1.1.1. *2π-periodic functions.* In this part of the course we deal with functions (as above) that are periodic.

We say a function $f : \mathbb{R} \to \mathbb{C}$ is *periodic* with *period* $T > 0$ if $f(x+T) = f(x)$ for all $x \in \mathbb{R}$. For example, $\sin x$, $\cos x$, $e^{ix}(= \cos x + i \sin x)$ are periodic with period 2π. For $k \in R \setminus \{0\}$, $\sin kx$, $\cos kx$, and e^{ikx} are periodic with period $2\pi/|k|$. Constant functions are periodic with period T, for any $T > 0$. We shall specialize to periodic functions with period 2π: we call them 2π-periodic functions, for short. Note that $\cos nx$, $\sin nx$ and e^{inx} are 2π-periodic for $n \in \mathbb{Z}$. (Of course these are also $2\pi/|n|$-periodic.)

Any half-open interval of length T is a *fundamental domain* of a periodic function f of period T. Once you know the values of f on the fundamental domain, you know them everywhere, because any point x in $\mathbb{R}$ can be written uniquely as $x = w + nT$ where $n \in \mathbb{Z}$ and w is in the fundamental domain. Thus $f(x) = f(w+(n-1)T+T) = \cdots = f(w+T) = f(w)$.

For 2π-periodic functions, we shall usually take the fundamental domain to be $]-\pi,\pi]$. By abuse of language, we shall sometimes refer to $[-\pi,\pi]$ as the fundamental domain. We then have to be aware that $f(\pi) = f(-\pi)$.

1.1.2. *Integrating the complex exponential function.* We shall need to calculate $\int_a^b e^{ikx}\,dx$, for $k \in \mathbb{R}$. Note first that when $k = 0$, the integrand is the constant function 1, so the result is $b - a$. For non-zero k, $\int_a^b e^{ikx}\,dx = \int_a^b (\cos kx + i \sin kx)\,dx = (1/k)[(\sin kx - i\cos kx)]_a^b = (1/ik)[(\cos kx + i \sin kx)]_a^b = (1/ik)[e^{ikx}]_a^b = (1/ik)(e^{ikb} - e^{ika})$. Note that this is exactly the result you would have got by treating i as a real constant and using the usual formula for integrating e^{ax}. Note also that the cases $k = 0$ and $k \neq 0$ have to be treated separately: this is typical.

Definition 1.1. Let $f : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic function which is Riemann integrable on $[-\pi,\pi]$. For each $n \in \mathbb{Z}$ we define the *Fourier coefficient* $\hat{f}(n)$ by

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

Remark 1.2. (i) $\hat{f}(n)$ is a complex number whose modulus is the amplitude and whose argument is the phase (of that component of the original function).

(ii) If f and g are Riemann integrable on an interval, then so is their product, so the integral is well-defined.

(iii) The constant before the integral is to divide by the length of the interval.

(iv) We could replace the range of integration by any interval of length 2π, without altering the result, since the integrand is 2π-periodic.

(v) Note the minus sign in the exponent of the exponential. The reason for this will soon become clear.

Example 1.3. (i) $f(x) = c$ then $\hat{f}(0) = c$ and $\hat{f}(n) = 0$ when $n \neq 0$.

(ii) $f(x) = e^{ikx}$, where k is an integer. $\hat{f}(n) = \delta_{nk}$.

(iii) f is 2π periodic and $f(x) = x$ on $]-\pi, \pi]$. (Diagram) Then $\hat{f}(0) = 0$ and, for $n \neq 0$,

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx}\, dx = \left[\frac{-xe^{-inx}}{2\pi i n}\right]_{-\pi}^{\pi} + \frac{1}{in}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\, dx = \frac{(-1)^n i}{n}.$$

Proposition 1.4 (Linearity)**.** *If f and g are 2π-periodic functions and c and d are complex constants, then, for all $n \in \mathbb{Z}$,*

$$(cf + dg)\widehat{\ }(n) = c\hat{f}(n) + d\hat{g}(n)\,.$$

Corollary 1.5. *If $p(x) = \sum_{-k}^{k} c_n e^{inx}$, then $\hat{p}(n) = c_n$ for $|n| \leqslant k$ and $= 0$, for $|n| \geqslant k$.*

$$p(x) = \sum_{n\in\mathbb{Z}} \hat{p}(n) e^{inx}\,.$$

This follows immediately from Ex. 1.3(ii) and Prop.1.4.

Remark 1.6. (i) This corollary explains why the minus sign is natural in the definition of the Fourier coefficients.

(ii) The first part of the course will be devoted to the question of how far this result can be extended to other 2π-periodic functions, that is, for which functions, and for which interpretations of infinite sums is it true that

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n) e^{inx}\,. \tag{1.1}$$

Definition 1.7. $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ is called the *Fourier series* of the 2π-periodic function f.

For real-valued functions, the introduction of complex exponentials seems artificial: indeed they can be avoided as follows. We work with (1.1) in the case of a finite sum: then we can rearrange the sum as

$$\begin{aligned} & \hat{f}(0) + \sum_{n>0} (\hat{f}(n)e^{inx} + \hat{f}(-n)e^{-inx}) \\ = \ & \hat{f}(0) + \sum_{n>0} [(\hat{f}(n) + \hat{f}(-n)) \cos nx + i(\hat{f}(n) - \hat{f}(-n)) \sin nx] \\ = \ & \frac{a_0}{2} + \sum_{n>0} (a_n \cos nx + b_n \sin nx) \end{aligned}$$

Here

$$\begin{aligned} a_n &= (\hat{f}(n) + \hat{f}(-n)) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(e^{-inx} + e^{inx})\,dx \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx \end{aligned}$$

for $n > 0$ and

$$b_n = i((\hat{f}(n) - \hat{f}(-n)) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx$$

for $n > 0$. $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx$, the constant chosen for consistency.

The a_n and b_n are also called Fourier coefficients: if it is necessary to distinguish them, we may call them *Fourier cosine* and *sine coefficients*, respectively.

We note that if f is real-valued, then the a_n and b_n are real numbers and so $\Re\hat{f}(n) = \Re\hat{f}(-n)$, $\Im\hat{f}(n) = -\Im\hat{f}(n)$: thus $\hat{f}(-n)$ is the complex conjugate of $\hat{f}(n)$. Further, if f is an even function then all the sine coefficients are 0 and if f is an odd function, all the cosine coefficients are zero. We note further that the sine and cosine coefficients of the functions $\cos kx$ and $\sin kx$ themselves have a particularly simple form: $a_k = 1$ in the first case and $b_k = 1$ in the second. All the rest are zero.

For example, we should expect the 2π-periodic function whose value on $]-\pi,\pi]$ is x to have just sine coefficients: indeed this is the case: $a_n = 0$ and $b_n = i(\hat{f}(n) - \hat{f}(-n)) = (-1)^{n+1}2/n$ for $n > 0$.

The above question can then be reformulated as "to what extent is $f(x)$ represented by the Fourier series $a_0/2 + \sum_{n>0}(a_n\cos x + b_n \sin x)$?" For instance how well does $\sum(-1)^{n+1}(2/n)\sin nx$ represent the 2π-periodic sawtooth function f whose value on $]-\pi,\pi]$ is given by $f(x) = x$. The easy points are $x = 0$, $x = \pi$, where the terms are identically zero. This gives the 'wrong' value for $x = \pi$, but, if we look at the periodic function near π, we see that it jumps from π to $-\pi$, so perhaps the mean of those values isn't a bad value for the series to converge to. We could conclude that we had defined the function incorrectly to begin with and that its value at the points $(2n+1)\pi$ should have been zero anyway. In fact one can show (ref.) that the Fourier series converges at all other points to the given values of f, but I shan't include the proof in this course. The convergence is not at all uniform (it can't be, because the partial sums are continuous functions, but the limit is discontinuous.) In particular we get the expansion

$$\frac{\pi}{2} = 2(1 - 1/3 + 1/5 - \cdots)$$

which can also be deduced from the Taylor series for $\tan^{-1}$.

1.2. **The vibrating string.** In this subsection we shall discuss the formal solutions of the wave equation in a special case which Fourier dealt with in his work.

We discuss the wave equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{K^2}\frac{\partial^2 y}{\partial t^2}, \tag{1.2}$$

subject to the boundary conditions

$$y(0,t) = y(\pi,t) = 0, \tag{1.3}$$

for all $t \geqslant 0$, and the initial conditions

$$\begin{aligned} y(x,0) &= F(x), \\ y_t(x,0) &= 0. \end{aligned}$$

This is a mathematical model of a string on a musical instrument (guitar, harp, violin) which is of length π and is plucked, i.e. held in the shape $F(x)$ and released at time $t=0$. The constant K depends on the length, density and tension of the string. We shall derive the formal solution (that is, a solution which assumes existence and ignores questions of convergence or of domain of definition).

1.2.1. *Separation of variables.* We first look (as Fourier and others before him did) for solutions of the form $y(x,t) = f(x)g(t)$. Feeding this into the wave equation (1.2) we get

$$f''(x)g(t) = \frac{1}{K^2} f(x)g''(t)$$

and so, dividing by $f(x)g(t)$, we have

$$\frac{f''(x)}{f(x)} = \frac{1}{K^2}\frac{g''(t)}{g(t)}. \tag{1.4}$$

The left-hand side is an expression in x alone, the right-hand side in t alone. The conclusion must be that they are both identically equal to the same constant C, say.

We have $f''(x) - Cf(x) = 0$ subject to the condition $f(0) = f(\pi) = 0$. Working through the method of solving linear second order differential equations tells you that the only solutions occur when $C = -n^2$ for some positive integer n and the corresponding solutions, up to constant multiples, are $f(x) = \sin nx$.

Returning to equation (1.4) gives the equation $g''(t)+K^2n^2g(t) = 0$ which has the general solution $g(t) = a_n \cos Knt + b_n \sin Knt$. Thus the solution we get through separation of variables, using the boundary conditions but ignoring the initial conditions, are

$$y_n(x,t) = \sin nx(a_n \cos Knt + b_n \sin Knt)\,,$$

for $n \geqslant 1$.

1.2.2. *Principle of Superposition.* To get the general solution we just add together all the solutions we have got so far, thus

$$y(x,t) = \sum_{n=1}^{\infty} \sin nx(a_n \cos Knt + b_n \sin Knt) \tag{1.5}$$

ignoring questions of convergence. (We can do this for a finite sum without difficulty because we are dealing with a linear differential equation: the iffy bit is to extend to an infinite sum.)

We now apply the initial condition $y(x,0) = F(x)$ (note F has $F(0) = F(\pi) = 0$). This gives

$$F(x) = \sum_{n=1}^{\infty} a_n \sin nx\,.$$

We apply the reflection trick: the right-hand side is a series of odd functions so if we extend F to a function G by reflection in the origin, giving

$$G(x) := \begin{cases} F(x) & ,\ \text{if } 0 \leqslant x \leqslant \pi; \\ -F(-x) & ,\ \text{if } -\pi < x < 0. \end{cases}$$

we have

$$G(x) = \sum_{n=1}^{\infty} a_n \sin nx\,,$$

for $-\pi \leqslant x \leqslant \pi$.

If we multiply through by $\sin rx$ and integrate term by term, we get

$$a_r = \frac{1}{\pi} \int_{-\pi}^{\pi} G(x) \sin rx\, dx$$

so, assuming that this operation is valid, we find that the a_n are precisely the sine coefficients of G. (Those of you who took Real Analysis 2 last year may remember that a sufficient condition for integrating term-by -term is that the series which is integrated is itself uniformly convergent.)

If we now assume, further, that the right-hand side of (1.5) is differentiable (term by term) we differentiate with respect to t, and set $t = 0$, to get

$$0 = y_t(x,0) = \sum_{n=1}^{\infty} b_n Kn \sin nx. \tag{1.6}$$

This equation is solved by the choice $b_n = 0$ for all n, so we have the following result

Proposition 1.8 (Formal). *Assuming that the formal manipulations are valid, a solution of the differential equation* (1.2) *with the given boundary and initial conditions is*

$$y(x,t) = \sum_{1}^{\infty} a_n \sin nx \cos Knt\,, \tag{2.11}$$

where the coefficients a_n *are the Fourier sine coefficients*

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} G(x) \sin nx \, dx$$

of the 2π *periodic function* G*, defined on* $]-\pi, \pi]$ *by reflecting the graph of* F *in the origin.*

Remark 1.9. This leaves us with the questions

(i) For which F are the manipulations valid?
(ii) Is this the only solution of the differential equation? (which I'm not going to try to answer.)
(iii) Is $b_n = 0$ all n the only solution of (1.6)? This is a special case of the **uniqueness** problem for trigonometric series.

1.3. **Historic: Joseph Fourier.** *Joseph Fourier*, Civil Servant, Egyptologist, and mathematician, was born in 1768 in Auxerre, France, son of a tailor. Debarred by birth from a career in the artillery, he was preparing to become a Benedictine monk (in order to be a teacher) when the French Revolution violently altered the course of history and Fourier's life. He became president of the local revolutionary committee, was arrested during the Terror, but released at the fall of Robespierre.

Fourier then became a pupil at the Ecole Normale (the teachers' academy) in Paris, studying under such great French mathematicians as Laplace and Lagrange. He became a teacher at the Ecole Polytechnique (the military academy).

He was ordered to serve as a scientist under Napoleon in Egypt. In 1801, Fourier returned to France to become Prefect of the Grenoble region. Among his most notable achievements in that office were the draining of some 20 thousand acres of swamps and the building of a new road across the alps.

During that time he wrote an important survey of Egyptian history ("a masterpiece and a turning point in the subject").

In 1804 Fourier started the study of the theory of heat conduction, in the course of which he systematically used the sine-and-cosine series which are named after him. At the end of 1807, he submitted a memoir on this work to the Academy of Science. The memoir proved controversial both in terms of his use of Fourier series and of his derivation of the heat equation and was not accepted at that stage. He was able to resubmit a revised version in 1811: this had several important new features, including the introduction of the Fourier transform. With this version of his memoir, he won the Academy's prize in mathematics. In 1817, Fourier was finally elected to the Academy of Sciences and in 1822 his 1811 memoir was published as "Théorie de la Chaleur".

For more details see *Fourier Analysis* by T.W. Körner, 475-480 and for even more, see the biography by J. Herivel *Joseph Fourier: the man and the physicist*.

What is Fourier analysis. The idea is to analyse functions (into sine and cosines or, equivalently, complex exponentials) to find the underlying frequencies, their strengths (and phases) and, where possible, to see if they can be recombined (synthesis) into the original function. The answers will depend on the original properties of the functions, which often come from physics (heat, electronic or sound waves). This course will give basically a mathematical treatment and so will be interested in mathematical classes of functions (continuity, differentiability properties).

2. Basics of Linear Spaces

> A person is solely the concentration of an infinite set of interrelations with another and others, and to separate a person from these relations means to take away any real meaning of the life.
>
> Vl. Soloviev

A space around us could be described as a three dimensional Euclidean space. To single out a point of that space we need a fixed *frame of references* and three real numbers, which are *coordinates* of the point. Similarly to describe a pair of points from our space we could use six coordinates; for three points—nine, end so on. This makes it reasonable to consider Euclidean (linear) spaces of an arbitrary finite dimension, which are studied in the courses of linear algebra.

The basic properties of Euclidean spaces are determined by its *linear* and *metric* structures. The *linear space* (or *vector space*) structure allows to add and subtract vectors associated to points as well as to multiply vectors by real or complex numbers (scalars).

The *metric space* structure assign a *distance*—non-negative real number—to a pair of points or, equivalently, defines a *length of a vector* defined by that pair. A metric (or, more generally a topology) is essential for definition of the core analytical notions like limit or continuity. The importance of linear and metric (topological) structure in analysis sometime encoded in the formula:

$$\textbf{Analysis} = \textbf{Algebra} + \textbf{Geometry}\,. \tag{2.1}$$

On the other hand we could observe that many sets admit a sort of linear *and* metric structures which are linked each other. Just few among many other examples are:

- The set of convergent sequences;
- The set of continuous functions on $[0,1]$.

It is a very *mathematical way of thinking* to declare such sets to be *spaces* and call their elements *points*.

But shall we lose all information on a particular element (e.g. a sequence $\{1/n\}$) if we represent it by a shapeless and size-less "point" without any inner configuration? Surprisingly not: all properties of an element could be now retrieved not from its *inner configuration* but from interactions with other elements through linear and metric structures. Such a "sociological" approach to all kind of mathematical objects was codified in the abstract *category theory*.

Another surprise is that starting from our three dimensional Euclidean space and walking far away by a road of abstraction to infinite dimensional Hilbert spaces we are arriving just to yet another picture of the surrounding space—that time on the language of *quantum mechanics*.

> The distance from Manchester to Liverpool is 35 miles—just about the mileage in the opposite direction!
>
> *A tourist guide to England*

2.1. **Banach spaces (basic definitions only).** The following definition generalises the notion of *distance* known from the everyday life.

Definition 2.1. A *metric* (or *distance function*) d on a set M is a function $d : M \times M \to \mathbb{R}_+$ from the set of pairs to non-negative real numbers such that:

(i) $d(x, y) \geqslant 0$ for all $x, y \in M$, $d(x, y) = 0$ implies $x = y$.
(ii) $d(x, y) = d(y, x)$ for all x and y in M.
(iii) $d(x, y) + d(y, z) \geqslant d(x, z)$ for all x, y, and z in M (*triangle inequality*).

Exercise 2.2. Let M be the set of UK's cities are the following function are metrics on M:

(i) $d(A, B)$ is the price of 2nd class railway ticket from A to B.
(ii) $d(A, B)$ is the off-peak driving time from A to B.

The following notion is a useful specialisation of metric adopted to the linear structure.

Definition 2.3. Let V be a (real or complex) vector space. A *norm* on V is a real-valued function, written $\|x\|$, such that

(i) $\|x\| \geqslant 0$ for all $x \in V$, and $\|x\| = 0$ implies $x = 0$.
(ii) $\|\lambda x\| = |\lambda| \, \|x\|$ for all scalar λ and vector x.
(iii) $\|x + y\| \leqslant \|x\| + \|y\|$ (*triangle inequality*).

A vector space with a norm is called a *normed space*.

The connection between norm and metric is as follows:

Proposition 2.4. *If* $\|\cdot\|$ *is a norm on* V*, then it gives a metric on* V *by* $d(x, y) = \|x - y\|$.

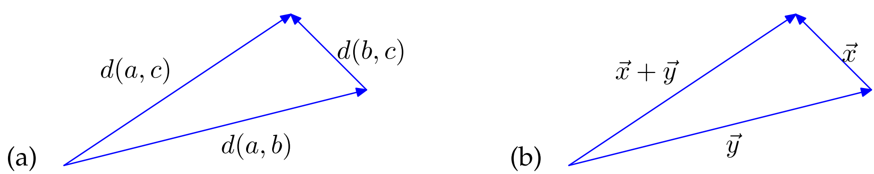

FIGURE 1. Triangle inequality in metric (a) and normed (b) spaces.

Proof. This is a simple exercise to derive items 2.1(i)–2.1(iii) of Definition 2.1 from corresponding items of Definition 2.3. For example, see the Figure 1 to derive the triangle inequality. □

An important notions known from real analysis are limit and convergence. Particularly we usually wish to have enough limiting points for all "reasonable" sequences.

Definition 2.5. A sequence $\{x_k\}$ in a metric space (M, d) is a *Cauchy sequence*, if for every $\epsilon > 0$, there exists an integer n such that $k, l > n$ implies that $d(x_k, x_l) < \epsilon$.

Definition 2.6. (M, d) is a *complete metric space* if every Cauchy sequence in M converges to a limit in M.

For example, the set of integers $\mathbb{Z}$ and reals $\mathbb{R}$ with the natural distance functions are complete spaces, but the set of rationals $\mathbb{Q}$ is not. The complete normed spaces deserve a special name.

Definition 2.7. A *Banach space* is a complete normed space.

Exercise* 2.8. A convenient way to define a norm in a Banach space is as follows. The *unit ball* Uin a normed space B is the set of x such that $\|x\| \leqslant 1$. Prove that:

(i) U is a *convex set*, i.e. $x, y \in U$ and $\lambda \in [0, 1]$ the point $\lambda x + (1 - \lambda)y$ is also in U.
(ii) $\|x\| = \inf\{\lambda \in \mathbb{R}_+ \mid \lambda^{-1}x \in U\}$.
(iii) U is closed if and only if the space is Banach.

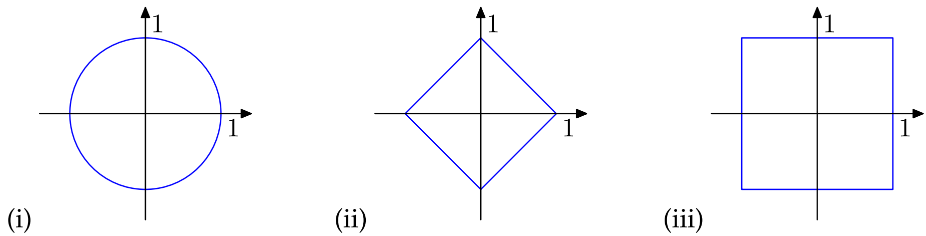

FIGURE 2. Different unit balls defining norms in $\mathbb{R}^2$ from Example 2.9.

Example 2.9. Here is some examples of normed spaces.

(i) ℓ_2^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by

$$\|(x_1, \dots, x_n)\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

(ii) ℓ_1^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by

$$\|(x_1, \dots, x_n)\|_1 = |x_1| + |x_2| + \dots + |x_n|.$$

(iii) ℓ_∞^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by
$$\|(x_1,\ldots,x_n)\|_\infty = \max(|x_1|,|x_2|,\cdots,|x_n|).$$
(iv) Let X be a topological space, then $C_b(X)$ is the space of continuous bounded functions $f : X \to \mathbb{C}$ with norm $\|f\|_\infty = \sup_X |f(x)|$.
(v) Let X be any set, then $\ell_\infty(X)$ is the space of all bounded (not necessarily continuous) functions $f : X \to \mathbb{C}$ with norm $\|f\|_\infty = \sup_X |f(x)|$.

All these normed spaces are also complete and thus are Banach spaces. Some more examples of both complete and incomplete spaces shall appear later.

—We need an extra space to accommodate this product!
A manager to a shop assistant

2.2. **Hilbert spaces.** Although metric and norm capture important geometric information about linear spaces they are not sensitive enough to represent such geometric characterisation as angles (particularly *orthogonality*). To this end we need a further refinements.

From courses of linear algebra known that the scalar product $\langle x,y\rangle = x_1y_1 + \cdots + x_ny_n$ is important in a space $\mathbb{R}^n$ and defines a norm $\|x\|^2 = \langle x,x\rangle$. Here is a suitable generalisation:

Definition 2.10. A *scalar product* (or *inner product*) on a real or complex vector space V is a mapping $V \times V \to \mathbb{C}$, written $\langle x,y\rangle$, that satisfies:
(i) $\langle x,x\rangle \geqslant 0$ and $\langle x,x\rangle = 0$ implies $x = 0$.
(ii) $\langle x,y\rangle = \overline{\langle y,x\rangle}$ in complex spaces and $\langle x,y\rangle = \langle y,x\rangle$ in real ones for all x, $y \in V$.
(iii) $\langle \lambda x,y\rangle = \lambda\langle x,y\rangle$, for all $x, y \in V$ and scalar λ. (What is $\langle x,\lambda y\rangle$?).
(iv) $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$, for all x, y, and $z \in V$. (What is $\langle x,y+z\rangle$?).

Last two properties of the scalar product is oftenly encoded in the phrase: "it is linear in the first variable if we fix the second and anti-linear in the second if we fix the first".

Definition 2.11. An *inner product space* V is a real or complex vector space with a scalar product on it.

Example 2.12. Here is some examples of inner product spaces which demonstrate that expression $\|x\| = \sqrt{\langle x,x\rangle}$ defines a norm.
(i) The inner product for $\mathbb{R}^n$ was defined in the beginning of this section. The inner product for $\mathbb{C}^n$ is given by $\langle x,y\rangle = \sum_1^n x_j\bar{y}_j$. The norm $\|x\| = \sqrt{\sum_1^n |x_j|^2}$ makes it ℓ_2^n from Example 2.9(i).
(ii) The extension for infinite vectors: let ℓ_2 be

$$\ell_2 = \{\text{sequences } \{x_j\}_1^\infty \mid \sum_1^\infty |x_j|^2 < \infty\}. \tag{2.2}$$

Let us equip this set with operations of term-wise addition and multiplication by scalars, then ℓ_2 is closed under them. Indeed it follows from the triangle inequality and properties of absolutely convergent series. From the standard Cauchy–Bunyakovskii–Schwarz inequality follows that the series $\sum_1^\infty x_j \bar{y}_j$ absolutely converges and its sum defined to be $\langle x, y\rangle$.

(iii) Let $C_b[a,b]$ be a space of continuous functions on the interval $[a,b] \in \mathbb{R}$. As we learn from Example 2.9(iv) a normed space it is a normed space with the norm $\|f\|_\infty = \sup_{[a,b]} |f(x)|$. We could also define an inner product:

$$\langle f, g\rangle = \int_a^b f(x)\bar{g}(x)\, dx \quad \text{and} \quad \|f\|_2 = \left(\int_a^b |f(x)|^2\, dx\right)^{\frac{1}{2}}. \tag{2.3}$$

Now we state, probably, the most important inequality in analysis.

Theorem 2.13 (Cauchy–Schwarz–Bunyakovskii inequality)**.** *For vectors x and y in an inner product space V let us define $\|x\| = \sqrt{\langle x, x\rangle}$ and $\|y\| = \sqrt{\langle y, y\rangle}$ then we have*

$$|\langle x, y\rangle| \leqslant \|x\|\,\|y\|\,, \tag{2.4}$$

with equality if and only if x and y are scalar multiple each other.

Proof. For any $x, y \in V$ and any $t \in \mathbb{R}$ we have:

$$0 < \langle x + ty, x + ty\rangle = \langle x, x\rangle + 2t\Re\langle y, x\rangle + t^2\langle y, y\rangle),$$

Thus the discriminant of this quadratic expression in t is non-positive: $(\Re\langle y, x\rangle)^2 - \|x\|^2\|y\|^2 \leqslant 0$, that is $|\Re\langle x, y\rangle| \leqslant \|x\|\,\|y\|$. Replacing y by $e^{i\alpha}y$ for an arbitrary $\alpha \in [-\pi, \pi]$ we get $\left|\Re(e^{i\alpha}\langle x, y\rangle)\right| \leqslant \|x\|\,\|y\|$, this implies the desired inequality. □

Corollary 2.14. *Any inner product space is a normed space with norm $\|x\| = \sqrt{\langle x, x\rangle}$ (hence also a metric space, Prop. 2.4).*

Proof. Just to check items 2.3(i)–2.3(iii) from Definition 2.3. □

Again complete inner product spaces deserve a special name

Definition 2.15. A complete inner product space is *Hilbert space*.

The relations between spaces introduced so far are as follows:

$$\begin{array}{ccccc} \text{Hilbert spaces} & \Rightarrow & \text{Banach spaces} & \Rightarrow & \text{Complete metric spaces} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \text{inner product spaces} & \Rightarrow & \text{normed spaces} & \Rightarrow & \text{metric spaces.} \end{array}$$

How can we tell if a given norm comes from an inner product?

Theorem 2.16 (Parallelogram identity)**.** *In an inner product space H we have for all x and $y \in H$ (see Figure 3):*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2\,. \tag{2.5}$$

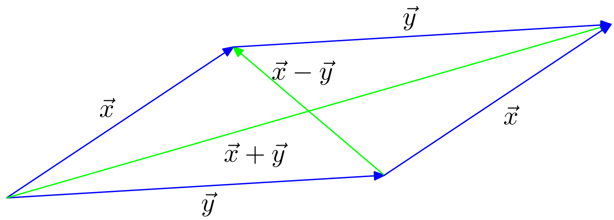

FIGURE 3. To the parallelogram identity.

Proof. Just by linearity of inner product:

$$\langle x+y, x+y\rangle + \langle x-y, x-y\rangle = 2\langle x,x\rangle + 2\langle y,y\rangle,$$

because the cross terms cancel out. □

Exercise 2.17. Show that (2.5) is also a sufficient condition for a norm to arise from an inner product. Namely, for a norm on a complex Banach space satisfying to (2.5) the formula

$$\begin{aligned} \langle x,y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) \\ &= \frac{1}{4}\sum_0^3 i^k \left\|x+i^k y\right\|^2 \end{aligned} \tag{2.6}$$

defines an inner product. What is a suitable formula for a real Banach space?

Divide and rule!

Old but still much used recipe

2.3. **Subspaces.** To study Hilbert spaces we may use the traditional mathematical technique of *analysis* and *synthesis*: we split the initial Hilbert spaces into smaller and probably simpler subsets, investigate them separately, and then reconstruct the entire picture from these parts.

As known from the linear algebra, a *linear subspace* is a subset of a linear space is its subset, which inherits the linear structure, i.e. possibility to add vectors and multiply them by scalars. In this course we need also that subspaces inherit topological structure (coming either from a norm or an inner product) as well.

Definition 2.18. By a *subspace* of a normed space (or inner product space) we mean a linear subspace with the same norm (inner product respectively). We write $X \subset Y$ or $X \subseteq Y$.

Example 2.19. (i) $C_b(X) \subset \ell_\infty(X)$ where X is a metric space.

(ii) Any linear subspace of $\mathbb{R}^n$ or $\mathbb{C}^n$ with any norm given in Example 2.9(i)–2.9(iii).

(iii) Let c_{00} be the *space of finite sequences*, i.e. all sequences (x_n) such that exist N with $x_n = 0$ for $n > N$. This is a subspace of ℓ_2 since $\sum_1^\infty |x_j|^2$ is a finite sum, so finite.

We also wish that the both inhered structures (linear and topological) should be in agreement, i.e. the subspace should be complete. Such inheritance is linked to the property be closed.

A subspace need *not* be closed—for example the sequence

$$x = (1, 1/2, 1/3, 1/4, \ldots) \in \ell_2 \qquad \text{because} \qquad \sum 1/k^2 < \infty$$

and $x_n = (1, 1/2, \ldots, 1/n, 0, 0, \ldots) \in c_{00}$ converges to x thus $x \in \overline{c_{00}} \subset \ell_2$.

Proposition 2.20. (i) *Any closed subspace of a Banach/Hilbert space is complete, hence also a Banach/Hilbert space.*
(ii) *Any complete subspace is closed.*
(iii) *The closure of subspace is again a subspace.*

Proof. (i) This is true in any metric space X: any Cauchy sequence from Y has a limit $x \in X$ belonging to $\bar{Y}$, but if Y is closed then $x \in Y$.
(ii) Let Y is complete and $x \in \bar{Y}$, then there is sequence $x_n \to x$ in Y and it is a Cauchy sequence. Then completeness of Y implies $x \in Y$.
(iii) If $x, y \in \bar{Y}$ then there are x_n and y_n in Y such that $x_n \to x$ and $y_n \to y$. From the triangle inequality:

$$\|(x_n + y_n) - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\| \to 0,$$

so $x_n + y_n \to x + y$ and $x + y \in \bar{Y}$. Similarly $x \in \bar{Y}$ implies $\lambda x \in \bar{Y}$ for any λ. □

Hence c_{00} is an *incomplete* inner product space, with inner product $\langle x, y \rangle = \sum_1^\infty x_k \bar{y}_k$ (this is a finite sum!) as it is not closed in ℓ_2.

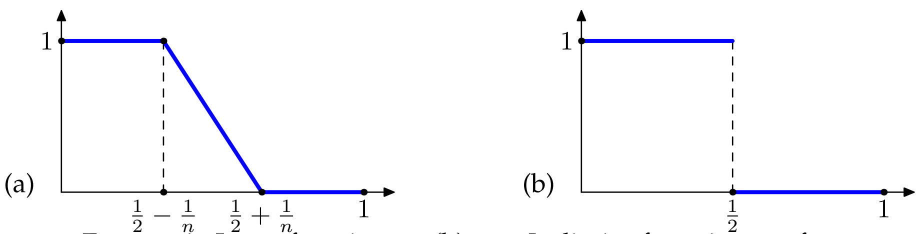

FIGURE 4. Jump function on (b) as a L_2 limit of continuous functions from (a).

Similarly $C[0,1]$ with inner product norm $\|f\| = \left(\int_0^1 |f(t)|^2\, dt\right)^{1/2}$ is incomplete—take the large space X of functions continuous on $[0,1]$ except for a possible jump at

$\frac{1}{2}$ (i.e. left and right limits exists but may be unequal and $f(\frac{1}{2}) = \lim_{t\to\frac{1}{2}+} f(t)$. Then the sequence of functions defined on Figure 4(a) has the limit shown on Figure 4(b) since:

$$\|f - f_n\| = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}+\frac{1}{n}} |f - f_n|^2 \, dt < \frac{2}{n} \to 0.$$

Obviously $f \in \overline{C[0,1]} \setminus C[0,1]$.

Exercise 2.21. Show alternatively that the sequence of function f_n from Figure 4(a) is a Cauchy sequence in $C[0,1]$ but has no continuous limit.

Similarly the space $C[a,b]$ is *incomplete* for any $a < b$ if equipped by the inner product and the corresponding norm:

$$\langle f, g \rangle = \int_a^b f(t)\bar{g}(t)\, dt \tag{2.7}$$

$$\|f\|_2 = \left(\int_a^b |f(t)|^2 \, dt \right)^{1/2}. \tag{2.8}$$

Definition 2.22. Define a Hilbert space $L_2[a,b]$ to be the smallest complete inner product space containing space $C[a,b]$ with the restriction of inner product given by (2.7).

It is practical to realise $L_2[a,b]$ as a certain space of "functions" with the inner product defined via an integral. There are several ways to do that and we mention just two:

(i) Elements of $L_2[a,b]$ are equivalent classes of Cauchy sequences $f^{(n)}$ of functions from $C[a,b]$.

(ii) Let integration be extended from the Riemann definition to the wider *Lebesgue integration* (see Section 13). Let L be a set of square integrable in Lebesgue sense functions on $[a,b]$ with a finite norm (2.8). Then $L_2[a,b]$ is a quotient space of L with respect to the equivalence relation $f \sim g \Leftrightarrow \|f - g\|_2 = 0$.

Example 2.23. Let the *Cantor function* on $[0,1]$ be defined as follows:

$$f(t) = \begin{cases} 1, & t \in \mathbb{Q}; \\ 0, & t \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

This function is *not* integrable in the Riemann sense but *does* have the Lebesgue integral. The later however is equal to 0 and as an L_2-function the Cantor function equivalent to the function identically equal to 0.

(iii) The third possibility is to map $L_2(\mathbb{R})$ onto a space of "true" functions but with an additional structure. For example, in *quantum mechanics* it is useful to work with the *Segal–Bargmann space* of analytic functions on $\mathbb{C}$ with the inner product:

$$\langle f_1, f_2 \rangle = \int_{\mathbb{C}} f_1(z)\bar{f}_2(z)e^{-|z|^2}\, dz.$$

Theorem 2.24. *The sequence space ℓ_2 is complete, hence a Hilbert space.*

Proof. Take a Cauchy sequence $x^{(n)} \in \ell_2$, where $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \ldots)$. Our proof will have three steps: identify the limit x; show it is in ℓ_2; show $x^{(n)} \to x$.

(i) If $x^{(n)}$ is a Cauchy sequence in ℓ_2 then $x_k^{(n)}$ is also a Cauchy sequence of numbers for any fixed k:

$$\left|x_k^{(n)} - x_k^{(m)}\right| \leqslant \left(\sum_{k=1}^{\infty} \left|x_k^{(n)} - x_k^{(m)}\right|^2\right)^{1/2} = \left\|x^{(n)} - x^{(m)}\right\| \to 0.$$

Let x_k be the limit of $x_k^{(n)}$.

(ii) For a given $\epsilon > 0$ find n_0 such that $\left\|x^{(n)} - x^{(m)}\right\| < \epsilon$ for all $n, m > n_0$. For any K and m:

$$\sum_{k=1}^{K} \left|x_k^{(n)} - x_k^{(m)}\right|^2 \leqslant \left\|x^{(n)} - x^{(m)}\right\|^2 < \epsilon^2.$$

Let $m \to \infty$ then $\sum_{k=1}^{K} \left|x_k^{(n)} - x_k\right|^2 \leqslant \epsilon^2$.

Let $K \to \infty$ then $\sum_{k=1}^{\infty} \left|x_k^{(n)} - x_k\right|^2 \leqslant \epsilon^2$. Thus $x^{(n)} - x \in \ell_2$ and because ℓ_2 is a linear space then $x = x^{(n)} - (x^{(n)} - x)$ is also in ℓ_2.

(iii) We saw above that for any $\epsilon > 0$ there is n_0 such that $\left\|x^{(n)} - x\right\| < \epsilon$ for all $n > n_0$. Thus $x^{(n)} \to x$.

Consequently ℓ_2 is complete. □

All good things are covered by a thick layer of chocolate (well, if something is not yet–it certainly will)

2.4. **Linear spans.** As was explained into introduction 2, we describe "internal" properties of a vector through its relations to other vectors. For a detailed description we need sufficiently many external reference points.

Let A be a subset (finite or infinite) of a normed space V. We may wish to upgrade it to a linear subspace in order to make it subject to our theory.

Definition 2.25. The *linear span* of A, write $\mathrm{Lin}(A)$, is the intersection of all linear subspaces of V containing A, i.e. the smallest subspace containing A, equivalently the set of all finite linear combination of elements of A. The *closed linear span* of A write $\mathrm{CLin}(A)$ is the intersection of all *closed* linear subspaces of V containing A, i.e. the smallest *closed* subspace containing A.

Exercise* 2.26. (i) Show that if A is a subset of finite dimension space then $\mathrm{Lin}(A) = \mathrm{CLin}(A)$.
(ii) Show that for an infinite A spaces $\mathrm{Lin}(A)$ and $\mathrm{CLin}(A)$could be different. (*Hint*: use Example 2.19(iii).)

Proposition 2.27. $\overline{\mathrm{Lin}(A)} = \mathrm{CLin}(A)$.

Proof. Clearly $\overline{\mathrm{Lin}(A)}$ is a closed subspace containing A thus it should contain $\mathrm{CLin}(A)$. Also $\mathrm{Lin}(A) \subset \mathrm{CLin}(A)$ thus $\overline{\mathrm{Lin}(A)} \subset \overline{\mathrm{CLin}(A)} = \mathrm{CLin}(A)$. Therefore $\overline{\mathrm{Lin}(A)} = \mathrm{CLin}(A)$. □

Consequently $\mathrm{CLin}(A)$ is the set of all limiting points of finite linear combination of elements of A.

Example 2.28. Let $V = C[a, b]$ with the sup norm $\|\cdot\|_\infty$. Then:
$\mathrm{Lin}\{1, x, x^2, \ldots\} = \{\text{all polynomials}\}$
$\mathrm{CLin}\{1, x, x^2, \ldots\} = C[a, b]$ by the Weierstrass approximation theorem proved later.

The following simple result will be used later many times without comments.

Lemma 2.29 (about Inner Product Limit)**.** *Suppose H is an inner product space and sequences x_n and y_n have limits x and y correspondingly. Then $\langle x_n, y_n \rangle \to \langle x, y \rangle$ or equivalently:*

$$\lim_{n\to\infty} \langle x_n, y_n \rangle = \left\langle \lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n \right\rangle .$$

Proof. Obviously by the Cauchy–Schwarz inequality:

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n - x, y_n \rangle + \langle x, y_n - y \rangle| \\ &\leqslant |\langle x_n - x, y_n \rangle| + |\langle x, y_n - y \rangle| \\ &\leqslant \|x_n - x\| \, \|y_n\| + \|x\| \, \|y_n - y\| \to 0, \end{aligned}$$

since $\|x_n - x\| \to 0$, $\|y_n - y\| \to 0$, and $\|y_n\|$ is bounded. □

3. Orthogonality

Pythagoras is forever!
The catchphrase from TV commercial of Hilbert Spaces course

As was mentioned in the introduction the Hilbert spaces is an analog of our 3D Euclidean space and theory of Hilbert spaces similar to plane or space geometry. One of the primary result of Euclidean geometry which still survives in high school curriculum despite its continuous nasty de-geometrisation is Pythagoras' theorem based on the notion of *orthogonality*[1].

[1]Some more "strange" types of orthogonality can be seen in the paper *Elliptic, Parabolic and Hyperbolic Analytic Function Theory–1: Geometry of Invariants*.

So far we was concerned only with distances between points. Now we would like to study angles between vectors and notably *right angles*. Pythagoras' theorem states that if the angle C in a triangle is right then $c^2 = a^2 + b^2$, see Figure 5 .

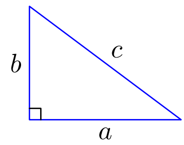

FIGURE 5. The Pythagoras' theorem $c^2 = a^2 + b^2$

It is a very *mathematical way of thinking* to turn this *property* of right angles into their *definition*, which will work even in infinite dimensional Hilbert spaces.

Look for a triangle, or even for a right triangle

A universal advice in solving problems from elementary geometry.

3.1. **Orthogonal System in Hilbert Space.** In inner product spaces it is even more convenient to give a definition of orthogonality not from Pythagoras' theorem but from an equivalent property of inner product.

Definition 3.1. Two vectors x and y in an inner product space are *orthogonal* if $\langle x, y\rangle = 0$, written $x \perp y$.

An *orthogonal sequence* (or *orthogonal system*) e_n (finite or infinite) is one in which $e_n \perp e_m$ whenever $n \neq m$.

An *orthonormal sequence* (or *orthonormal system*) e_n is an orthogonal sequence with $\|e_n\| = 1$ for all n.

Exercise 3.2.
(i) Show that if $x \perp x$ then $x = 0$ and consequently $x \perp y$ for any $y \in H$.
(ii) Show that if all vectors of an orthogonal system are non-zero then they are linearly independent.

Example 3.3. These are orthonormal sequences:
(i) Basis vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ in $\mathbb{R}^3$ or $\mathbb{C}^3$.
(ii) Vectors $e_n = (0, \ldots, 0, 1, 0, \ldots)$ (with the only 1 on the nth place) in ℓ_2. (Could you see a similarity with the previous example?)
(iii) Functions $e_n(t) = 1/(\sqrt{2}\pi)e^{int}$, $n \in \mathbb{Z}$ in $C[0, 2\pi]$:

$$\langle e_n, e_m\rangle = \int_0^{2\pi} \frac{1}{2\pi} e^{int} e^{-imt}\, dt = \begin{cases} 1, & n = m; \\ 0, & n \neq m. \end{cases} \tag{3.1}$$

Exercise 3.4. Let A be a subset of an inner product space V and $x \perp y$ for any $y \in A$. Prove that $x \perp z$ for all $z \in \mathrm{CLin}(A)$.

Theorem 3.5 (Pythagoras'). *If* $x \perp y$ *then* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. *Also if* $e_1, \ldots, e_n$ *is orthonormal then*

$$\left\| \sum_1^n a_k e_k \right\|^2 = \left\langle \sum_1^n a_k e_k, \sum_1^n a_k e_k \right\rangle = \sum_1^n |a_k|^2 .$$

Proof. A one-line calculation. □

The following theorem provides an important property of Hilbert spaces which will be used many times. Recall, that a subset K of a linear space V is *convex* if for all $x, y \in K$ and $\lambda \in [0,1]$ the point $\lambda x + (1-\lambda) y$ is also in K. Particularly any subspace is convex and any unit ball as well (see Exercise 2.8(i)).

Theorem 3.6 (about the Nearest Point). *Let* K *be a non-empty convex closed subset of a Hilbert space* H. *For any point* $x \in H$ *there is the unique point* $y \in K$ *nearest to* x.

Proof. Let $d = \inf_{y\in K} d(x,y)$, where $d(x,y)$—the distance coming from the norm $\|x\| = \sqrt{\langle x, x\rangle}$ and let y_n a sequence points in K such that $\lim_{n\to\infty} d(x, y_n) = d$. Then y_n is a Cauchy sequence. Indeed from the parallelogram identity for the parallelogram generated by vectors $x - y_n$ and $x - y_m$ we have:

$$\|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - y_n - y_m\|^2 .$$

Note that $\|2x - y_n - y_m\|^2 = 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \geqslant 4d^2$ since $\frac{y_n+y_m}{2} \in K$ by its convexity. For sufficiently large m and n we get $\|x - y_m\|^2 \leqslant d + \epsilon$ and $\|x - y_n\|^2 \leqslant d + \epsilon$, thus $\|y_n - y_m\| \leqslant 4(d^2 + \epsilon) - 4d^2 = 4\epsilon$, i.e. y_n is a Cauchy sequence.

Let y be the limit of y_n, which exists by the completeness of H, then $y \in K$ since K is closed. Then $d(x,y) = \lim_{n\to\infty} d(x, y_n) = d$. This show the existence of the nearest point. Let y' be another point in K such that $d(x, y') = d$, then the parallelogram identity implies:

$$\|y - y'\|^2 = 2\|x - y\|^2 + 2\|x - y'\|^2 - \|2x - y - y'\|^2 \leqslant 4d^2 - 4d^2 = 0.$$

This shows the uniqueness of the nearest point. □

Exercise* 3.7. The essential rôle of the parallelogram identity in the above proof indicates that the theorem does not hold in a general Banach space.

(i) Show that in $\mathbb{R}^2$ with either norm $\|\cdot\|_1$ or $\|\cdot\|_\infty$ form Example 2.9 the nearest point could be non-unique;

(ii) Could you construct an example (in Banach space) when the nearest point does not exists?

Liberte, Egalite, Fraternite!

A longstanding ideal approximated in the real life by something completely different