass on the wire is $w(t)dt$. If $w(t)$ is a positive function define

$$L^2([a,b], w(t)dt) = \{f : [a,b] \to \mathbb{C} \mid \int_a^b |f(t)|^2 \, w(t)dt \, < \, \infty \}$$

with the inner product given by

$$\langle f, g\rangle_w = \int_a^b \overline{f(t)} \, g(t) \, w(t)dt$$

Note that $f \in L^2([a,b], w(t)dt)$ if and only if $\frac{f}{\sqrt{w}} \in L^2[a,b]$.

2.6. **The Banach space $C[a,b]$.**

Recall: a function continuous on a closed interval does have an absolute maximum and an absolute minimum.

The space of continuous function on $[a,b]$ is closed with respect to **the sup norm** $\|.\|_\infty$:

$$\|f\| = \|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$$

In other words, completeness of $C[a,b]$ in the sup-norm means that: if f_n are continuous on $[a,b]$ and if $f_n \to f$ in the sense that $\lim_{n\to\infty} \sup_{x\in[a,b]} |f_n(x) - f(x)| = 0$ then f is continuous.

2.7. The Banach spaces L^p.

The space

$$L^p[a,b] = \{f : [a,b] \to \mathbb{C}(\text{or } \mathbb{R}) \mid \int_a^b |f(t)|^p \, dt \, < \, \infty \}$$

is complete in the L^p norm $\|f\|_p = (\int_a^b |f|^p)^{1/p}$.

L^p are called the Lebesgue[4] spaces (only L^2 is a Hilbert space).

2.8. Closed sets, dense sets.

Definition 22. *Let (M,d) be a complete metric space.*

A subset $F \subset M$ is called closed *if it contains the limits of the Cauchy sequences in F: if $f_n \in F$ and $f_n \to f$ then $f \in F$.*

Examples: for $M = \mathbb{R}$ the intervals (a,b), $(a,+\infty)$ are not closed sets, but the intervals $[a,b]$, $[a,+\infty)$ are closed. For $M = \mathbb{R}^2$ the disk $x^2 + y^2 \leq 1$ is closed, but the disk $x^2 + y^2 < 1$ is not closed, and neither is the punctured disk $0 < x^2 + y^2 \leq 1$.

Definition 23. *Given a complete metric space (M,d) and a set $S \subset M$ we call the set $\overline{S}$* the closure of S in M *if $\overline{S}$ contains all the limits of the Cauchy sequences sequences $(f_n)_n \subset S$.*

Note that $S \subset \overline{S}$ (since if $f \in S$ we can take the trivial Cauchy sequence constantly equal to f, $f_n = f$ for all n.

Examples: for $M = \mathbb{R}$ the closure of the open interval (a,b) is the closed interval $[a,b]$. For $M = \mathbb{R}^2$ the closure of the (open) disk $x^2 + y^2 < 1$, and of the punctured disk $0 < x^2 + y^2 \leq 1$ is the closed disk $x^2 + y^2 \leq 1$.

Definition 24. *Given a complete metric space (M,d) and a set $S \subset M$ we say that* a set S is dense in M if $\overline{S} = M$.

It is worth repeating: S is dense in M means that any $f \in M$ can be approximated by elements of S: there exists $f_n \in S$ such that $\lim_{n\to\infty} f_n = f$.

Examples: for $M = \mathbb{R}$ the rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$ (why?). For $M = \ell^2$ the sequences that terminate[5] are dense in ℓ^2.

Dense sets have important practical consequences: when we need to establish a property of M that is preserved when taking limits (e.g. equalities or inequalities), then we can prove the property on S and then, by taking limits, we find it on M.

[4]Henri Lebesgue (1875-1941) was a French mathematician most famous for Lebesgue's theory of integration.

[5]These are the sequences $x = (x_1, x_2, \dots)$ such that $x_n = 0$ for all n large enough.

2.9. **Sets dense in the Hilbert space L^2.**

$C[a,b]$ is dense (in the L^2 norm) in $L^2[a,b]$ (by our construction of L^2).

In fact, we can assume functions as smooth as we wish, and still obtain dense spaces: the smaller space $C^1[a,b]$, of functions which have a continuous derivative, is also dense in $L^2[a,b]$, and so is $C^r[a,b]$, functions with r continuous derivatives, for any r (including $r=\infty$).

Even the smaller space $\mathcal{P}$ (of polynomials) is dense in $L^2[a,b]$.

Another type of (inner product sub)spaces dense in L^2 are those consisting of functions satisfying zero boundary conditions, for example

$$\{f \in C[a,b] \,|\, f(a) = 0\} \tag{20}$$

(Why: any function f in (20) can be approximated in the L^2-norm by functions in (20): consider the sequence of functions f_n which equal f for $x \geq a + \frac{1}{n}$ and whose graph is the segment joining $(a,0)$ to $(a+\frac{1}{n}, f(a+\frac{1}{n}))$ for $a \leq x < a + \frac{1}{n}$. Then the L^2 norm of $f_n - f$ converges to zero, much like in the Example 3. of §2.3.)

Similarly,

$$C_0[a,b] = \{f \in C[a,b] \,|\, f(a) = f(b) = 0\}$$

is dense in $L^2[a,b]$, and so is the following very useful space

$$C_0^1[a,b] = \{f \,|\, f' \text{ continuous on } [a,b],\ f(a) = f(b) = 0,\ f'(a) = f'(b) = 0\}$$

2.10. **Polynomials are dense in the Banach space $C[a,b]$.** The space of polynomials is dense in $C[a,b]$ in the sup norm. (This is the Weierstrass Theorem: any continuous function f on $[a,b]$ can be uniformly approximated by polynomials, in the sense that there is a sequence of polynomials p_n such that $\lim_{n\to\infty} \sup_{[a,b]} |f - p_n| = 0$.)

Please note that closure and density are relative to a norm. For example, $C[a,b]$ is not closed in the L^2 norm, but it is closed in the sup-norm.

3. Hilbert Spaces

In a Hilbert space we can do linear algebra (since it is a vector space), geometry (since we have lengths and angles) and calculus (since it is complete). And they are all combined when we write series expansions.

Recall:

Definition 25. *A Hilbert space H is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product.*

Recall the two fundamental examples: the space of sequences ℓ^2, and the space of square integrable function L^2.

3.1. **When does a norm come from an inner product?**

In every Hilbert space the **parallelogram identity** holds: for any $f, g \in H$

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2) \tag{21}$$

(in a parallelogram the sum of the squares of the sides equals the sum of the squares of the diagonals).

Relation (21) is proved by a direct calculation (expanding the norms in terms of the inner products and collecting the terms).

The remarkable fact is that, if the parallelogram identity holds in a Banach space, then its norm actually comes from an inner product, and can recover the inner product only in terms of the norm by
the polarization identity:
for complex spaces

$$\langle f, g\rangle = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\right)$$

and for real spaces

$$\langle f, g\rangle = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2\right)$$

3.2. **The inner product is continuous.**

This means that we can take limits inside the inner product:

Theorem 26. *If $f_n \in H$, with $f_n \to f$ then $\langle f_n, g\rangle \to \langle f, g\rangle$.*

Indeed: $|\langle f_n, g\rangle - \langle f, g\rangle| = |\langle f_n - f, g\rangle| \leq$ (by Cauchy-Schwartz) $\|f_n - f\| \, \|g\| \to 0$. (Recall: $f_n \to f$ means that $\|f_n - f\| \to 0$.)

3.3. **Orthonormal bases.**

Consider a Hilbert space H.

Just like in linear algebra we define:

Definition 27. *If $\langle f, g\rangle = 0$ then $f, g \in H$ are called orthogonal ($f \perp g$).*

and

Definition 28. *A set $B \subset H$ is called orthonormal if all $f \neq g \in B$ are orthogonal ($f \perp g$) and unitary ($\|f\| = 1$).*

Note that as in linear algebra, an orthonormal set is a linearly independent set (why?).

And departing from linear algebra:

Definition 29. *A set $B \subset H$ is called an* **orthonormal basis** *for H if it is an orthonormal set and it is complete in the sense that the span of B is* dense *in H: $\overline{Sp(B)} = H$.*

Note that in the above B is not a basis in the sense of linear algebra (unless H is finite-dimensional). But like in linear algebra:

Theorem 30. *Any Hilbert space admits an orthonormal basis; furthermore, any two orthonormal bases of the same space have the same cardinality, called the Hilbert dimension of the space.*

From here on we will only consider Hilbert spaces which admit a (finite or) countable orthonormal basis.

These are the Hilbert spaces usually encountered in applications.

It can be proved that this condition is equivalent to the existence of a countable set S dense in H. This condition is often easier to check, and the property is called "H is separable". Application: $L^2[a,b]$ is separable (why?) therefore $L^2[a,b]$ has a countable orthonormal basis. Many physical problems are solved by finding special orthonormal basis of $L^2[a,b]$!

[For your amusement: it is quite easy to construct a Hilbert space with an uncountable basis, e.g. just like we took $\ell^2(\mathbb{Z}_+)$ we could take $\ell^2(\mathbb{R})$.]

We assume from now on that our Hilbert spaces H do have a countable orthonormal basis.

The following theorem shows that many of the properties of inner product vector spaces which are finite dimensional (think $\mathbb{R}^n$ or $\mathbb{C}^n$ with the standard basis) are very similar for Hilbert spaces; the main difference is that in infinite dimensions instead of sums we have series - which means sums followed by limits, see for example the infinite linear combination (22).

Theorem 31. *Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space with a countable basis.*

Let $u_1, \ldots, u_n, ..$ be an orthonormal basis.

The following hold.

(i) Let $c_1, \ldots, c_n, ..$ be scalars such that $(c_1, \ldots, c_n, ..) \in \ell^2$. Then the series

$$f = \sum_{n=1}^{\infty} c_n u_n \tag{22}$$

converges and its sum $f \in \mathcal{H}$.

Moreover

$$\|f\|^2 = \sum_{n=1}^{\infty} |c_n|^2$$

(ii) Conversely, every $f \in \mathcal{H}$ has an expansion (22). The scalars c_n satisfty $c_n = \langle u_n, f\rangle$ and are called **generalized Fourier coefficients** *of f.*

Therefore any $f \in \mathcal{H}$ can be written as a **generalized Fourier series***:*

$$f = \sum_{n=1}^{\infty} \langle u_n, f\rangle\, u_n \tag{23}$$

and **Parseval's identity** *holds:*

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle u_n, f\rangle|^2 \tag{24}$$

As a consequence, **Bessel's inequality** *holds:*

$$\|f\|^2 \geq \sum_{n\in J} |f_n|^2 \qquad \text{for any } J \subset \mathbb{Z}_+ \tag{25}$$

(iii) If $f, g \in \mathcal{H}$ then

$$\langle f, g\rangle = \sum_{n=1}^{\infty} \langle f, u_n\rangle \, \langle u_n, g\rangle \tag{26}$$

Proof. We prove (i), (ii); (iii) is similar. Since, by definition $\sum_{n=1}^{\infty} = \lim_{N\to\infty}\sum_{n=1}^{N}$, the series is convergent if and only if it is Cauchy, meaning

$$(*)\|\sum_{n=1}^{N+P} c_n u_n - \sum_{n=1}^{N} c_n u_n\|^2 = \|\sum_{n=N+1}^{N+P} c_n u_n\|^2$$

converges to zero uniformly in N, P. But

$$\|\sum_{n=N+1}^{N+P} c_n u_n\|^2 = \sum_{n=N+1}^{N+P} |c_n|^2 \|u_n\|^2 = \sum_{n=N+1}^{N+P} |c_n|^2$$

(inequality would follow by Cauchy-Schwarz. Equality holds because $\|a + b\|^2 = \langle a+b, a+b\rangle = \|a\|^2 + \langle a, b\rangle + \langle b, a\rangle + \|b\|^2$). Now the the last term in (*) converges to zero uniformly in N, P since we assumed $(c_1, \ldots, c_n, ..) \in \ell^2$. Parseval follows in the same way.

For (ii), let $f = \sum_{n=1}^{\infty} c_n u_n = \lim_{n\to\infty} S_N := \sum_{n=1}^{N} c_n u_n$ Clearly,

$$\langle u_j, S_N\rangle = c_j$$

and by Theorem 39,

$$\langle u, f\rangle = \lim_{N\to\infty} \langle u_j, S_N\rangle = c_j$$

□

Note that (separable) Hilbert spaces are essentially ℓ^2, since given an orthonormal basis $u_1, u_2, \ldots$, the elements $f \in H$ can be identified with the sequence of their generalized Fourier coefficients $(c_1, c_2, c_3, \ldots)$ which, by (24), belongs to ℓ^2.

Remark. If $v_1, \ldots, v_n, \ldots$ form an orthogonal basis, but not an orthonormal basis (some v_n are not unit vectors) then one can produce an orthonormal basis by setting $u_n = v_n/\|v_n\|$, which used in formula (23) gives the expansion of f in terms of v_n as

$$f = \sum_{n=1}^{\infty} \frac{\langle v_n, f\rangle}{\|v_n\|^2} \, v_n \tag{27}$$

3.4. **Generalized Fourier series in L^2.** If $H = L^2[a,b]$ then (22) means that

$$\lim_{N\to\infty} \int_a^b \left|f(x) - \sum_{n=1}^{N} c_n u_n(x)\right|^2 dx = 0 \tag{28}$$

(we took the square of the norm rather than the norm).

Formula (28) is often expressed as

(29) f equals $\sum_{n=1}^{\infty} c_n u_n$ in the least square sense, or *in mean square*

see also §3.8.

Only if $f(x)$ and $u_n(x)$ are smooth enough it is true that the series $\sum c_n u_n$ is point-wise convergent, meaning that

$$f(x) = \sum_{n=1}^{\infty} c_n u_n(x) \tag{30}$$

(precise conditions can be given) But this is **not** the case for every $f \in L^2$. In general, the series on the right side of (30) may not converge for all x, and even if it converges, it may not equal $f(x)$.

3.5. **Example of orthonormal basis in $L^2[a,b]$: Fourier series.**

The functions $\sin x$, $\cos x$ have period 2π. The same is true for any linear combination of $\sin nx$, $\cos nx$: any **trigonometric polynomial**[6]

$$T(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left[a_n \cos(nx) + b_n \sin(nx)\right] \tag{31}$$

is a periodic function, period 2π.

What if instead of a finite sum in (31) we consider a series? Periodicity should survive taking limits (if the sequence of functions is not oscillating too wildly).

Exercise. Check that the functions 1, $\sin(nx)$, $\cos(nx)$, $(n = 1, 2, 3\ldots)$ form an orthogonal system in $L^2[-\pi, \pi]$. Then normalize these functions, to obtain an orthonormal system.

Moreover

Theorem 32. *The functions* 1, $\sin(nx)$, $\cos(nx)$, $(n = 1, 2, 3\ldots)$ *form a basis for the real-valued Hilbert space* $L^2([-\pi, \pi], \mathbb{R})$

[6] It is convenient to write the constant term as $a_0/2$ rather than a_0 such that formula (33) applies for all $n \geq 0$, and not need a separate formula for a_0.

The proof of this important and deep theorem is not included here.

Is is clear that if we allow the scalars a_n, b_n to be complex numbers, then the same trigonometric monomials form a basis for the complex-valued Hilbert space $L^2([-\pi,\pi],\mathbb{C})$.

It follows that if $f(x) \in L^2[-\pi,\pi]$ then $f(x)$ can be approximated (in mean squre) by the partial sums of a Fourier series:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \tag{32}$$

Exercise. Assuming that $f \in L^2[-\pi,\pi]$ has its Fourier series expansion (32) verify that the Fourier coefficients a_n and b_n are given by the formulas

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\, \cos(nx)\, dx \ , \ b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\, \sin(nx)\, dx \tag{33}$$

A more concise and compact formula can be written using Euler's formula $e^{inx} = \cos(nx) + i\sin(nx)$.

Exercise. Check that the functions e^{inx}, $n \in \mathbb{Z}$ form an orthogonal system in the complex valued functions $L^2[-\pi,\pi]$. Normalize them to obtain an orthonormal system.

An orthonormal basis for the complex Hilbert space$L^2[-\pi,\pi]$ is thus found, and any complex valued function in $L^2[-\pi,\pi]$ can be expanded in a Fourier series

$$\sum_{n=-\infty}^{\infty} \hat{f}_n e^{inx}$$

with the Fourier coefficients given by

$$\hat{f}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-inx}\, dx$$

(why?) The series converges in square mean (i.e. in the L^2 norm), and, moreover, pointwise if $f(x)$ is "smooth enough" (precise mathematical formulations form the topic of one domain of mathematics, harmonic analysis).

Exercise. Show that the Fourier coefficients $a_n, b_n, \hat{f}_n$ are related by the formulas

$$a_n = \hat{f}_n + \hat{f}_{-n} \text{ for } n = 0, 1, 2, \ldots, \quad b_n = i(\hat{f}_n - \hat{f}_{-n}) \text{ for } n = 1, 2, \ldots$$

Find the conditions on $\hat{f}_n$ equivalent to the fact that function $f(x)$ is real valued.

On a general interval $[a,b]$ the Fourier series is generated by expansion in terms of $exp(2\pi inx/(b-a))$ (with $n \in \mathbb{Z}$) as it is easily seen by a rescaling of x.

Exercise. Use a linear change of the variable x (i.e. setting $x = cx+d$ for suitable scalars c, d) to show that the Fourier series of a function $g \in L^2[a,b]$ has the form

$$g = \sum_{n=-\infty}^{\infty} \hat{g}_n e^{2\pi i n x/(b-a)}$$

and find the formula that expresses $\hat{g}_n$ in terms of $g(x)$.

3.6. **Other bases for L^2: orthogonal polynomials.**

Recall that polynomials are dense in $L^2[a,b]$. The fact that every function in L^2 can be approximated by polynomials is extremely useful in applications. It would be even better to have an orthonormal basis of polynomials.

The Legendre orthogonal polynomials are polynomials which form a basis for the vector space $\mathcal{P}$ and orthogonal in $L^2[-1,1]$.

Exercise. Use a Gram-Schmidt process on the polynomials $1, x, x^2, x^3$ to obtain an orthonormal set; these are the first four Legendre polynomials.

More generally, consider weighted L^2 spaces, $L^2([a,b],\, w(x)dx)$. As we noted, the polynomials form a dense set, and they are spanned by $1, x, x^2, \ldots$ The Gram-Schmidt process (with respect to the weighted inner product) produces a sequence of orthonormal polynomials which span a dense set in the weighted L^2, hence it forms an orthonormal basis for $L^2([a,b],\, w(x)dx)$.

Orthogonal polynomials (with respect to a given weight, on a given interval) have been playing a fundamental role in many areas of mathematics and its applications, and are invaluable in approximations.

Exercise. The Laguerre orthogonal polynomials are orthogonal in the weighted $L^2([0,+\infty), e^{-x}dx)$. Use a Gram-Schmidt process on the polynomials $1, x, x^2, x^3$ to obtain an orthonormal set; these are the first four Laguerre polynomials.

3.7. **Orthogonal complements, The Projection Theorem.** The constructions are quite similar to the finite-dimensional case. One important difference is that we often need to assume that subspaces are closed, or otherwise take their closure.

Definition 33. *If S is a subset of a Hilbert space H its orthogonal is*

$$S^\perp = \{f \in H \,|\, \langle s, f\rangle = 0, \text{ for all } s \in S\}$$

Remark that $S^\perp$ is a **closed subspace**, therefore it is a Hilbert space itself.

Why: as in linear algebra, $S^\perp$ is a vector subspace; to see that it is closed take a sequence $f_n \in S^\perp$ such that $f_n \to f$, and show that $f \in S^\perp$. Indeed, for any $s \in S$ we have $0 = \langle f_n, s\rangle \to \langle f, s\rangle$ (by Theorem 36) so $f \in S^\perp$. $\Box$

Definition 34. *If V is a* closed subspace *of H, then $V^\perp$ is called the orthogonal complement of V.*

Note that if V is a vector subspace (not necessarily closed) then $(V^\perp)^\perp = \overline{V}$ (another point to be careful about in infinite diemensions).

3.7.1. *Orthogonal Projections.* Like in finite dimensions, in Hilbert spaces the projection f_V of f onto a subspace V is the vector that minimizes the distance between f and V. Note however: when we try to upgrade a statement from finite to infinite dimensions we often need to assume that our subspaces are closed, and if they are not, to replace them by their closure.

Theorem 35. *Given V a* closed subspace *of H and $f \in H$ there exists a unique $f_V \in V$ which minimizes the distance from f to V:*

$$\|f - f_V\| = \min\{\|f - g\| \,|\, g \in V\}$$

f_V is called the **orthogonal projection of f onto** *V.*

The outline of the proof is as follows: let $g_n \in V$ so that $\|f - g_n\| \to \inf\{\|f-g\| \,|\, g \in V\}$. A calculation (expanding $\|g_n - g_m\|^2$ and using Cauchy-Schwartz) yields that the sequence $(g_n)_n$ is Cauchy. Since V was assumed closed, then g_n converges, and f_V is its limit. $\square$

The linear operator $P_V : H \to H$ which maps f to f_V is called the **orthogonal projection onto** V.

Like in finite dimensions, H decomposes as a direct sum between a subspace V and its orthogonal complement: but V must to be **closed**:

Theorem 36. *(The Projection Theorem)*

Let V be a **closed subspace** *of H.*

Every $f \in H$ can be written uniquely as $f = f_V + f_\perp$, with $f_V \in V$ and $f_\perp \in V^\perp$.

In other words, $H = V \oplus V^\perp$.

The Projection Theorem follows by proving that $f - f_V$ is orthogonal to V (which follows after calculations not included here). $\square$

3.8. **Least squares approximation via subspaces.** Consider an orthonormal basis of H: $u_1, u_2, u_3, \ldots$. Any given $f \in H$ can be written as a (convergent) series

$$f = \sum_{n=1}^{\infty} \hat{f}_n u_n, \qquad \text{where } \hat{f}_n = \langle u_n, f\rangle$$

We can approximate f by finite sums, with control of the errors, in the following way.

It is clear that the N^{th} partial Fourier sum:

$$f^{[N]} = \sum_{n=1}^{N} \hat{f}_n u_n$$

represents the orthogonal projection of f onto $V_N = \mathrm{Sp}(u_1, u_2, \ldots, u_N)$ (why?). Since V_N is closed (it is finite-dimesional, hence it is essentially $\mathbb{R}^N$ or $\mathbb{C}^N$) then by Theorem 35 we have that

$$\|f - f^{[N]}\| = \min\{\|f - g\| \,|\, g \in V_N\} = \min\{\, E_N(c_1, \ldots, c_N) \big| \, c_k \in \mathbb{C}\}$$

$$\text{where} \;\; E_N(c_1, \ldots, c_N) = \|f - \sum_{n=1}^{N} c_n u_n\|$$

The function $E_N^2(c_1, \ldots, c_N)$ is called *Gauss' mean squared error.* For example, for $H = L^2[a, b]$ we have

$$E_N^2(c_1, \ldots, c_N) = \int_a^b \left| f(x) - \sum_{n=1}^{N} c_n u_n(x) \right|^2 dx$$

Note that $\left| f(x) - \sum_{n=1}^{N} c_n u_n(x) \right|^2 \equiv s(x)$ is the squared error at x, and then $\frac{1}{b-a} \int_a^b s(x)dx$ is its mean.

Assume the Hilbert space is over the reals. To minimize the squared error (solving the least squares problem) we look for stationary points, hence we solve the system

$$\frac{\partial E_N^2}{\partial c_k} = 0, \;\; k = 1, \ldots, N$$

whose solution is (try it!) $c_k = \langle u_k, f\rangle = \hat{f}_k$ (for $k = 1, \ldots, N$).

The error in approximating f by $f^{[N]}$ is $E_N(\hat{f}_1, \ldots, \hat{f}_N)$ therefore

$$E_N^2(\hat{f}_1, \ldots, \hat{f}_N) = \|f - f^{[N]}\|^2 = \| \sum_{n=N+1}^{\infty} \hat{f}_n u_n\|^2 = \sum_{n=N+1}^{\infty} |\hat{f}_n|^2 \leq \|f\|^2$$

4. Linear operators in Hilbert spaces

Let H be a Hilbert space over the scalar field $F = \mathbb{R}$ or $\mathbb{C}$, with a countable orthonormal basis $u_1, u_2, u_3, \ldots$.

In order to extend the concept of a matrix to infinite dimensions, it is preferable to look at the linear operator associated to it.

Definition 37. *An operator $T : H \to H$ is called linear if $T(f + g) = Tf + Tg$ and $T(cf) = cTf$ for all $f, g \in H$ and $c \in F$.*

The linearity conditions are often written more compactly as

$$T(cf + dg) = cTf + dTg \quad \text{for all } f, g \in H, \text{ and } c, d \in F$$

The adjoint T^* of an operator T is defined as in the finite-dimensional case, by $\langle Tf, g\rangle = \langle f, T^*g\rangle$ for all f, g.

We will see that it is sometimes convenient to consider operators T which are defined on a domain $D(T) \subset H$ which is smaller than the whole Hilbert space H. In such cases T^* is (usually) also defined on a domain $D(T^*) \subset H$ and we will call T^* the *adjoint* of T.

We will call the operator T *formally selfadjoint* (or symmetric) if

$$\langle Tf, g\rangle = \langle f, Tg\rangle \quad \text{for all } f \in D(T),\ g \in D(T^*)$$

We call T *selfadjoint* if it is formally selfadjoint and $D(T) = D(T^*)$.

4.1. Shift operators on ℓ^2.

The examples below illustrate a very important distinction of infinite dimension: for a linear operator to be invertible we need to check both that its kernel is null (meaning that the operator is one-to-one) and that its range is the whole space (the operator is onto). We will see that the two conditions are not equivalent in infinite dimensions, and that an isometry, while it clearly is one-to-one, need not be onto.

1. Consider the left shift operator $S : \ell^2 \to \ell^2$:

$$S(x_1, x_2, x_3, x_4, \ldots) = (x_2, x_3, x_4, \ldots)$$

which is clearly linear.

Obviously $\mathrm{Ker}(S) = Sp(e_1)$ (S is not one-to-one) and $\mathrm{Ran}(S) = \ell^2$ (S is onto).

2. Consider the right shift operator R on ℓ^2:

$$R(x_1, x_2, x_3, x_4, \ldots) = (0, x_1, x_2, x_3, \ldots)$$

The operator R is clearly linear. Note that R is an isometry, since

$$\|Rx\|^2 = \sum_{n=1}^{\infty} |x_n|^2 = \|x\|^2$$

thus we have $\mathrm{Ker}(R) = \{0\}$ (R is one-to-one). But R is not onto since $\mathrm{Ran}(R) = \{x \in \ell^2 \mid x_1 = 0\}$.

4.2. Unitary operators. Isomorphic Hilbert spaces.

Definition 38. *A linear operator $U : H \to H$ is called* unitary *if $UU^* = U^*U = I$ (i.e. its inverse equals it adjoint).*

This definition is similar to the finite dimensional case, and, it follows that if U is unitary, then U is an isometry. Example 2 in §4.1 shows that the converse is not true in infinite dimensions (an isometry may not be onto). It can be shown that

Theorem 39. *If a linear operator is an isometry, and is onto, then it is unitary.*

We noted that if H is a Hilbert space with a countable orthonormal basis then H is essentially ℓ^2. Mathematically, this is stated as follows:

Theorem 40. *Let H be a Hilbert space with a countable orthonormal basis. Then H* is isomorphic to *ℓ^2 in the sense that there exists a unitary operator $U : \ell^2 \to H$.*

To be precise, if H is a complex Hilbert space, then we consider the complex ℓ^2 (sequences of complex numbers), while if H is a real Hilbert space, then we consider the real ℓ^2 (sequences of real numbers.)

The proof of Theorem 40 is immediate: the operator U is constructed in the obvious way. Denoting by $u_1, u_2, \ldots$ an orthonormal basis of H we define

$$U(a_1, a_2, \ldots) = a_1 u_1 + a_2 u_2 + \ldots = \sum_{n=1}^{\infty} a_n u_n \quad \text{for each } (a_1, a_2, \ldots) \in \ell^2$$

We need to show that $\sum_{n=1}^{\infty} a_n u_n \in H$, that U is one-to-one (clearly true by the definition of the Hilbert space basis) and onto (intuitively clear but requires an argument). Finally, by the Parseval's identity U is an isometry, and by Theorem 39 it is unitary. □

Other examples of unitary operators.

It can be shown that the Fourier transform is a unitary operator on $L^2(\mathbb{R})$.

4.3. **Integral operators.**

4.3.1. *Illustration - solutions of a first order differential equation.* The simplest integral operator takes f to $\int_a^x f(s)\, ds$ (it is clearly linear!). More general integral operators have the form

$$Kf(x) = \int_a^b G(x,s) f(s)\, ds \tag{34}$$

where $G(x,s)$ (called "integral kernel") is continuous, or has jump discontinuities.

(Note the similarity of (34) with the action of a matrix on a vector: if $G = (G_{xs})_{x,s}$ is a matrix, and $f = (f_s)_s$ is a vector, the $(Gf)_x = \sum_s G_{xs} f_s$.)

Integral operators appear as solutions of nonhomogenous differential equations; then G is called the Green's functions.

As a first example, consider the differential problem

$$\frac{dy}{dx} + y = f(x), \qquad y(0) = 0 \tag{35}$$

where we look for solutions $y(x)$ for x in a finite interval, say $x \in [0,1]$.

It is well known that the general solution of (35) is obtained by multiplying by the integrating factor (exp of the integral of the coefficient of y) giving

$$\frac{d}{dx}\left(e^x y\right) = e^x f(x) \quad \text{therefore} \quad y(x) = Ce^{-x} + e^{-x} \int e^x f(x)\, dx$$

Imposing the initial condition $y(0) = 0$ we obtain the solution

$$y(x) = e^{-x} \int_0^x e^s f(s)\, ds \tag{36}$$

which has the form (34) for the (Green's function of the problem)

$$G(x,s) = e^{-x}\, e^s\, \chi_{[0,x]}(s) \tag{37}$$

where $\chi_{[0,x]}$ is the characteristic function of the interval $[0,x]$, defined as

$$\chi_{[0,x]}(s) = \begin{cases} 1 & \text{if } s \in [0,x] \\ 0 & \text{if } s \notin [0,x] \end{cases} \tag{38}$$

In many applications it is important to understand how solutions of (35) depend on the nonhomogeneous term f, for example, if a small change of f produces only a small change of the solution $y = Kf$. This is the case if

$$\text{there is a constant } C \text{ so that } \quad \|Kf\| \leq C\,\|f\| \quad \text{for all } f \tag{39}$$

If an operator K satisfies (39) then K is called *bounded*. In fact, condition (39) is equivalent to the fact that the linear operator K is *continuous*.

It is not hard to see that the integral operator K given by (34), (37) is bounded on $L^2[0,1]$; this means that the mean squared average of Kf does not exceed a constant times the mean squared average of f (the estimates are shown in §4.3.2 below).

We can find better estimates, point-wise rather than in average, since the linear operator K is also bounded as an operator on the Banach space $C[0,1]$, and this means that the maximum of $|Kf(x)|$ does not exceed C times the maximum of $|f|$ (see the proof of this estimate below in §4.3.3).

While the estimate in $C[0,1]$ (in the sup-norm) is stronger and more informative than the estimate in $L^2[0,1]$ (in square-average), we have to pay a price: we can only obtain them for continuous forcing terms f.

General principle: **integral operators do offer good estimates**.

4.3.2. *Proof that K is bounded on $L^2[0,1]$ (Optional).* Let us denote the norm in $L^2[0,1]$ by $\|\cdot\|_{L^2}$, and the inner product by $\langle\cdot,\cdot\rangle_{L^2}$ to avoid possible confusions.

To see that K is bounded on $L^2[0,1]$ note that $Kf(x)$ is, for each fixed x, an L^2 inner product, and using the Cauchy-Schwartz inequality

$$|Kf(x)| = |\langle G(x,\cdot), f(\cdot)\rangle_{L^2}| \leq \|G(x,\cdot)\|_{L^2}\,\|f\|_{L^2} \tag{40}$$

We can calculate using (37)

$$\|G(x,\cdot)\|_{L^2}^2 = \int_0^1 G(x,s)^2\,ds = \int_0^x e^{2s-2x}ds = \frac{1}{2}\left(1-e^{-2x}\right) < \frac{1}{2}$$

which used in (40) gives

$$|Kf(x)| \leq \frac{1}{2}\,\|f\|_{L^2} \tag{41}$$

for any $f \in L^2[0,1]$. Then from (41) we obtain

$$\|Kf\|_{L^2}^2 = \int_0^1 |Kf(x)|^2 dx \leq \frac{1}{4}\,\|f\|_{L^2}^2 \tag{42}$$

so K satisfies (39) in the L^2-norm for $C = 1/2$. □

4.3.3. *Proof that K is bounded on $C[0,1]$ (Optional).* Recall that $C[0,1]$ is a Banach space with the sup-norm $\|\cdot\|_\infty$. Note that we obviously have $|f(x)| \leq \|f\|_\infty$ for any continuous f.

We need to use the following (very useful) inequality:

$$\Big|\int_a^b F(s)\,ds\Big| \leq \int_a^b |F(s)|\,ds$$

which is intuitively clear if we think of the definite integral of a function as the signed area between the graph of the function and the x-axis.

We then have, for any $f \in C[0,1]$,

$$|Kf(x)| = \Big|e^{-x}\int_0^x e^s f(s)\,ds\Big| \leq e^{-x}\int_0^x e^s\,|f(s)|\,ds$$

$$\leq \|f\|_\infty\, e^{-x}\int_0^x e^s\,ds = \|f\|_\infty\,(1-e^{-x}) \leq \|f\|_\infty$$

hence K satisfies (39) in the sup-norm for $C = 1$. □

4.3.4. *Remarks about the δ function.* This δ is "defined" by the property $\delta(x) = 0$ if $x = 0$ and $\int_{-\infty}^{\infty}\delta(s)ds = 1$.

The first remark is that ... there is no such function. There are objects more general than functions, they are called *distributions* and δ is a distribution.

However, in engineering and physics applications it is treated as though it were a function. This leads to correct results in many cases, not all, and one has to be careful.

For example, an alternative method for finding the Green's function of (35) is by solving $\frac{dy}{dx} + y = \delta(x-s)$ where δ is Dirac delta "function". Its solution turns out to be $G(x,s)$ in (37). Then, using the fundamental property of the delta function that

$$f(x) = \int \delta(x-s)f(s)ds \tag{43}$$

by superposition of solutions we obtain (36).

The study of distributions and of their applications form a separate chapter in mathematics.

4.4. **Differential operators in $L^2[a,b]$.** The simplest differential operator is $\frac{d}{dx}$, which takes f to $\frac{df}{dx}$. It is clearly linear.

There are two fundamental differences between the differential operator $\frac{d}{dx}$ and the integral operators (34):

1) we cannot define the derivative on all the functions in L^2, and

2) $\frac{d}{dx}$ is not a bounded operator: we cannot estimate the derivative of a function by only knowing the magnitude of the function (not in the sup-norm, and not even in average). Indeed, you can easily imagine a smooth function, whose values are close to zero and never exceed 1, but which has

a very narrow and sudden spike. The narrower the spike, the higher the derivative, even if the average of the function remains small.

4.4.1. *Illustration on a first order problem.* Consider again the problem (35). It can be written as

$$Ly = f, \quad y(0) = 0$$

where L is the differential operator

$$L = \frac{d}{dx} + 1, \quad \text{therefore } Ly = \left(\frac{d}{dx} + 1\right) y = \frac{dy}{dx} + y$$

If we are interested to allow the nonhomogeneous term f to have jump discontinuities, since $y' = f - y$ then y' will have jump discontinuities (at such a point y' may not be defined). An example of a such function is:

$$\frac{d}{dx}|x| = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$$

A good domain for L is then

$$D = \{y \in L^2[0,1] \,|\, y' \in L^2[0,1],\, y(0) = 0\}$$

Note that the initial condition was incorporated in D. Note also that D is dense in $L^2[0,1]$.

Note that the integral operator K found in §4.3.1 is the inverse of the differential operator L (defined on D).

4.5. **A second order boundary value problem.** Quite often separable partial differential equations lead to second order boundary value problems, illustrated here on a simple example.

Problem: Find the values of the constant λ for which the equation

$$y'' + \lambda y = 0 \tag{44}$$

has non-identically zero solutions for $x \in [0, \pi]$ satisfying the boundary conditions

$$y(0) = 0, \; y(\pi) = 0 \tag{45}$$

Equation (44) models the simple vibrating string: $x \in [0,1]$ represents the position on the string, and $y(x)$ is the displacement at position x. The boundary conditions (45) mean that the endpoints $x = 0$ and $x = 1$ of the string are kept fixed.

4.5.1. *Formulation of the problem using a differential operator.* Unbounded operators are trickier. If A is a bounded operator, defined everywhere, and for all x, y $\langle x, Ay\rangle = \langle Ax, y\rangle$, then A is self-adjoint.

But if A is unbounded, and the condition above holds for all x, y *where A is defined*, then A is called *formally self-adjoint* or *symmetric*. *Formally self-adjoint* is not enough to endure self-adjointness. For now, we'll work, once more, formally.

Equation (44) can be written as

$$Ly = \lambda y, \quad \text{where } L = -\frac{d^2}{dx^2} \tag{46}$$

with the domain (dense in $L^2[0,\pi]$)

$$D = \{y \in L^2[0,\pi] \,|\, y'' \in L^2[0,\pi],\ y(0) = 0,\ y(\pi) = 0\} \tag{47}$$

(it must be noted that D needs to be more precisely specified, stating conditions on y and y').

A solution of (44), (45) is a solution of (46) in the domain (47). If this solution is not (identically) zero, then it is an eigenfunction of L, corresponding to the eigenvalue λ.

Note that L is formally self-adjoint on D since for $y, g \in D$ we have

$$\langle Ly, g\rangle = \int_0^\pi \overline{Ly(x)}g(x)\,dx = -\int_0^1 \overline{y''(x)}g(x)\,dx$$

and integration by parts gives (since $g(0) = 0 = g(\pi)$)

$$= -\overline{y'(x)}g(x)\big|_0^\pi + \int_0^\pi \overline{y'(x)}g'(x)\,dx = \int_0^\pi \overline{y'(x)}g'(x)\,dx$$

and integrating by parts again (and using $y(0) = 0 = y(\pi)$)

$$= \overline{y(x)}g(x)\big|_0^\pi - \int_0^\pi \overline{y(x)}g''(x)\,dx = \langle y, Lg\rangle$$

(Note that the boundary terms vanish by virtue of the boundary conditions.)

It can be shown, just like in the finite-dimensional case, that since L is formally self-adjoint then its eigenvalues are real and the eigenfunctions of L corresponding to different eigenvalues are orthogonal (see §4.5.2).

Remark that $L = -\frac{d^2}{dx^2}$ is positive definite, motivating the choice of a negative sign in front of the second derivative. Indeed, for $y \neq 0$

$$\langle Ly, y\rangle = -\int_0^1 \overline{y''(x)}y(x)\,dx = \int_0^\pi \overline{y'(x)}y'(x)\,dx = \int_0^\pi |y'(x)|^2\,dx > 0$$

where the last inequality is strict because $\int_0^\pi |y'(x)|^2\,dx = 0$ implies $y' = 0$, therefore y is constant, and due to the zero boundary conditions then $y = 0$.

Then the eigenvalues of L are positive, like in the finite-dimensional case. Indeed, if $Lf = \lambda f$ $(f \neq 0)$ then $\langle f, Lf\rangle = \langle f, \lambda f\rangle$ hence

$$-\int_0^1 \overline{f(x)}\,f''(x)\,dx = \lambda\int_0^1 |f(x)|^2\,dx$$

then integrating by parts and using the fact that $f(0) = f(1) = 0$ we obtain

$$\int_0^1 |f'(x)|^2\,dx = \lambda\int_0^1 |f(x)|^2\,dx$$

which implies $\lambda > 0$.

Let us find the eigenvalues and eigenfunctions of L. Since $\lambda \neq 0$ the general solution of (44) is $y = C_1 e^{\sqrt{-\lambda}x} + C_2 e^{-\sqrt{-\lambda}x}$ (note that $\sqrt{-\lambda} = i\sqrt{\lambda}$ since $\lambda > 0$). Since $y(0) = 0$ it follows that $C_2 = -C_1$, so, choosing $C_1 = 1$, $y(x) = e^{i\sqrt{\lambda}x} - e^{-i\sqrt{\lambda}x}$. The condition $y(\pi) = 0$ implies that $e^{2i\sqrt{\lambda}\pi} = 1$ therefore $\sqrt{\lambda} \in \mathbb{Z}$ so $\lambda = \lambda_n = n^2$, $n = 1, 2, \ldots$ and the corresponding eigenfunctions are $y_n(x) = e^{inx} - e^{-inx} = 2i\sin(nx)$.

A proof is needed to show that the eigenfunctions are complete, i.e. they form a basis for the Hilbert space $L^2[0, \pi]$. The question of finding the eigenvalues, eigenfunctions, proving their completeness and finding their properties, is the subject of study in Sturm-Liouville theory.

4.5.2. *Review.* Let A be a selfadjoint operator.

1) If λ is an eigenvalue of A, then $\lambda \in \mathbb{R}$.

Indeed, $Av = \lambda v$ for some $v \neq 0$. Then on one hand we have

$$\langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2$$

and on the other hand

$$\langle v, Av\rangle = \langle Av, v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\|v\|^2$$

therefore $\lambda\|v\|^2 = \overline{\lambda}\|v\|^2$ so $(\lambda - \overline{\lambda})\|v\|^2 = 0$ therefore $\lambda = \overline{\lambda}$ so $\lambda \in \mathbb{R}$.

2) If $\lambda \neq \mu$ are to eigenvalues of A, $Av = \lambda v$ ($v \neq 0$), and $Au = \mu u$ ($u \neq 0$) then $u \perp v$.

Indeed, on one hand

$$\langle u, Av\rangle = \langle u, \lambda v\rangle = \lambda\langle u, v\rangle$$

and on the other hand

$$\langle u, Av\rangle = \langle Au, v\rangle = \langle \mu u, v\rangle = \mu\langle u, v\rangle$$

therefore $\lambda\langle u, v\rangle = \mu\langle u, v\rangle$ so $(\lambda - \mu)\langle u, v\rangle = 0$ therefore $\langle u, v\rangle = 0$.

4.5.3. *Solution of a boundary value problem using an integral operator.* Let us consider again equation (46) on the domain (47) and rewrite it as $(L - \lambda)y = 0$. We have non-zero solutions y (eigenfunctions of L) only when $Ker(L - \lambda) \neq \{0\}$.

Let us consider the totally opposite case, and look at complex numbers z for which $L - z$ is invertible. (Recall that, in general, the condition that the kernel be zero does not guarantee invertibility of an operator in infinite dimensions.)

Let us invert $L - z$; this means that for any f we solve $(L - z)y = f$ giving $y = (L - z)^{-1}f$. The operator $(L - z)^{-1}$ is called *the resolvent* of L.

To find y we solve the differential equation

$$y'' + zy = -f \tag{48}$$

with the boundary conditions

$$y(0) = 0, \ y(\pi) = 0 \tag{49}$$

Recall that the general solution of (48) is given by

$$C_1 y_1 + C_2 y_2 - y_1 \int \frac{y_2}{W} f + y_2 \int \frac{y_1}{W} f$$

where y_1, y_2 are two linearly independent solutions of the homogeneous equation $y'' + zy = 0$ and $W = W[y_1, y_2] = y_1' y_2 - y_1 y_2'$ is their Wronskian.

We need to distinguish the cases $z = 0$ and $z \neq 0$.

I. If $z = 0$, then $y_1 = 1, y_2 = x$ and we can easily solve the problem (48), (49).

II. If $z \neq 0$, then $y_1 = \exp(kx)$ and $y_2 = \exp(-kx)$ where $k = \sqrt{-z}$ (note that k may be a complex number; in particular, if $z > 0$ then $k = i\sqrt{z}$). Their Wronskian is $W = 2k$, so the general solution of (48) has the form

$$C_1 e^{kx} + C_2 e^{-kx} - \frac{e^{kx}}{2k} \int_0^x e^{-ks} f(s)\, ds + \frac{e^{-kx}}{2k} \int_0^x e^{ks} f(s)\, ds$$

The boundary condition $y(0) = 0$ implies $C_2 = -C_1$ therefore

$$y(x) = C\left(e^{kx} - e^{-kx}\right) - \frac{e^{kx}}{2k} \int_0^x e^{-ks} f(s)\, ds + \frac{e^{-kx}}{2k} \int_0^x e^{ks} f(s)\, ds$$

$$\equiv C\left(e^{kx} - e^{-kx}\right) + \int_0^\pi g(x,s)\, f(s)\, ds \tag{50}$$

where

$$g(x,s) = \frac{-1}{2k}\left(e^{kx-ks} - e^{-kx+ks}\right) \chi_{[0,x]}(s)$$

Imposing the boundary condition $y(\pi) = 0$ we obtain that C must satisfy

$$C\left(e^{k\pi} - e^{-k\pi}\right) + \int_0^\pi g(\pi,s)\, f(s)\, ds = 0$$

Solving for C and substituting in (50) we obtain that the solution of the problem (48), (49) has the form (34)

$$y(x) = \int_0^\pi G(x,s)\, f(s)\, ds \ = \ (L-z)^{-1} f$$

where $G(x,s)$ is the Green function of the problem

$$G(x,s) = g(x,s) \ - \ \frac{e^{kx} - e^{-kx}}{e^{k\pi} - e^{-k\pi}}\, g(\pi,s) \tag{51}$$

Note that G is not defined if $e^{k\pi} - e^{-k\pi} = 0$ which means for $ik \in \mathbb{Z}$. Since $k = \sqrt{-z}$ ($z \neq 0$), this means that for $z = n^2$ ($n = 1, 2, \ldots$) the Green function is undefined (for these values the denominator of G vanishes): the resolvent $(L-z)^{-1}$ does not exist for $z = \lambda_n = n^2$.

Note also that $G(x,s)$ is continuous (the discontinuity of $\chi_{[0,x]}(s)$ at $s = x$ does not result in a discontinuity of $G(x,s)$ at $s = x$ because $g(x,x) = 0$).

It is clear that if z is real then $(L-z)^{-1}$ is self-adjoint because L is self-adjoint.

The study of operators on Hilbert spaces is the topic of Functional Analysis. In one of its chapters it is proved that integral operators (34) (and other similar operators, called *compact operators*) which are selfadjoint are very much like selfadjoint matrices, in that they have real eigenvalues μ_n and the corresponding eigenfunctions u_n form an orthonormal basis for the Hilbert space. The infinite dimensionality of the Hilbert space implies that there are infinitely many eigenvalues: a countable set, which, moreover, tend to zero: $\mu_n \to 0$. (Zero may also be an eigenvalue.)

Let us see how the eigenvalues μ_n and eigenfunctions of the resolvent $(L-z)^{-1}$ are related to the eigenvalues λ_n and eigenfunctions y_n of L.

We have $Ly_n = \lambda_n y_n$ hence $(L-z)y_n = (\lambda_n - z)y_n$ for any number z. If $L-z$ is invertible (we saw that this is the case for $z \neq \lambda_k$ for all k) then $y_n = (\lambda_n - z)(L-z)^{-1}y_n$ so y_n is an eigenfunctions of the resolvent $(L-z)^{-1}$ corresponding to the eigenvalue $\mu_n = (\lambda_n - z)^{-1}$.

Note that $\lambda_n \to \infty$ (since $\mu_n \to 0$).

Note the following quite general facts:

∘ Remark the functional calculus aspect: if λ_n are the eigenvalues of L then $(\lambda_n - z)^{-1}$ are the eigenvalues of $(L-z)^{-1}$ and they correspond to the same eigenvectors.

∘ Note again that the eigenvalues of L appear as values of z for which the Green function has zero denominators.