

Alessandro Candolini

---

# Quantum Mechanics

A Non-Relativistic Primer

August 13, 2023

Alessandro Candolini: *Quantum Mechanics* © august 13, 2023.

E-MAIL: [alessandro.candolini@gmail.com](mailto:alessandro.candolini@gmail.com)

“ In the good old days, theorizing was like sailing between islands of experimental evidence. And, if the trip was not in the vicinity of the shoreline (which was strongly recommended for safety reasons) sailors were continuously looking forward, hoping to see land — the sooner the better.

Nowadays, some theoretical physicists (let us call them sailors) [have] found a way to survive and navigate in the open sea of pure theoretical constructions. Instead of the horizon, they look at stars, which tell them exactly where they are. Sailors are aware of the fact that the stars will never tell them where the new land is, but they may tell them their position on the globe.

Theoreticians become sailors simply because they just like it. Young people, seduced by captains forming crews to go to a Nuevo El Dorado [...] soon realize that they will spend all their life at sea. Those who do not like sailing desert the voyage, but for the true potential sailors the sea becomes their passion. They will probably tell the alluring and frightening truth to their students — and the proper people will join their ranks. ”

— Andrei Losev



## CONTENTS

---

Preface	vii
<b>I The basis</b>	
1 Schwinger's own way to teach quantum mechanics	3
1.1 Introduction	3
2 Linear operators in Hilbert spaces	5
2.1 Metrics space	5
2.2 Banach and Hilbert spaces	6
2.3 A glimpse at convex analysis	16
2.4 Canonical prototypes of Banach and Hilbert spaces	18
2.4.1 The Hilbert spaces $\mathbb{R}^N$ and $\mathbb{C}^N$	18
2.4.2 The Hilbert space $l^2$	19
2.4.3 The Hilbert space of square integrable function	19
2.4.4 Hölder and Minkovsky inequalities and p-norms	23
2.5 Orthogonality	24
2.6 Complete orthonormal sets and their characterization	25
2.7 The Fouries series in the space of square integrable functions	25
2.8 Linear operators: first properties	25
2.9 Riesz representation theorem	25
3 The rules of the game	27
4 One-dimensional quantum systems	29
<b>II The core</b>	
5 Angular momentum	33
6 Hydrogen atom	35
7 Perturbation theory	37
8 Scattering	39
<b>III Advanced topics</b>	
9 Path integrals	43
10 Semiclassical quantum mechanics	45
11 Supersymmetric Quantum Mechanics	47
12 Second quantization formalism	49
<b>IV Appendices</b>	
A Appendix A	53
Bibliography	55



## PREFACE

---

London, August 13, 2023

Alessandro Candolini





## Part I

### THE BASIS



## SCHWINGER'S OWN WAY TO TEACH QUANTUM MECHANICS

---

“ I presume that all of you have already been exposed to some undergraduate course in Quantum Mechanics, one that leans heavily on de Broglie waves and the Schroedinger equation. I have never thought that this simple wave approach was acceptable as a general basis for the whole subject, and I intend to move immediately to replace it in your mind by a foundation that *is* perfectly general. ”

J. Schwinger, *Quantum Mechanics. Symbolism of Atomic Measurements*  
Schwinger [2001].

### 1.1 INTRODUCTION

COMPARED with other traditional areas of physics, quantum mechanics is not easy. It often lacks physical intuition and it relies on heavy mathematical background from the very beginning. As we shall see shortly, topics like Heisenberg's uncertainty principle, the intriguing role of probability in predicting the outcomes of a quantum measurement, entanglement and superposition of quantum states, etc asks immediately for a theoretical framework capable of supporting and describing such behaviors. The phase space in Hamiltonian classical dynamics simply not suitable because it is not able to include such phenomena, are built in the formalism (at least in its traditional construction) does not offer the tools. it is not enough flexible. Naturally leading from the very beginning to adopt .

For example, incompatibility between physical measurable quantities like position and momentum, which is at the root of uncertainty principle, can be described in terms of *non-commutative* objects, leading naturally to adopt matrices or operators (instead of usual numbers) to describe position and momentum. As we shall see, the canonical commutation relations imposed even further requirements: they can't be realized in finite-dimensional vector spaces, so we need to look at infinite-dimensional (Hilbert) spaces.

Can we avoid Hilbert space? As we shall see in the last part, other descriptions are possible, in particular the path integral approach to quantum mechanics due to Feynman.

In chapter 3, we will present and thoroughly discuss all the postulates of non-relativistic quantum mechanics. They have been distilled in the Twenties and Thirties of '900 after at least two decades of strong and fascinating efforts, theoretical and experimental progress, wrong attempts, mistakes, etc made by the intellectual fathers of quantum mechanics to eventually capture the essential rules underlying the microscopic behavior of particles.

For example, consider classical electrodynamics. One can ask: are fields "necessary"?

### REFERENCES FOR § 1

SCHWINGER, JULIAN

2001 *Quantum Mechanics. Symbolism of Atomic Measurements*, Springer-Verlag, Berlin, ISBN: 3-540-41408-8. (Cit. on p. 3.)



Figure 1.1: Julian Seymour Schwinger (February 12, 1918 – July 16, 1994)

The operator formulation of standard non-relativistic quantum mechanics is rooted in the spectral theory of linear operators in complex Hilbert spaces. In particular, the spectral decomposition of self-adjoint operators (whether bounded or unbounded) is a key ingredient in formulating some of the basic rules of quantum mechanics. This chapter aims to provide the necessary mathematical groundwork in functional analysis to approach non-relativistic quantum mechanics in full generality. This chapter primarily focuses on the mathematical aspects rather than the physical implications. Readers already acquainted with these concepts may choose to skip this chapter. Familiarity with basic linear algebra (in particular, linear spaces) is assumed.

In formulating quantum mechanics, physicists often rely on Dirac's notation, a potent mnemonic technique that, at first glance, lacks strict mathematical rigor. Nonetheless, it is feasible to create a correspondence between Dirac's notation and the rigorous theorems of functional analysis, as we will delve into in subsequent sections.

This chapter is based on [Reed and Simon \[1980\]](#). Some other books that have been helpful while writing this chapter includes: [Berberian \[1976\]](#), [Debnath and Mikusiński \[2005\]](#), [Helmberg \[1969\]](#), [Hutson and Pym \[1980\]](#), and [Teschl \[2009\]](#).

## 2.1 METRICS SPACE

Metrics spaces are sets endowed with an additional algebraic structure that attaches significance to the concept of "distance" between each pair of elements in the set. Metric spaces provide a natural context in which to develop concepts like continuity, convergence, and more.

**DEFINITION 2.1 (distance):** Let  $A$  be a set. A "distance" (or "metrics") over  $A$  is any application  $d: A \times A \rightarrow [0, +\infty]$ , satisfying the following requirements:

- (a)  $d(\varphi, \psi) \geq 0$ ,  $\forall (\varphi, \psi) \in A \times A$  (nonnegativity);
- (b)  $d(\varphi, \psi) = 0$  if and only if  $\varphi = \psi$  (i.e., the only element in  $A$  at zero distance from  $\psi$  is  $\psi$  itself);
- (c)  $d(\varphi, \psi) = d(\psi, \varphi)$ ,  $\forall (\varphi, \psi) \in A \times A$  (i.e., the distance is a symmetric function of its arguments);
- (d)  $d(\varphi, \psi) \leq d(\psi, \eta) + d(\eta, \varphi)$ ,  $\forall (\varphi, \psi, \eta) \in A \times A \times A$  (the so-called "triangle inequality").

Elements of  $A$  are often referred to as "points" in this context.

The next sections will provide concrete examples of metrics spaces where the distance arises from other algebraic structures (i.e., norms and inner products). Without pretending to extensively develop the theory of metrics spaces in this section, it's worth noting how the familiar notions of continuity, convergence, etc can be entirely formulated in the context of metrics spaces.

**DEFINITION 2.2 (continuity):** Let  $(A, d)$  and  $(A', d')$  be metric spaces and  $f: A \rightarrow A'$  a function.  $f$  is "continuous" at a point  $\varphi \in A$  (with respect

to the given metrics) if for any positive real number  $\varepsilon > 0$ , there exists a positive real number  $\delta > 0$  such that for every  $\psi \in A$ ,  $d(\varphi, \psi) < \delta$  implies  $d'(f(\varphi), f(\psi)) < \varepsilon$ . Formally,

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall \psi \in A)(d(\varphi, \psi) < \delta \Rightarrow (d'(f(\varphi), f(\psi)) < \varepsilon). \quad (2.1)$$

The notion can be further generalised to functions between *topological spaces* in which there is generally no formal notion of distance.

## 2.2 BANACH AND HILBERT SPACES

Unless stated otherwise, let  $\mathbb{K}$  denote either the field of real numbers  $\mathbb{R}$  or the field of complex numbers  $\mathbb{C}$ . (It is possible to develop the theory also for the skew-field of quaternions, but this case will not be taken into account here in order to avoid dealing with the non-commutativity of the quaternionic product.)

**DEFINITION 2.3 (norm):** Let  $X$  be any vector space over  $\mathbb{K}$ . A “norm” on  $X$  is any application  $X \rightarrow \mathbb{R}$  hereafter denoted by  $\|\cdot\|$  satisfying the following properties:

- (a)  $\|\varphi\| \geq 0$ ,  $\forall \varphi \in X$  (nonnegativity);
- (b)  $\|\varphi\| = 0$  if and only if  $\varphi = 0$  (faithfulness)
- (c)  $\|\lambda\varphi\| = |\lambda| \|\varphi\|$ ,  $\forall \lambda \in \mathbb{K}$  and  $\forall \varphi \in X$  (positive homogeneity);
- (d)  $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ ,  $\forall (\varphi, \psi) \in X \times X$  (subadditivity).

A norm on  $X$  can be viewed as a way to define a “length” of the vectors belonging to  $X$ .

*Remark.* Item (a) is redundant: it follows from items (b) to (d). For every  $\varphi \in X$ , item (c) implies  $\|-\varphi\| = \|\varphi\|$ , from item (d) it follows that  $\|\varphi + (-\varphi)\| \leq \|\varphi\| + \|\varphi\|$ , and item (b) implies  $\|\varphi - \varphi\| = 0$ , so  $\|\varphi\| \geq 0$ .

*Remark.*  $X$  is not required to be finite-dimensional. It can be infinite-dimensional. (Actually, the infinite-dimensional case is the case we are most interested on.)

A “normed vector space” is a vector space equipped with a norm, as formalized by the following definition.

**DEFINITION 2.4 (normed vector space):** A “normed vector space” over  $\mathbb{K}$  is a pair  $(X, \|\cdot\|)$  where  $X$  is a vector space over  $\mathbb{K}$  and  $\|\cdot\|$  is any norm on  $X$ .

If  $(X, \|\cdot\|)$  is a normed vector space, the norm  $\|\cdot\|$  *induces* a metric (i. e., a notion of distance) and thus a topology on  $X$ . This is proved by the following theorem.

**THEOREM 2.1:** Let  $(X, \|\cdot\|)$  be any normed vector space over  $\mathbb{K}$ . Let  $d: X \times X \rightarrow \mathbb{R}$  be the function defined by

$$d(\varphi, \psi) = \|\varphi - \psi\|, \quad (2.2)$$

for every pair of vectors  $(\varphi, \psi) \in X \times X$ . Then,  $(X, d)$  is a metric space. The metric in theorem 2.1 is called the “metric induced by the norm” on  $X$ .

*Proof.* Let's show that  $d$  in theorem 2.1 satisfies all items (a) to (d) above. Item (a) follows from property (a) of the norm. Item (b) follows from property (b) of the norm, in fact  $\|\varphi - \psi\| = 0$  if and only if  $\varphi - \psi = 0$ , i.e., if and only if  $\varphi = \psi$ . Item (c) follows from property (c) of the norm, since

$$\begin{aligned} d(\varphi, \psi) &= \|\varphi - \psi\| = \|-(\psi - \varphi)\| = |-1| \|\psi - \varphi\| \\ &= \|\psi - \varphi\| = d(\psi, \varphi), \quad \forall (\varphi, \psi) \in X \times X. \end{aligned}$$

Item (d) follows from property (d) of the norm, since

$$\begin{aligned} d(\varphi, \psi) &= \|\varphi - \psi\| = \|\varphi - \eta + \eta - \psi\| \leq \|\varphi - \eta\| + \|\eta - \psi\| \\ &= d(\varphi, \eta) + d(\eta, \psi), \quad \forall (\varphi, \psi, \eta) \in X \times X \times X. \end{aligned}$$

This completes the proof. ■

The metric induced by a norm fulfills the following extra properties, the proof of which is straightforward and is left to the Reader as exercise:

1.  $d(\varphi + \eta, \psi + \eta) = d(\varphi, \psi)$  (translation invariance), and
2.  $d(\lambda\varphi, \lambda\psi) = |\lambda| d(\varphi, \psi)$  (homogeneity),

for all  $(\varphi, \psi, \eta) \in X \times X \times X$  and  $\lambda \in \mathbb{K}$ .

EXERCISE 2.1: Prove items 1 and 2 above.

Theorem 2.1 shows that any normed vector space is naturally endowed with a notion of distance. Please notice however that a normed vector space could be equipped also with distances other than the one induced by the norm; such distances are not necessarily related to the norm; furthermore, theorem 2.1 is not the only one possible distance built from the norm (see ?? 2.2).

EXERCISE 2.2: Let  $(X, \|\cdot\|)$  be a normed vector space. Let  $d: X \times X \rightarrow \mathbb{R}$  be the function defined by

$$d(\varphi, \psi) = \frac{\|\varphi - \psi\|}{1 + \|\varphi - \psi\|}, \quad \forall (\varphi, \psi) \in X \times X$$

Check whether  $d$  defines a metric on  $X$ .

Now that we have a “natural” notion of distance in normed vector spaces, i.e., theorem 2.1, all the concepts defined for metric spaces applies automatically, in particular, to normed linear spaces. Among these concepts are: continuity, limits, convergence, compactness, completeness, open sets etc.

LEMMA 2.1: Let  $(X, \|\cdot\|)$  be a normed vector space over  $\mathbb{K}$ . The following holds:

$$\left| \|\psi\| - \|\varphi\| \right| \leq \|\psi - \varphi\|, \quad (2.3)$$

for every  $(\psi, \varphi) \in X \times X$ .

*Proof.* The key ingredient is the triangle inequality of the norm. For all  $(\varphi, \psi) \in X \times X$ , we have

$$\|\varphi\| = \|\varphi - \psi + \psi\| \leq \|\varphi - \psi\| + \|\psi\|$$

and

$$\|\psi\| = \|\psi - \varphi + \varphi\| \leq \|\varphi - \psi\| + \|\varphi\|.$$

Thus,

$$\begin{aligned}\|\varphi\| - \|\psi\| &\leq \|\varphi - \psi\| \\ \|\psi\| - \|\varphi\| &\leq \|\varphi - \psi\|.\end{aligned}$$

which can be recasted in the single expression lemma 2.1. ■

Lemma 2.1 implies the continuity of the norm, as stated in the following theorem.

**THEOREM 2.2:** Let  $(X, \|\cdot\|)$  be any normed vector space over  $\mathbb{K}$ . The norm  $\|\cdot\|$  is continuous on  $X$  (with respect to the metric induced by the norm).

*Proof.* Lemma 2.1 implies that  $\|\cdot\|$  is Lipschitz and every Lipschitz function is continuous. ■

As a consequence of the continuity of the norm, if  $(X, \|\cdot\|)$  is a normed vector space and  $(\varphi_k)_{k \in \mathbb{N}}: \mathbb{N} \rightarrow X$  is a sequence in  $X$  convergent to some  $\varphi \in X$ , i.e.,

$$\lim_{k \rightarrow +\infty} \varphi_k = \varphi,$$

then

$$\|\varphi\| = \left\| \lim_{k \rightarrow +\infty} \varphi_k \right\| = \lim_{k \rightarrow +\infty} \|\varphi_k\|. \quad (2.4)$$

**DEFINITION 2.5** (equivalence of the norms): Let  $X$  be a vector space over  $\mathbb{K}$  and  $\|\cdot\|_1: X \rightarrow \mathbb{R}$  and  $\|\cdot\|_2: X \rightarrow \mathbb{R}$  two norms on  $X$ . The two norms are said to be “equivalent” if there exists a pair of strictly positive real numbers  $\lambda$  and  $\mu$  such that

$$\begin{aligned}\alpha \|\varphi\|_1 &\leq \|\varphi\|_2 \\ &\leq \beta \|\varphi\|_1,\end{aligned} \quad (2.5)$$

for all  $\varphi \in X$ .

Equivalent norms define the same notions of continuity and convergence and for many purposes do not need to be distinguished. As we shall prove later, for finite-dimensional (real or complex) vector space, all norms are equivalent. On the other hand, in the case of infinite-dimensional vector spaces, not all norms are equivalent and we need to specify which norm we are using.

*Completeness of a metric space*

Let us recall an important fact about Cauchy sequences. Give any metric space  $(X, d)$ , a convergent sequence in  $X$  is also a Cauchy sequence. The converse of this statement is not generally true however, i.e., there exists metric spaces where some Cauchy sequences do not need to converge.

A typical counter-example works as follow. Let  $(x_n)_{n \geq 1}$  be a sequence in  $\mathbb{R}$  such that for every  $n \geq 1$  we have:  $x_n \in \mathbb{Q}$  and  $\sqrt{2} - \frac{1}{n} < x_n < \sqrt{2} + \frac{1}{n}$ . (It is always possible to construct such sequence since  $\mathbb{Q}$  is dense in  $\mathbb{R}$ .) It is possible to prove that  $\lim_{n \rightarrow +\infty} x_n = \sqrt{2}$ . Being convergent in  $\mathbb{R}$ ,  $x_n$  is of Cauchy type in  $\mathbb{R}$ . The same sequence can be defined in  $\mathbb{Q}$  and it is of Cauchy type in  $\mathbb{Q}$ , but it is not convergent in  $\mathbb{Q}$  since  $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$ .

A metric space is “complete” if all Cauchy sequences are convergent.

*Banach space*

**DEFINITION 2.6** (Banach space): Let  $(X, \|\cdot\|)$  be a normed vector space over  $\mathbb{K}$ . If  $(X, d)$  (where  $d$  is the metric induced by the norm) is complete,  $(X, \|\cdot\|)$  is called a “Banach space”.



We are almost ready to introduce the notion of Hilber space, which is the setting where we will develop the operator formulation of non-relativistic quantum mechanics. The first ingredient is the “inner product”, defined below.

In the following, for every  $z \in \mathbb{C}$  we will denote the “complex conjugate” of  $z$  with  $z^*$  and the “modulus” of  $z$  with  $|z|$ . Remember that  $z = \Re z + \imath \Im z$  (where  $\Re z$  and  $\Im z$  are the real and imaginary parts of  $z$ , respectively),  $z^* = \Re z - \imath \Im z$  and  $|z|^2 = zz^*$ . The inverse of  $z \neq 0$  is  $z^{-1} = z^*/|z|^2$ . We have  $z \in \mathbb{R}$  if and only if  $z^* = z$ . The complex conjugation satisfies the “involution” property:  $(z^*)^* = z$ . Furthermore,  $(z_1 z_2)^* = z_1^* z_2^*$ ,  $(z_1 \pm z_2)^* = z_1^* \pm z_2^*$ , and  $|z_1 z_2| = |z_1| |z_2|$ .

Complex numbers  
notation

**DEFINITION 2.7 (inner product):** Let  $X$  be a vector space over  $\mathbb{K}$ . A “inner product” on  $X$  is any application  $X \times X \rightarrow \mathbb{K}$ , hereafter denoted by  $\langle \cdot | \cdot \rangle$ , satisfying the following properties:

Inner product

- (a)  $\langle \varphi | \psi + \eta \rangle = \langle \varphi | \psi \rangle + \langle \varphi | \eta \rangle, \quad \forall (\varphi, \psi, \eta) \in X \times X \times X;$
- (b)  $\langle \varphi | \lambda \psi \rangle = \lambda \langle \varphi | \psi \rangle, \quad \forall (\varphi, \psi) \in X \times X \text{ and } \lambda \in \mathbb{K};$
- (c)  $\langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^*, \quad \forall (\varphi, \psi) \in X \times X;$
- (d)  $\langle \varphi | \varphi \rangle \geq 0, \quad \forall \varphi \in X;$
- (e)  $\langle \varphi | \varphi \rangle = 0$  if and only if  $\varphi = 0$ .

Several remarks are in order.

*Remark.* Items (a) and (b) are equivalent to say that  $\langle \cdot | \cdot \rangle$  is linear on the second component, i. e.,  $\langle \cdot | \cdot \rangle$  satisfies items (a) and (b) if and only if

$$\langle \varphi | \lambda \psi + \mu \eta \rangle = \lambda \langle \varphi | \psi \rangle + \mu \langle \varphi | \eta \rangle,$$

for all  $(\varphi, \psi, \eta) \in X \times X \times X$  and for all  $(\lambda, \mu) \in \mathbb{K} \times \mathbb{K}$ . Item (c) together with items (a) and (b) implies that the inner product in general is *conjugate-linear* (or *anti-linear*) on the *first* component, namely

$$\langle \lambda \varphi + \mu \eta | \psi \rangle = \lambda^* \langle \varphi | \psi \rangle + \mu^* \langle \eta | \psi \rangle,$$

for all  $(\varphi, \psi, \eta) \in X \times X \times X$  and for all  $(\lambda, \mu) \in \mathbb{K} \times \mathbb{K}$ . Of course, if  $\mathbb{K} = \mathbb{R}$ ,  $\lambda^* = \lambda$ ,  $\mu^* = \mu$  and the inner product becomes linear also on the first component (thus, it is *bilinear*); but this is not the case if  $\mathbb{K} = \mathbb{C}$ , where complex conjugation appears.

*Remark.* Convention (b) is a matter of choice. Some authors prefer the different convention:

$$\langle \lambda \varphi | \psi \rangle = \lambda \langle \varphi | \psi \rangle, \quad \forall (\varphi, \psi) \in X \times X \text{ and } \lambda \in \mathbb{K};$$

with this latter convention, the inner product would become linear on the first component and conjugate-linear on the second one. The convention of having the inner product linear on the second component is the one most often employed by physicists, and the one used in this notes.

*Remark.* Regarding items (d) and (e), one may wonder what does mean  $\langle \varphi | \varphi \rangle \geq 0$ , since we expect  $\langle \varphi | \varphi \rangle \in \mathbb{K}$ , and if  $\mathbb{K} = \mathbb{C}$  it might seem that the inequality does not make sense. Actually, from item (c)

$$\langle \varphi | \varphi \rangle = \langle \varphi | \varphi \rangle^*, \quad \forall \varphi \in X,$$

*Remark.* Some authors prefer the notation  $(\varphi, \psi)$  instead of  $\langle \varphi | \psi \rangle$ ; the notation  $\langle \varphi | \psi \rangle$  is closer to the one used by physicists and it is the first step towards the introduction of Dirac's notation. (Dirac's notation is more than simply writing the inner product this way; we will discuss this point in connection with the spectral theorem of linear operators.) In Dirac notation the vector  $\psi$  is denoted by  $|\psi\rangle$ , and it is called "ket"; there is a "kind of conjugation" (more on this later on) that converts the analogous "ket"  $|\varphi\rangle$  to a so-called "bra"  $\langle \varphi|$  and the inner product is considered as a product between a "bra" and a "ket" (resulting in a "bracket"! ). Of course, this is just a suggestive naming convention.

*Inner product space*

**DEFINITION 2.8** (inner product space): A "inner product space" over  $\mathbb{K}$  is a pair  $(X, \langle \cdot | \cdot \rangle)$ , where  $X$  is a vector space over  $\mathbb{K}$  and  $\langle \cdot | \cdot \rangle : X \times X \rightarrow \mathbb{K}$  is an inner product on  $X$ .

The following lemma will be useful later on.

**LEMMA 2.2:** Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space over  $\mathbb{K}$ . If  $\psi = 0$ , then

$$\langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle = 0, \quad (2.6)$$

for every  $\varphi \in X$ .

The statement of this lemma looks rather trivial. However, a technical proof is given below. The linearity of the inner product is a key ingredient.

*Proof.* By linearity of the inner product,

$$\langle \varphi | \psi + \psi \rangle = \langle \varphi | \psi \rangle + \langle \varphi | \psi \rangle, \quad \forall (\varphi, \psi) \in X \times X.$$

In particular, if  $\psi = 0$  we have  $\psi + \psi = \psi$  and

$$\langle \varphi | \psi + \psi \rangle = \langle \varphi | \psi \rangle, \quad \forall \varphi \in X \text{ and } \psi = 0.$$

Thus

$$\langle \varphi | \psi \rangle + \langle \varphi | \psi \rangle = \langle \varphi | \psi \rangle, \quad \forall \varphi \in X \text{ and } \psi = 0,$$

which is an equation in  $\mathbb{K}$  for the unknown  $\langle \varphi | \psi \rangle$ .  $\langle \varphi | \psi \rangle$  is a solution of this equation *if and only if*  $\langle \varphi | \psi \rangle = 0$ . ■

*Norm induced by inner product*

Any inner product space is naturally endowed with a norm coming from the inner product. Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space over  $\mathbb{K}$ . Let  $\|\cdot\| : X \rightarrow \mathbb{R}$  defined by

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle}, \quad \forall \psi \in X. \quad (2.7)$$

Observe that such  $\|\cdot\|$  in eq. (2.7) is well-defined, since  $\langle \psi | \psi \rangle \geq 0$  for every  $\psi \in X$ . The square root is not ambiguous: it is not a square root of a complex number; it is the square root of a positive real number, we don't need to specify a branch for the square root function. We shall prove in a moment that  $\|\cdot\|$  is indeed a norm on  $X$ , this justifies the notation  $\|\cdot\|$ . Such norm is called the "norm induced by the inner product". Before proving this, we need a preliminary but extremely important result, which goes under the name of Cauchy-Schwarz inequality.

Cauchy-Schwarz inequality is of major importance. It is a key ingredient in several proofs of functional analysis. It has important implications also outside the realm of analysis. For example, the general formulation of the Heisenberg uncertainty principle in quantum mechanics (or the analogous time-bandwidth uncertainty principle for temporal signal transmission) is derived using the Cauchy-Schwarz inequality.

**THEOREM 2.3** (Cauchy-Schwarz inequality): Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space. The following holds:

Cauchy-Schwarz  
inequality

$$|\langle \varphi | \psi \rangle| \leq \|\varphi\| \|\psi\|, \quad \forall (\varphi, \psi) \in X \times X. \quad (2.8)$$

Equality holds if and only if the vectors are linearly dependent.

*Remark.* In theorem 2.3 we are using the definition eq. () but it is important to emphasize that we are *not* using (in both the statement and in the proof of Cauchy-Schwarz inequality) the fact that eq. () is a norm. We don't know at this point that eq. () defines a norm, we will prove that in the next theorem, using the Cauchy-Schwarz inequality.

*Proof.* We distinguish two cases:  $\psi = 0$  and  $\psi \neq 0$ .

If  $\psi = 0$ , then  $\langle \varphi | \psi \rangle = 0$  (see lemma 2.2) and  $\langle \psi | \psi \rangle = \|\psi\|^2 = 0$ , thus the inequality becomes  $0 \leq 0$ , which is satisfied.

Let us now consider  $\psi \neq 0$ . As a preliminary step, consider for every  $\lambda \in \mathbb{K}$

$$\begin{aligned} \langle \varphi - \lambda\psi | \varphi - \lambda\psi \rangle &= \langle \varphi | \varphi \rangle - \lambda \langle \varphi | \psi \rangle - \lambda^* \langle \psi | \varphi \rangle \\ &\quad + \lambda^* \lambda \langle \psi | \psi \rangle = \langle \varphi | \varphi \rangle - 2\Re(\lambda \langle \varphi | \psi \rangle) + \lambda^* \lambda \langle \psi | \psi \rangle. \end{aligned}$$

From item (d) in definition 2.7, the left-hand side of this equation is positive:

$$\langle \varphi - \lambda\psi | \varphi - \lambda\psi \rangle \geq 0$$

and it is zero if and only if  $\varphi - \lambda\psi = 0$  (i. e., if  $\varphi$  and  $\psi$  are linearly dependent). Thus

$$\langle \varphi | \varphi \rangle - 2\Re(\lambda \langle \varphi | \psi \rangle) + \lambda^* \lambda \langle \psi | \psi \rangle \geq 0,$$

for every  $\varphi \in X$ ,  $\psi \in X \setminus \{0\}$  and  $\lambda \in \mathbb{K}$ . Choose

$$\lambda = \frac{\langle \varphi | \psi \rangle^*}{\langle \psi | \psi \rangle}, \quad \psi \neq 0, \quad (2.9)$$

which makes sense since we are discussing the case  $\psi \neq 0$ . Plugin into the previous equation yields

$$\langle \varphi | \varphi \rangle - 2\Re\left(\frac{\langle \varphi | \psi \rangle^*}{\langle \psi | \psi \rangle} \langle \varphi | \psi \rangle\right) + \frac{\langle \varphi | \psi \rangle^*}{\langle \psi | \psi \rangle} \frac{\langle \varphi | \psi \rangle}{\langle \psi | \psi \rangle} \langle \psi | \psi \rangle \geq 0,$$

hence

$$\langle \varphi | \varphi \rangle - \frac{\langle \varphi | \psi \rangle^* \langle \varphi | \psi \rangle}{\langle \psi | \psi \rangle} \geq 0,$$

from which it follows

$$\langle \varphi | \varphi \rangle \langle \psi | \psi \rangle \geq |\langle \varphi | \psi \rangle|^2$$

(using the fact that  $\langle \psi | \psi \rangle > 0$ ). Taking the square root of both sides (notice that both sides are surely positive) yields the expected result. The equality holds if and only if  $\varphi - \lambda\psi = 0$ , i. e., if the two vectors are linearly dependent. ■

**EXERCISE 2.3:** For every fixed  $(\varphi, \psi) \in X \times X$ , consider  $f: \mathbb{C} \rightarrow \mathbb{R}$  defined by

$$\begin{aligned} f(\lambda) &= \|\varphi - \lambda\psi\|^2 \\ &= |\lambda|^2 \|\psi\|^2 - \lambda \langle \varphi | \psi \rangle - \lambda^* \langle \varphi | \psi \rangle^* + \|\varphi\|^2, \end{aligned} \quad (2.10)$$

as a function of  $\lambda \in \mathbb{C}$ . The positivity of the norm ensures that  $f(\lambda) \geq 0$  and  $f(\lambda) = 0$  if and only if  $\varphi = \lambda\psi$ , as already discussed in the proof of Cauchy-Schwarz inequality. Show that  $\lambda$  given by eq. () is a minimum of  $f$ . This fact can be used to explain the choice in eq. ().

**THEOREM 2.4:** Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space over  $\mathbb{K}$ .  $(X, \|\cdot\|)$  with  $\|\cdot\|$  defined by eq. () is a normed vector space over  $\mathbb{K}$ .

*Proof.* We need to check that eq. () makes sense and that it satisfies items (a) to (d) of the definition of the norm. The only non-trivial property is the subadditivity. It can be proved using the Cauchy-Schwarz inequality. In fact, for every  $\varphi, \psi$  in  $X$ ,

$$\begin{aligned} \|\varphi + \psi\|^2 &= \langle \varphi + \psi | \varphi + \psi \rangle \\ &= \langle \varphi | \varphi \rangle + \langle \varphi | \psi \rangle + \langle \psi | \varphi \rangle + \langle \psi | \psi \rangle \\ &= \|\varphi\|^2 + 2\Re\langle \varphi | \psi \rangle + \|\psi\|^2 \\ &\leq \|\varphi\|^2 + 2|\langle \varphi | \psi \rangle| + \|\psi\|^2 \\ &\leq \|\varphi\|^2 + 2\|\varphi\|\|\psi\| + \|\psi\|^2 \\ &= (\|\varphi\| + \|\psi\|)^2; \end{aligned}$$

taking the square root (all quantities involved are positive real numbers) yields the expected result. ■

**THEOREM 2.5:** Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space over  $\mathbb{K}$ .  $\langle \cdot | \cdot \rangle$  is a continuous function of both arguments (with respect to the topology induced by the inner product).

*Proof.* Let us discuss the continuity on the second argument (the continuity on the first argument can be handled in a similar way).

We need to show that  $\forall(\varphi, \psi) \in X \times X$  the following holds: for all real numbers  $\varepsilon > 0$ , there exists a real number  $\delta > 0$  such that, for all  $\eta \in X$ , if  $d(\psi, \eta) < \delta$  then  $\tilde{d}(\langle \varphi | \psi \rangle, \langle \varphi | \eta \rangle) < \varepsilon$ , where  $\tilde{d}$  is the euclidean distance in  $\mathbb{R}$ .

Notice that  $d(\psi, \eta) < \delta$  means

$$\|\psi - \eta\| < \delta,$$

and  $\tilde{d}(\langle \varphi | \psi \rangle, \langle \varphi | \eta \rangle) < \varepsilon$  means

$$|\langle \varphi | \psi \rangle - \langle \varphi | \eta \rangle| < \varepsilon.$$

Using linearity of the inner product and Cauchy-Schwarz inequality yields

$$|\langle \varphi | \psi \rangle - \langle \varphi | \eta \rangle| = |\langle \varphi | \psi - \eta \rangle| \leq \|\varphi\| \|\psi - \eta\| \leq \delta \|\varphi\|.$$

If  $\varphi = 0$ , the result is 0 and we are done. Otherwise, if  $\varphi \neq 0$ , it is enough to choose  $0 < \delta < \varepsilon / \|\varphi\|$ . ■

As a consequence of the continuity of the inner product, if  $(X, \langle \cdot | \cdot \rangle)$  is a inner product space and  $(\psi_k)_{k \geq 1}$  is a sequence in  $X$  such that  $\sum_{k \in \mathbb{N}} \psi_k$  is convergent to some  $\psi \in X$ , i.e.,

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \psi_k = \psi,$$

then

$$\langle \varphi | \psi \rangle = \left\langle \varphi \left| \lim_{n \rightarrow +\infty} \sum_{k=1}^n \psi_k \right. \right\rangle = \lim_{n \rightarrow +\infty} \left\langle \varphi \left| \sum_{k=1}^n \psi_k \right. \right\rangle = \lim_{n \rightarrow +\infty} \sum_{k=1}^n \langle \varphi | \psi_k \rangle. \quad (2.11)$$

**EXERCISE 2.4:** Justify all steps in eq. ().

DEFINITION 2.9 (Hilbert space): Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space over  $\mathbb{K}$ . If the metric space  $(X, d)$  (with the distance arising from the inner product) is complete,  $(X, \langle \cdot | \cdot \rangle)$  is called an “Hilbert space”.

Hilbert space

In short: Banach spaces are complete normed vector spaces and Hilbert spaces are complete inner product spaces. Hilbert spaces are a special case of Banach space, where the norm comes from an inner product. The underlying inner product inducing the norm introduces extra features (in particular, some related to the notion of orthogonality) which are not present in general Banach spaces (see next section for details on this).

In the next sections, we will focus on Hilbert spaces only, and we will not consider incomplete inner product space. One reason for this is that any incomplete metric space admit a completion (more on this later, when discussing BLT theorem).

Hereafter, if  $(X, \langle \cdot | \cdot \rangle)$  is an inner product space, if not explicitly stated, we will also intend that  $\|\cdot\|$  denotes the norm induced by the inner product, and that the convergence, etc refers to the distance induced by that norm.

Is it possible to distinguish if in a normed vector space, the norm is coming from some underlying inner space? The answer is yes, and we are going to prove immediately an interesting criterion to perform this check.

The following result is basic to establish the later theorem. It will also be useful later, when discussing positive operators.

LEMMA 2.3 (polarization identity): Let  $(X, \langle \cdot | \cdot \rangle)$  be an inner product space. In the case  $\mathbb{K} = \mathbb{R}$ ,

$$\langle \varphi | \psi \rangle = \frac{1}{4} \left( \|\varphi + \psi\|^2 - \|\varphi - \psi\|^2 \right), \quad \forall (\varphi, \psi) \in X \times X. \quad (2.12)$$

In the case  $\mathbb{K} = \mathbb{C}$ ,

$$\begin{aligned} \langle \varphi | \psi \rangle = & \frac{1}{4} \left( \|\varphi + \psi\|^2 - \|\varphi - \psi\|^2 \right) \\ & - \frac{-i}{4} \left( \|\varphi + i\psi\|^2 - \|\varphi - i\psi\|^2 \right), \quad \forall (\varphi, \psi) \in X \times X. \end{aligned} \quad (2.13)$$

Remark. Equation (2.12) can be recasted in a shorter form:

$$\boxed{\langle \varphi | \psi \rangle = \frac{1}{4} \sum_{k=0}^3 (-i)^k \|\varphi + i^k \psi\|^2.} \quad (2.14)$$

Remark. Different signs are possible in literature if a different convention is chosen for property (b) of definition 2.7, namely, if the inner product is chosen to be linear on the first component instead of the second one.

Proof. We discuss the case  $\mathbb{K} = \mathbb{C}$ . The proof can be easily modified for the case  $\mathbb{K} = \mathbb{R}$ .

By straightforward computation,

$$\|\varphi + \psi\|^2 = \|\varphi\|^2 + 2\Re \langle \varphi | \psi \rangle + \|\psi\|^2 \quad (2.15)$$

$$\|\varphi - \psi\|^2 = \|\varphi\|^2 - 2\Re \langle \varphi | \psi \rangle + \|\psi\|^2 \quad (2.16)$$

$$\|\varphi + i\psi\|^2 = \|\varphi\|^2 + 2\Im \langle \varphi | \psi \rangle + \|\psi\|^2 \quad (2.17)$$

$$\|\varphi - i\psi\|^2 = \|\varphi\|^2 - 2\Im \langle \varphi | \psi \rangle + \|\psi\|^2 \quad (2.18)$$

Subtracting eqs. (2.14) and (2.15) we get

$$\|\varphi + \psi\|^2 - \|\varphi - \psi\|^2 = 4\Re\langle\varphi|\psi\rangle;$$

subtracting eqs. (2.16) and (2.17) we get

$$\begin{aligned} & \|\varphi + \sqrt{-1}\langle\varphi|\psi\rangle\psi\| \\ & - \|\varphi - \sqrt{-1}\langle\varphi|\psi\rangle\psi\| = -4\Im\langle\varphi|\psi\rangle. \end{aligned}$$

Thus

$$\begin{aligned} & \frac{1}{4} \left( \|\varphi + \psi\|^2 - \|\varphi - \psi\|^2 \right) \\ & - \frac{\sqrt{-1}\langle\varphi|\psi\rangle}{4} \left( \|\varphi + \sqrt{-1}\langle\varphi|\psi\rangle\psi\| \right. \\ & \quad \left. - \|\varphi - \sqrt{-1}\langle\varphi|\psi\rangle\psi\| \right) = \Re\langle\varphi|\psi\rangle \\ & \quad + \sqrt{-1}\langle\varphi|\psi\rangle \\ & = \langle\varphi|\psi\rangle. \end{aligned}$$

This completes the proof.  $\blacksquare$

The following result allows a complete characterization of inner-product spaces among norm vector spaces. It is called Jordan-von Neumann theorem\*.

*Jordan-von Neumann  
theorem  
Parallelogram law*

**THEOREM 2.6** (Jordan-von Neumann theorem): Let  $(X, \|\cdot\|)$  be a normed vector space over  $\mathbb{K}$ . The norm  $\|\cdot\|$  comes from an inner product *if and only if* the following identity holds:

$$\|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 = 2\|\varphi\|^2 + 2\|\psi\|^2, \quad \forall(\varphi, \psi) \in X \times X. \quad (2.19)$$

Theorem 2.6 is known as “parallelogram law”.

*Proof.* Let’s prove first that if  $(X, \langle\cdot|\cdot\rangle)$  is a Hilbert space, then the norm induced by the inner product fullfills theorem 2.6. This is the straightforward part of the proof. By definition of the norm induced by the inner product,

$$\begin{aligned} \|\varphi + \psi\| &= \langle\varphi + \psi|\varphi + \psi\rangle \\ &= \langle\varphi|\varphi\rangle + \langle\varphi|\psi\rangle + \langle\psi|\varphi\rangle + \langle\psi|\psi\rangle \\ &= \|\varphi\|^2 + 2\Re\langle\varphi|\psi\rangle + \|\psi\|^2 \\ \|\varphi - \psi\| &= \langle\varphi - \psi|\varphi - \psi\rangle \\ &= \langle\varphi|\varphi\rangle - \langle\varphi|\psi\rangle - \langle\psi|\varphi\rangle + \langle\psi|\psi\rangle \\ &= \|\varphi\|^2 - 2\Re\langle\varphi|\psi\rangle + \|\psi\|^2 \end{aligned}$$

for all  $(\varphi, \psi) \in X \times X$ ; summing the two we have theorem 2.6.

Now, let’s prove the converse: if  $(X, \|\cdot\|)$  is a Banach space whose norm satisfies theorem 2.6, then the norm comes from an inner product. We discuss the case  $\mathbb{K} = \mathbb{C}$ , the case  $\mathbb{C} = \mathbb{R}$  is analogous. Let  $\langle\cdot|\cdot\rangle : X \times X \rightarrow \mathbb{K}$  be the application defined by

$$\begin{aligned} \langle\varphi|\psi\rangle &= \frac{1}{4} \left( \|\varphi + \psi\|^2 - \|\varphi - \psi\|^2 \right) \\ & - \frac{\sqrt{-1}\langle\varphi|\psi\rangle}{4} \left( \|\varphi + \sqrt{-1}\langle\varphi|\psi\rangle\psi\|^2 \right. \\ & \quad \left. - \|\varphi - \sqrt{-1}\langle\varphi|\psi\rangle\psi\|^2 \right), \end{aligned}$$

\* See Jordan and von Neumann, 1935.

for all  $(\varphi, \psi) \in X \times X$ . We will prove now that this application indeed is a inner product on  $X$ , i.e., we need to prove that it satisfies items (a) to (e) in definition 2.7. That this formula induced the norm is known from lemma 2.3.

Ad (a). For every  $(\varphi, \psi, \eta) \in X \times X \times X$ , by applying the parallelogram law we get

$$2 \|\varphi + \psi\|^2 + 2 \|\eta\|^2 = \|\varphi + \psi + \eta\|^2 + \|\varphi + \psi - \eta\|^2, \quad (2.20a)$$

$$2 \|\varphi - \psi\|^2 + 2 \|\eta\|^2 = \|\varphi - \psi + \eta\|^2 + \|\varphi - \psi - \eta\|^2, \quad (2.20b)$$

$$2 \|\varphi + \eta\|^2 + 2 \|\psi\|^2 = \|\varphi + \psi + \eta\|^2 + \|\varphi - \psi + \eta\|^2, \quad (2.20c)$$

$$2 \|\varphi - \eta\|^2 + 2 \|\psi\|^2 = \|\varphi + \psi - \eta\|^2 + \|\varphi - \psi - \eta\|^2. \quad (2.20d)$$

Subtracting eqs. (2.20a) and (2.20b) we get

$$2 \|\varphi + \psi\|^2 - 2 \|\varphi - \psi\|^2 = \|\varphi + \psi + \eta\|^2 + \|\varphi + \psi - \eta\|^2 - \|\varphi - \psi + \eta\|^2 - \|\varphi - \psi - \eta\|^2; \quad (2.21)$$

subtracting eqs. (2.20b) and (2.20c) we get

$$2 \|\varphi + \eta\|^2 - 2 \|\varphi - \eta\|^2 = \|\varphi + \psi + \eta\|^2 + \|\varphi - \psi + \eta\|^2 - \|\varphi + \psi - \eta\|^2 - \|\varphi - \psi - \eta\|^2; \quad (2.22)$$

summing eqs. (2.20d) and (2.21) we get

$$\begin{aligned} & 2 \|\varphi + \psi\|^2 - 2 \|\varphi - \psi\|^2 + 2 \|\varphi + \eta\|^2 - 2 \|\varphi - \eta\|^2 \\ &= \|\varphi + \psi + \eta\|^2 + \|\varphi + \psi - \eta\|^2 - \|\varphi - \psi + \eta\|^2 - \|\varphi - \psi - \eta\|^2 \\ &\quad + \|\varphi + \psi + \eta\|^2 + \|\varphi - \psi + \eta\|^2 - \|\varphi + \psi - \eta\|^2 - \|\varphi - \psi - \eta\|^2 \\ &= 2 \|\varphi + \psi + \eta\|^2 - 2 \|\varphi - \psi - \eta\|^2. \end{aligned} \quad (2.23)$$

Ad (b). This is the tricky part. Let

$$S = \{ z \in \mathbb{C} \mid \langle \varphi | z\psi \rangle = z \langle \varphi | \psi \rangle \}.$$

Notice that  $0 \in S$  and  $1 \in S$ . Furthermore, for every  $z$  and  $w$  in  $S$ , also  $z \pm w \in S$ , so  $\mathbb{Z} \subseteq S$ . Now, observe that for every  $z \in \mathbb{Z}$ ,  $w \in \mathbb{Z} \setminus \{0\}$ ,

$$z \langle \varphi | \psi \rangle = \langle \varphi | z\psi \rangle = \left\langle \varphi \left| w \frac{z}{w} \psi \right. \right\rangle = w \left\langle \varphi \left| \frac{z}{w} \psi \right. \right\rangle, \quad \forall (\varphi, \psi) \in X \times X,$$

yielding

$$\frac{z}{w} \langle \varphi | \psi \rangle = \left\langle \varphi \left| \frac{z}{w} \psi \right. \right\rangle,$$

so also  $\mathbb{Q} \subseteq S$ . Since  $\mathbb{Q}$  is dense in  $\mathbb{R}$  we *would like* to conclude that  $\mathbb{R} \subseteq S$ . This however requires the continuity of  $\langle \cdot | \cdot \rangle$  and we can't rely on the fact that the inner product is continuous on both components because we do not know whether  $\langle \cdot | \cdot \rangle$  is a inner product yet. Upgrading from  $\mathbb{Q}$  to  $\mathbb{R}$  requires that we prove that the Cauchy-Schwarz inequality holds using only the tools at our disposal so far. Fortunately, the usual trick works just fine. So, we can conclude that  $\mathbb{R} \in S$ . Finally, since a direct computation shows that

$$\langle \varphi | -_s \text{kmath_i maginary_u nit : ni} \psi \rangle = -_s \text{kmath_i maginary_u nit : ni} \langle \varphi | \psi \rangle,$$

we conclude that  $S = \mathbb{C}$ .

Ad (c). This is straightforward:

$$\begin{aligned} \langle \psi | \varphi \rangle &= \frac{1}{4} \left\{ \|\psi + \varphi\|^2 - \|\psi - \varphi\|^2 \right. \\ &\quad \left. - \|\varphi - \psi\|^2 \right\} \\ &= \frac{\|\psi + \varphi\|^2 - \|\psi - \varphi\|^2 - \|\varphi - \psi\|^2}{4} \end{aligned}$$

Ad (d). Also this is straightforward:

$$\begin{aligned} \langle \psi | \psi \rangle &= \frac{1}{4} \left\{ \|\psi + \psi\|^2 - \|\psi - \psi\|^2 \right. \\ &\quad \left. - \|\psi - \psi\|^2 \right\} \\ &= \frac{\|\psi + \psi\|^2 - \|\psi - \psi\|^2 - \|\psi - \psi\|^2}{4} \end{aligned}$$

Ad (e). Since we have already shown that  $\langle \psi | \psi \rangle = \|\psi\|^2$ ,  $\langle \psi | \psi \rangle = 0$  if and only if  $\|\psi\| = 0$ , that is if and only if  $\psi = 0$ . ■

Subsequent authors after Jordan and von Neumann have found norm conditions weaker than theorem 2.6 which characterize inner product spaces amongst normed vector spaces. See, e.g., [Reznick \[1978\]](#) and references therein; a recent paper is [Scholtes \[2010\]](#), which also contains references to the relevant literature on this subject.

### 2.3 A GLIMPSE AT CONVEX ANALYSIS

“Many inequalities in physics and mathematics have their origin in the notion of convexity.”

Folklore

Convex analysis is an important branch of analysis, with implications in optimization, etc\*. Here, we give general definitions for arbitrary convex sets and convex functions defined on convex sets. We apply these definitions to real-valued functions of one real variable as a special case and we use this to derive Young and Jensen inequalities, which will prove useful in later proofs. Jensen inequality plays a major role in information theory, entropy etc.

Convex set

**DEFINITION 2.10 (Convex set):** Let  $X$  be a vector space over  $\mathbb{K}$  and  $\Omega \subseteq X$  a non-empty subset of  $X$ .  $\Omega$  is “convex” if for every  $(\varphi, \psi) \in \Omega \times \Omega$ ,

$$\vartheta \varphi + (1 - \vartheta) \psi \in \Omega, \quad (2.24)$$

for every  $\vartheta \in [0, 1]$ <sup>†</sup> As a convention, the empty set is considered to be convex by definition.

*Remark.* Definition 2.10 is a “weighted average” of  $\varphi$  and  $\psi$ .

*Remark.* In order definition 2.10 to be meaningful, we need the notion of addition and scalar multiplication for a real number. The natural setting where these two operations are defined are the vector spaces. This is the

\* See, e.g., [Rockafellar, 1970](#).

<sup>†</sup> Of course, if  $\vartheta = 0$  we get  $\psi$  and if  $\vartheta = 1$  we get  $\varphi$ , so we could have considered equivalently  $]0, 1[$  instead of  $[0, 1]$  in definition 2.10.



reason why in definition 2.10  $X$  is a vector space over  $\mathbb{R}$  or  $\mathbb{C}$ . The notion of convexity may be generalized to more general settings if certain properties of convexity are taken as axioms, but we will not discuss this approach here.

*Remark.*  $X$  itself is convex. Furthermore, let  $\psi \in X$ ; then,  $\{\psi\}$  is a convex subset of  $X$ .

The relation with the usual “geometric” definition of convex set can be understood by introducing the notion of “line segment” for arbitrary real or complex vector spaces.

**DEFINITION 2.11** (line segment): Let  $X$  be a vector space over  $\mathbb{K}$  and consider  $(\varphi, \psi) \in X \times X$ . The set

*Line segment*

$$[\varphi, \psi] = \{ \vartheta \varphi + (1 - \vartheta) \psi \mid \forall \vartheta \in [0, 1] \} \quad (2.25)$$

is called the “(straight) line segment” joining  $\varphi$  and  $\psi$  (or with endpoints  $\varphi$  and  $\psi$ ).

*Remark.* Of course,  $[\varphi, \psi] = [\psi, \varphi]$ . Check as exercise. (Hint: It is enough to replace  $\lambda = 1 - \vartheta$ .)

Using definition 2.11 we can restate definition 2.10 in the following equivalent way:

**LEMMA 2.4:** Let  $X$  be a vector space over  $\mathbb{K}$  and  $\Omega \subseteq X$  any non-empty subset of  $X$ .  $\Omega$  is convex if and only if for every pair of points  $(\varphi, \psi) \in \Omega \times \Omega$ ,  $[\varphi, \psi] \subseteq \Omega$ .

*Alternative definition of convex set*

**EXERCISE 2.5:** Prove ??.

Before proving more details on convex sets, let us see some examples in  $\mathbb{R}$  and  $\mathbb{R}^N$ . First of all, recall the definition of “real interval”.

**DEFINITION 2.12** (real interval): Let  $I \subseteq \mathbb{R}$  a non-empty subset of  $\mathbb{R}$ .  $I$  is a “real interval” if  $(\varphi, \psi) \in I \times I$  and  $\eta \in \mathbb{R}$  such that  $\varphi \leq \eta \leq \psi$  implies  $\eta \in I$ . The empty set is regarded as a interval.

This definition covers different typologies of real intervals:  $[\varphi, \psi]$ ,  $] \varphi, \psi[$ ,  $] \varphi, \psi]$ ,  $[ \varphi, \psi[$ , etc. It covers also the case of a single real number  $I = \{\psi\}$  and the case where one of the two endpoints or both of them are  $\pm\infty$ .

**LEMMA 2.5:** Let  $\Omega \subseteq \mathbb{R}$  a subset of  $\mathbb{R}$ .  $\Omega$  is convex if and only if  $\Omega$  is a real interval.

**THEOREM 2.7** (Young inequality): For every  $A, B \geq 0$  and  $0 \leq \vartheta \leq 1$ ,

*Young inequality*

$$\boxed{A^\vartheta B^{1-\vartheta} \leq \vartheta A + (1 - \vartheta) B,} \quad (2.26)$$

*Remark.* As a special case, for  $\vartheta = 1/2$  we recover the arithmetic-geometric mean (AGM) inequality for two positive real numbers, namely

$$\sqrt{AB} \leq \frac{A+B}{2}, \quad A, B \geq 0. \quad (2.27)$$

The general arithmetic-geometric mean inequality with  $n \in \mathbb{N}$  terms will be proven later, as a special case of the Jensen inequality. There are other

interesting ways to prove the AGM inequality for two terms, for example rearranging the square terms:

$$0 \leq (A - B)^2 = (A + B)^2 - 4AB, \quad \forall A, B \in \mathbb{R},$$

yielding

$$AB \leq \left( \frac{A+B}{2} \right)^2, \quad \forall A, B \in \mathbb{R}.$$

If  $A, B \geq 0$  we can take the square root and find eq. (2.26). There are also several other proofs for AGM inequality with  $n \in \mathbb{N}$  terms (e.g., proof by induction).

*Remark.* Geometric meaning of eq. (2.26):  $AB$  is the area of a rectangle with sides of length  $A$  and  $B$  and  $\frac{A+B}{2}$  is the semi-perimeter of such rectangle.  $\sqrt{AB}$  can be viewed as the length of the side of a square having the same area of the rectangle. Re-arranging the terms we have  $4\sqrt{AB} \leq 2A + 2B$  which means that the square has the smallest perimeter amongst all rectangles of equal area.

*Remark.* Young inequality is often stated in a slightly different form. Let

$$\frac{1}{p} + \frac{1}{q} = 1,$$

put  $\vartheta = 1/p$  so  $1 - \vartheta = 1/q$  and set  $A = a^p$  and  $B = b^q$ . This way theorem 2.7 reads

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (2.28)$$

*Proof.* There are various proof, mainly based on convexity. Our proof uses the fact that the exponential function is convex on  $\mathbb{R}$ , thus

$$\exp \vartheta \varphi + (1 - \vartheta) \psi \leq \vartheta \exp \varphi + (1 - \vartheta) \exp \psi \quad (2.29)$$

for all  $\vartheta \in [0, 1]$  and  $(\varphi, \psi) \in \mathbb{R} \times \mathbb{R}$ . Using  $A = \exp \varphi$  and  $B = \exp \psi$  (this makes sense because  $A$  and  $B$  are positive) yields theorem 2.7. ■

Another possible proof goes as follows. Consider  $f_\vartheta: ]0, +\infty[ \rightarrow \mathbb{R}$  defined by

$$f_\vartheta(x) = x^\vartheta - \vartheta x + (\vartheta - 1), \quad \forall x > 0,$$

where by definition

$$x^\vartheta = \exp \vartheta \log x, \quad \forall x > 0.$$

Notice that when  $\vartheta = 0$  or  $\vartheta = 1$  the function is constant, in particular  $f_0(x) = f_1(x) = 0$ . Study the function and show that it has a maximum for  $x = 1$ . Since  $f_\vartheta(1) = 0$ , we get  $f_\vartheta(x) \leq 0$ .

## 2.4 CANONICAL PROTOTYPES OF BANACH AND HILBERT SPACES

### 2.4.1 The Hilbert spaces $\mathbb{R}^N$ and $\mathbb{C}^N$

Let  $N \in \mathbb{N} \setminus \{0\}$  be any positive integer number. Consider the vector space  $\mathbb{C}^N$ , whose elements are  $N$ -ple of complex numbers, endowed with the usual operations of elementwise addition and elementwise multiplication of an  $N$ -ple by a complex number. For every

$$\vec{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_N \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix},$$

in  $\mathbb{C}^N$ , define an application  $\langle \cdot | \cdot \rangle : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{C}$  as follows:

$$\langle \vec{\varphi} | \vec{\psi} \rangle = \sum_{k=1}^N \varphi_k^* \psi_k. \quad (2.30)$$

In a similar manner one can define the usual (Euclidean) inner product on  $\mathbb{R}^N$ , which takes the same form of eq. (2.30) without complex conjugation (because in that case all numbers involved are real numbers).

**THEOREM 2.8:**  $(\mathbb{C}^N, \langle \cdot | \cdot \rangle)$ , where  $\langle \cdot | \cdot \rangle$  is defined by eq. (2.30), is an inner product space.

*Proof.* It is straightforward to show that  $\langle \cdot | \cdot \rangle$  defined in eq. (2.30) satisfies Items (a) to (e) of definition 2.7. ■

**THEOREM 2.9:**  $(\mathbb{C}^N, \langle \cdot | \cdot \rangle)$ , where  $\langle \cdot | \cdot \rangle$  is defined by eq. (2.30), is an Hilbert space.

*Proof.* We need to check completeness. ■

#### 2.4.2 The Hilbert space $\ell^2$

Let  $\ell^2$  be the space of all and only the sequences  $(\varphi_k)_{k \geq 1}$  in  $\mathbb{K}$  such that the series  $\sum_{k=1}^{+\infty} |\varphi_k|^2$  is convergent in  $\mathbb{R}$ .

It is easy to show that, with the usual addition and multiplication for an element of  $\mathbb{K}$ ,  $\ell^2$  is a vector space over  $\mathbb{K}$ . Define

$$\langle \varphi | \psi \rangle = \sum_{k=1}^{+\infty} \varphi_k^* \psi_k, \quad (2.31)$$

for every  $\varphi = (\varphi_k)_{k \geq 1}$  and  $\psi = (\psi_k)_{k \geq 1}$  in  $\ell^2$ .

**THEOREM 2.10:**  $(\ell^2, \langle \cdot | \cdot \rangle)$ , where  $\langle \cdot | \cdot \rangle$  is defined by section 2.4.2, is an inner product space.

*Proof.* Similar to the proof that  $\mathbb{C}^N$  is an inner product space, with some care to handle the limits of the various sums. ■

**THEOREM 2.11:**  $(\ell^2, \langle \cdot | \cdot \rangle)$ , where  $\langle \cdot | \cdot \rangle$  is defined by section 2.4.2, is an Hilbert space.

#### 2.4.3 The Hilbert space of square integrable function

Let  $\Omega \subseteq \mathbb{R}^N$  be an arbitrary non-empty subset of  $\mathbb{R}^N$  for some  $N \in \mathbb{N}$  and let  $\tilde{\mathcal{L}}^2(\Omega)$  be the set consisting of all and only the functions defined on  $\Omega$  and taking values on  $\mathbb{K}$ , i. e.,  $\psi: \Omega \rightarrow \mathbb{K}$ , such that their square modulus is Lebesgue-integrable over  $\Omega$ , i. e., the following integral exists and it is finite (in the sense of Lebesgue):

$$\int_{\Omega} |\psi|^2.$$

*Remark.* Even if  $\mathbb{K} = \mathbb{C}$ , the modulus is real-valued so the integral above is always an integral of a real function.

It is not difficult to make  $\tilde{L}^2(\Omega)$  a vector space. Our goal would be to endow  $\tilde{L}^2(\Omega)$  also with an inner product. A problem would arise however, due to the fact that there are non-zero functions in  $\tilde{L}^2(\Omega)$  whose square modulus integral is zero, and this would ultimately make not possible to satisfy property (e) of the inner product.

To overcome this difficulty, a slightly more technical construction is needed. The idea is to “identify” two functions when they are equal almost everywhere (i. e., everywhere but on a set of zero Lebesgue measure). The formal construction involves working with equivalence classes and proceeds as follows. First, we will equip  $\tilde{L}^2(\Omega)$  with an equivalence relation, which allows to identify two functions whenever they are equal almost everywhere (i. e., when they differ only on a set of Lebesgue zero measure). Then, the quotient space with respect to this equivalence relation can be made a vector space over  $\mathbb{K}$  (with suitable definitions of addition and multiplication by a element of  $\mathbb{K}$ ) and can be equipped with an inner product. We will see that, with this inner product, the quotient space becomes an Hilbert space.

As a preliminary step, consider the subset  $M \subseteq \tilde{L}^2(\Omega)$  defined as follows:

$$M = \left\{ \psi \in \tilde{L}^2(\Omega) \mid \int_{\Omega} |\psi|^2 = 0 \right\}$$

The following lemma is not strictly necessary to be mentioned, but it makes the next proofs clearer.

**LEMMA 2.6:** For every  $(\varphi, \psi) \in M \times M$  and for every  $\lambda \in \mathbb{C}$ ,  $\varphi + \psi$  and  $\lambda\psi$  belong to  $M$ .

In words: any linear combination of vectors of  $M$  is an element of  $M$  itself.

*Proof.* By linearity of the integral,

$$\int_{\Omega} |\lambda\psi|^2 = \int_{\Omega} |\lambda|^2 |\psi|^2 = |\lambda|^2 \int_{\Omega} |\psi|^2,$$

thus  $\lambda\psi \in M$ . Using the monotonicity of the integral,

$$0 \leq \int_{\Omega} |\varphi + \psi|^2 \leq \int_{\Omega} |\psi|^2 + \int_{\Omega} |\varphi|^2 = 0,$$

thus  $\psi + \varphi \in M$ . ■

Let  $\sim$  be the relation on  $\tilde{L}^2(\Omega)$  defined in this way: for every  $(\varphi, \psi) \in \tilde{L}^2(\Omega) \times \tilde{L}^2(\Omega)$ ,  $\varphi \sim \psi$  if  $\varphi - \psi \in M$ . In words:  $\varphi \sim \psi$  if the two functions agree outside a set of (Lebesgue) zero measure.

**LEMMA 2.7:**  $\sim$  is an equivalence relation over  $\tilde{L}^2(\Omega)$ .

*Proof.* Let's check the equivalence relation properties:

**REFLEXIVITY :** for every  $\psi \in \tilde{L}^2(\Omega)$ ,  $\psi - \psi = 0$  (where  $0$  denotes the identically zero function, i. e., the function which takes value zero everywhere on  $\Omega$ ) and thus  $\psi - \psi \in M$ ;

**SYMMETRY :** for every  $\psi$  and  $\varphi$  in  $\tilde{L}^2(\Omega)$ , if  $\psi \sim \varphi$  also  $\varphi \sim \psi$ ; in fact,  $\psi \sim \varphi$  means  $\psi - \varphi \in M$  and, for lemma 2.6, also  $\varphi - \psi = (-1)(\psi - \varphi) \in M$ ;

**TRANSITIVITY :** for every  $\psi$ ,  $\varphi$  and  $\eta$  in  $\tilde{L}^2(\Omega)$ , if  $\psi \sim \eta$  and  $\eta \sim \varphi$ , then  $\psi \sim \varphi$ ; in fact, if  $\psi - \eta \in M$  and  $\eta - \varphi \in M$ , then also  $(\psi - \eta) + (\eta - \varphi) \in M$  by lemma 2.6.

■

We introduce the following notation: for every  $\psi$  in  $\tilde{L}^2(\Omega)$ , let  $[\psi]$  denote the equivalence class of  $\psi$  under  $\sim$ ; furthermore, let  $L^2(\Omega)$  denote the quotient space (i.e., the space of all possible equivalence classes) of  $\tilde{L}^2(\Omega)$  by  $\sim$ .

Define addition and multiplication by a scalar constant in  $L^2(\Omega)$  in the following way. For every  $[\psi]$  and  $[\varphi]$  in  $L^2(\Omega)$  and  $\lambda \in \mathbb{K}$ , put

$$\begin{aligned} [\psi] + [\varphi] &= [\psi + \varphi], \\ \lambda[\psi] &= [\lambda\psi]. \end{aligned}$$

First of all, let us check that these operations are well-defined.

LEMMA 2.8: For every  $\psi$  and  $\tilde{\psi}$  in  $[\psi]$  and for every  $\varphi$ ,  $\tilde{\varphi}$  in  $[\varphi]$ ,

$$\begin{aligned} [\psi] + [\varphi] &= [\tilde{\psi}] + [\tilde{\varphi}], \\ \lambda[\psi] &= \lambda[\tilde{\psi}]. \end{aligned}$$

*Proof.* There exists  $\eta$  and  $\xi$  in  $M$  such that  $\tilde{\psi} = \psi + \eta$  and  $\tilde{\varphi} = \varphi + \xi$ ; then,  $\tilde{\psi} + \tilde{\varphi} = (\psi + \eta) + (\varphi + \xi) = (\psi + \varphi) + (\eta + \xi)$ , where  $\eta + \xi \in M$ . Thus  $(\tilde{\psi} + \tilde{\varphi}) - (\psi + \varphi) = \eta + \xi \in M$ ,  $(\tilde{\psi} + \tilde{\varphi}) \sim (\psi + \varphi)$ . In the same way one proves that  $\lambda\tilde{\psi} \sim \lambda\psi$ . ■

THEOREM 2.12:  $(L^2(\Omega), +, \cdot)$  is a vector space over  $\mathbb{K}$ , where the addition and multiplication are those defined above.

It is a standard result of linear algebra that the quotient space with the above definitions of addition and multiplication is a vector space. More generally, this result applies to every quotient space, no matter what is the underlying set and what is the specific equivalence relation. We leave the proof of theorem 2.12 as exercise.

EXERCISE 2.6: Prove theorem 2.12. Generalize the proof to arbitrary quotient spaces under a generic equivalence relation.

In  $L^2(\Omega)$ , define an application  $\langle \cdot | \cdot \rangle$  by letting

$$\langle [\varphi] | [\psi] \rangle = \int_{\Omega} \varphi^* \psi, \quad (2.32)$$

where on the right hand side  $\varphi$  is any function belonging to  $[\varphi]$  and  $\psi$  is any function belonging to  $[\psi]$ .

*Remark.* In order to simplify the notation, we will denote the equivalence class  $[\psi]$  containing  $\psi$  by  $\psi$  itself.

*Remark.* Given  $\varphi: \Omega \rightarrow \mathbb{K}$ , the function  $\varphi^*: \Omega \rightarrow \mathbb{K}$  is defined by  $\varphi^*(x) = (\varphi(x))^*$  for all  $x \in \Omega$ .

We need to show that such  $\langle \cdot | \cdot \rangle$  is well-defined, i.e.: (a) show that the integral exists and is convergent, and (b) show that the integral is independent from the choice of  $\psi \in [\psi]$  and  $\varphi \in [\varphi]$ .

Let us recall the following theorem from the theory of Lebesgue integration.

THEOREM 2.13: Let  $\varphi: \Omega \rightarrow \mathbb{K}$  be Lebesgue integrable on  $\Omega$  and let  $\psi: \Omega \rightarrow \mathbb{K}$  be any function satisfying

$$|\psi(x)| \leq |\varphi(x)|,$$

for all  $x \in \Omega$ . Then,  $\psi$  is Lebesgue integrable on  $\Omega$ .

To show that the integral exists, notice that for every  $(z, w) \in \mathbb{C} \times \mathbb{C}$ , we have

$$|zw| \leq \frac{1}{2}|z|^2 + \frac{1}{2}|w|^2; \quad (2.33)$$

in fact,

$$0 \leq (|z| - |w|)^2 = |z|^2 + |w|^2 - 2|zw|.$$

Thus,

$$|\varphi^*(x)\psi(x)| \leq \frac{1}{2}|\varphi(x)|^2 + \frac{1}{2}|\psi(x)|^2, \quad \forall x \in \Omega,$$

thus the integral of  $|\varphi^*(x)\psi(x)|$  exists and is absolutely convergent, and this ensures the convergence of the integral of  $\varphi^*(x)\psi(x)$ .

**EXERCISE 2.7:** Show that the integral in eq. () is independent from the choice of  $\psi \in [\psi]$  and  $\varphi \in [\varphi]$ . Hint: write  $\psi = \psi - \tilde{\psi} + \tilde{\psi}$  and remember that

**THEOREM 2.14:**  $(L^2(\Omega), \langle \cdot | \cdot \rangle)$ , where  $\langle \cdot | \cdot \rangle$  is defined by eq. (), is a inner product space.

*Proof.* We need to show that  $\langle \cdot | \cdot \rangle$  defined by eq. () satisfies items (a) to (e) in definition 2.7.

Linearity of  $\langle \cdot | \cdot \rangle$  follows immediately from the linearity of the integral:

$$\begin{aligned} \langle \varphi | \psi + \eta \rangle &= \int_{\Omega} \varphi^* (\psi + \eta) = \int_{\Omega} \varphi^* \psi + \int_{\Omega} \varphi^* \eta = \langle \varphi | \psi \rangle + \langle \varphi | \eta \rangle \\ \langle \varphi | \lambda \psi \rangle &= \int_{\Omega} \varphi^* (\lambda \psi) = \lambda \int_{\Omega} \varphi^* \psi = \lambda \langle \varphi | \psi \rangle \end{aligned}$$

Furthermore,

$$\langle \psi | \varphi \rangle = \int_{\Omega} \psi^* \varphi = \int_{\Omega} (\psi \varphi^*)^* = \left( \int_{\Omega} \varphi^* \psi \right)^* = \langle \varphi | \psi \rangle^*$$

The fact that complex conjugation can be moved outside the integration symbol is rigorously justified as follows. Let  $\psi: \Omega \rightarrow \mathbb{C}$  be a complex-valued function defined on  $\Omega$ , and set  $\psi_1, \psi_2: \Omega \rightarrow \mathbb{R}$  by letting  $\psi_1 = \Re \psi$  and  $\psi_2 = \Im \psi$ . By definition,  $\psi$  is Lebesgue integrable on  $\Omega$  if and only if  $\psi_1$  and  $\psi_2$  are Lebesgue integrable on  $\Omega$  and in that case we put

$$\begin{aligned} \int_{\Omega} \psi &= \int_{\Omega} \psi_1 + i \int_{\Omega} \psi_2 \\ &: \text{ni} \int_{\Omega} \psi_2 \end{aligned}$$

Clearly, if  $\psi$  is Lebesgue integrable,  $\psi^*$  is Lebesgue integrable and

$$\begin{aligned} \int_{\Omega} \psi^* &= \int_{\Omega} \psi_1 - i \int_{\Omega} \psi_2 \\ &: \text{ni} \int_{\Omega} \psi_2 \\ &= \left( \int_{\Omega} \psi \right)^*. \end{aligned}$$

Finally, the monotonicity of the integral ensures that

$$\langle \psi | \psi \rangle = \int_{\Omega} \psi^* \psi = \int_{\Omega} |\psi|^2 \geq 0,$$

and  $\int_{\Omega} |\psi|^2 = 0$  if and only if  $\psi$  belongs to the class of equivalence of the function which is identically zero on  $\Omega$ , i. e.,  $\psi = 0$ . ■

THEOREM 2.15 (Riesz-Fischer):  $(L^2(\Omega), \langle \cdot | \cdot \rangle)$  is an Hilbert space.

*Proof.* We need to check completeness. ■

#### 2.4.4 Hölder and Minkovsky inequalities and p-norms

We will discuss the p-norms for  $\mathbb{R}^N$ ,  $\mathbb{C}^N$ ,  $l^p$  and  $L^p(\Omega)$ . All these examples are generalization of those discussed in previous sections.

The Banach spaces  $l^p$  and  $L^p(\Omega)$

We know that for  $p = 2$ ,  $l^2$  is a Hilbert space. This is also the only value of  $p$  for which  $l^p$  is an Hilbert space.

THEOREM 2.16: If  $p \neq 2$ ,  $l^p$  is not a Hilbert space.

*Proof.* By absurdum, let's suppose  $l^p$  is an Hilbert space for  $p \neq 2$ . Then the parallelogram law applies for all  $\varphi = (\varphi_n)_{n \geq 0}$  and  $\psi = (\psi_n)_{n \geq 0}$  in  $l^p$ . Select

$$\varphi_n = \begin{cases} 1 & \text{if } n = 1 \\ 1 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}, \quad \psi_n = \begin{cases} 1 & \text{if } n = 1 \\ -1 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.$$

and so

$$\varphi_n + \psi_n = \begin{cases} 2 & \text{if } n = 1 \\ 0 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}, \quad \varphi_n - \psi_n = \begin{cases} 0 & \text{if } n = 1 \\ 2 & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}.$$

We have

$$\|\varphi + \psi\| = \left( \sum_{n=1}^{+\infty} |\varphi_n + \psi_n|^p \right)^{\frac{1}{p}} = (2^p)^{\frac{1}{p}} = 2,$$

$$\|\varphi - \psi\| = \left( \sum_{n=1}^{+\infty} |\varphi_n - \psi_n|^p \right)^{\frac{1}{p}} = (2^p)^{\frac{1}{p}} = 2,$$

$$\|\varphi\| = \left( \sum_{n=1}^{+\infty} |\varphi_n|^p \right)^{\frac{1}{p}} = (1^p + 1^p)^{\frac{1}{p}} = 2^{\frac{1}{p}},$$

$$\|\psi\| = \left( \sum_{n=1}^{+\infty} |\psi_n|^p \right)^{\frac{1}{p}} = (1^p + 1^p)^{\frac{1}{p}} = 2^{\frac{1}{p}},$$

and the parallelogram law

$$\|\varphi - \psi\|^2 + \|\varphi + \psi\|^2 = 2\|\varphi\|^2 + 2\|\psi\|^2$$

reads in this case

$$2^2 + 2^2 = 2 \cdot 2^{\frac{2}{p}} + 2 \cdot 2^{\frac{2}{p}},$$

that is

$$2 = 2^{\frac{2}{p}}.$$

This equation is true if and only if  $\frac{2}{p} = 1$ , i.e., if  $p = 2$ , which is not the case. ■

## 2.5 ORTHOGONALITY

The notion of inner product allows to naturally equip an inner product space with a notion of orthogonality between vectors. Some of the following results (e.g., Pitagorean theorem) can be used to some extent to export the notion of orthogonality to general Banach spaces, but here we will restrict ourselves to Hilbert spaces.

Orthogonality of two  
vectors

DEFINITION 2.13: Let  $(X, \langle \cdot, \cdot \rangle)$  be any inner product space. For every  $(\varphi, \psi) \in X \times X$ ,  $\varphi$  and  $\psi$  are said to be mutually “orthogonal” if

$$\langle \varphi | \psi \rangle = 0. \quad (2.34)$$

The extension to a (possibly not-countable) set of vectors is trivial and it is formalized by the following definition.

Orthogonality of a set  
of vectors

DEFINITION 2.14: Let  $(X, \langle \cdot, \cdot \rangle)$  be any inner product space and  $W \subseteq X$  a non-empty subset of  $X$ .  $W$  is a set of mutually orthogonal vectors if

$$\langle \varphi | \psi \rangle = 0, \quad \forall (\varphi, \psi) \in W \times W \text{ with } \varphi \neq \psi. \quad (2.35)$$

If moreover

$$\langle \varphi | \varphi \rangle = 1, \quad \forall \varphi \in W \quad (2.36)$$

the set  $W$  is said to be “orthonormal”.

In particular, given a sequence  $\mathbb{N} \rightarrow X$  of vectors of  $X$ , we say that it is “orthonormal” if the image set is an orthonormal set.

THEOREM 2.17 (Pythagorean theorem): Let  $(X, \langle \cdot, \cdot \rangle)$ .

1. for every orthogonal pair of vectors  $\varphi$  and  $\psi$  in  $X$ ,

$$\|\varphi + \psi\|^2 = \|\varphi\|^2 + \|\psi\|^2;$$

2. for every orthogonal finite set of  $N \in \mathbb{N}$  vectors  $\varphi_1, \varphi_2, \dots, \varphi_N$  in  $X$ ,

$$\left\| \sum_{k=1}^N \varphi_k \right\|^2 = \sum_{k=1}^N \|\varphi_k\|^2$$

3. if  $(\varphi_k)_{k \in \mathbb{N}}$  is a sequence  $\mathbb{N} \rightarrow X$  of mutually orthogonal vectors of  $X$ , then  $\sum_{k=1}^{+\infty} \varphi_k$  is convergent in  $X$  if and only if  $\sum_{k=1}^{+\infty} \|\varphi_k\|^2$  is convergent in  $\mathbb{R}$  and in that case we have the following identity:

$$\left\| \sum_{k=1}^{+\infty} \varphi_k \right\|^2 = \sum_{k=1}^{+\infty} \|\varphi_k\|^2.$$

*Proof.* In the first case the proof is straightforward, we have already found the same expression in the proof of the parallelogram identity, we have by direct computation that

$$\begin{aligned} \|\varphi + \psi\|^2 &= \langle \varphi + \psi | \varphi + \psi \rangle \\ &= \langle \varphi | \varphi \rangle + \langle \varphi | \psi \rangle + \langle \psi | \varphi \rangle + \langle \psi | \psi \rangle \\ &= \|\varphi\|^2 + 2 \underbrace{\Re \langle \varphi | \psi \rangle}_{=0} + \|\psi\|^2 \\ &= \|\varphi\|^2 + \|\psi\|^2. \end{aligned}$$



In the second case, the proof is by induction on  $N$ . The case  $N = 2$  has been shown explicitly in the first part of the proof. Now, let's check that

$$\left\| \sum_{k=1}^N \varphi_k \right\|^2 = \sum_{k=1}^N \|\varphi_k\|^2$$

for a given  $N$  implies

$$\left\| \sum_{k=1}^{N+1} \varphi_k \right\|^2 = \sum_{k=1}^{N+1} \|\varphi_k\|^2$$

for  $N + 1$ . We have

The third case shows a classic trick of using the Cauchy criterion of convergence.

■

## 2.6 COMPLETE ORTHONORMAL SETS AND THEIR CHARACTERIZATION

DEFINITION 2.15: Let  $(X, \langle \cdot | \cdot \rangle)$  a Hilbert space and  $(\varphi_k)_k$  a family of vectors in  $X$ .  $(\varphi_k)_k$  is called a “complete orthonormal set” if

1.  $(\varphi_k)_k$  is an orthonormal set in  $X$ ;
2. for every  $\psi \in X$ ,

$$\langle \varphi_k | \psi \rangle = 0, \quad \forall k$$

if and only if  $\psi = 0$ .

## 2.7 THE FOURIER SERIES IN THE SPACE OF SQUARE INTEGRABLE FUNCTIONS

## 2.8 LINEAR OPERATORS: FIRST PROPERTIES

## 2.9 RIESZ REPRESENTATION THEOREM

## REFERENCES FOR § 2

BERBERIAN, STERLING K.

1976 *Introduction to Hilbert Space*, Chelsea Publishing Company, New York. (Cit. on p. 5.)

DEBNATH, L. and P. MIKUSIŃSKI

2005 *Hilbert Spaces with Applications*, Elsevier Academic Press, Amsterdam, The Netherlands, ISBN: 9780122084386. (Cit. on p. 5.)

HELMBERG, GILBERT

1969 *Introduction to Spectral Theory in Hilbert Space*, North-Holland series in applied mathematics and mechanics, American Elsevier Publishing Company, Amsterdam, The Netherlands, ISBN: 0-7204-2356-2. (Cit. on p. 5.)

HUTSON, V. and J. S. PYM

1980 *Applications of Functional Analysis and Operator Theory*, Academic Press, London. (Cit. on p. 5.)

JORDAN, PASCUAL and JOHN VON NEUMANN

- 1935 "On inner products in linear, metric spaces", *Annals of Mathematics*, 36, 3 [July 1935], pp. 719-723. (Cit. on p. 14.)

REED, MARTIN and BERRY SIMON

- 1980 *Functional Analysis*, 2nd ed., Methods of Modern Mathematical Physics, Academic Press, London, vol. 1, ISBN: 0-12-585050-6. (Cit. on p. 5.)

REZNICK, BRUCE

- 1978 "Banach spaces which satisfy linear identities." *Pacific J. Math.*, 74, 1, pp. 221-233. (Cit. on p. 16.)

ROCKAFELLAR, TYRRELL R.

- 1970 *Convex Analysis*, Princeton University Press, Princeton, New Jersey, ISBN: 978-0-691-01586-6. (Cit. on p. 16.)

SCHOLTES, SEBASTIAN

- 2010 *A characterisation of inner product spaces by the maximal circumradius of spheres*, Feb. 3, 2010, arXiv: 1202.0503 [math]. (Cit. on p. 16.)

TESCHL, GERALD

- 2009 *Mathematical Methods in Quantum Mechanics with Applications to Schrödinger Operators*, Graduate Studies in Mathematics, AMS, Providence, Rhode Island, USA, vol. 99, ISBN: 978-0-8218-4660-5 (alk. paper), <http://www.ams.org/bookpages/gsm-99>. (Cit. on p. 5.)

THE RULES OF THE GAME

---

“ Quantum phenomena do not occur in a Hilbert space. They occur in a laboratory. ”

Asher Peres *Quantum theory: concepts and methods*







## Part II

### THE CORE























## Part III

### ADVANCED TOPICS





















## Part IV

### APPENDICES







## BIBLIOGRAPHY

---

### REFERENCES FOR § 1

SCHWINGER, JULIAN

- 2001 *Quantum Mechanics. Symbolism of Atomic Measurements*, Springer-Verlag, Berlin, ISBN: 3-540-41408-8. (Cit. on p. 3.)

### REFERENCES FOR § 2

BERBERIAN, STERLING K.

- 1976 *Introduction to Hilbert Space*, Chelsea Publishing Company, New York. (Cit. on p. 5.)

DEBNATH, L. and P. MIKUSIŃSKI

- 2005 *Hilbert Spaces with Applications*, Elsevier Academic Press, Amsterdam, The Netherlands, ISBN: 9780122084386. (Cit. on p. 5.)

HELMBERG, GILBERT

- 1969 *Introduction to Spectral Theory in Hilbert Space*, North-Holland series in applied mathematics and mechanics, American Elsevier Publishing Company, Amsterdam, The Netherlands, ISBN: 0-7204-2356-2. (Cit. on p. 5.)

HUTSON, V. and J. S. PYM

- 1980 *Applications of Functional Analysis and Operator Theory*, Academic Press, London. (Cit. on p. 5.)

JORDAN, PASCUAL and JOHN VON NEUMANN

- 1935 “On inner products in linear, metric spaces”, *Annals of Mathematics*, 36, 3 [July 1935], pp. 719-723. (Cit. on p. 14.)

REED, MARTIN and BERRY SIMON

- 1980 *Functional Analysis*, 2nd ed., Methods of Modern Mathematical Physics, Academic Press, London, vol. 1, ISBN: 0-12-585050-6. (Cit. on p. 5.)

REZNICK, BRUCE

- 1978 “Banach spaces which satisfy linear identities.” *Pacific J. Math.*, 74, 1, pp. 221-233. (Cit. on p. 16.)

ROCKAFELLAR, TYRRELL R.

- 1970 *Convex Analysis*, Princeton University Press, Princeton, New Jersey, ISBN: 978-0-691-01586-6. (Cit. on p. 16.)

SCHOLTES, SEBASTIAN

- 2010 *A characterisation of inner product spaces by the maximal circumradius of spheres*, Feb. 3, 2010, arXiv: [1202.0503 \[math\]](#). (Cit. on p. 16.)

TESCHL, GERALD

- 2009 *Mathematical Methods in Quantum Mechanics with Applications to Schrödinger Operators*, Graduate Studies in Mathematics, AMS, Providence, Rhode Island, USA, vol. 99, ISBN: 978-0-8218-4660-5 (alk. paper), <http://www.ams.org/bookpages/gsm-99>. (Cit. on p. 5.)