Alessandro Candolini

Quantum Mechanics

A Non-Relativistic Primer

August 26, 2023

e-mail: 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 where 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 bacause they just like it. Young people, seduced by capitans forming crews to go to a Nuevo El Dorando [...] 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 become 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

There are all sort of books out there about non-relativistic quantum mechanics. Some focus on the physical theory, while others delve deeper into the mathematical intricacies or the philosophical aspects. Certain volumes span across various applications,Äîranging from condensed matter and computational physics to statistical mechanics and nuclear physics. A few lay the groundwork for eventual explorations into relativistic quantum field theory, whereas others shine a spotlight on quantum information, quantum computation, and even the unresolved queries encompassing the foundational interpretation of quantum field theory. This merely scratches the surface of the diversity present.

Attempting to providing a single, detailed, comprehesive account of all the ramification of quantum mechanics is a hopeless task. So, any attempt at writing some notes about quantum mechanics will have to make certain design choices about the selection of topics to be presented. Inevitably, this involves making subjective choices influenced by personal inclinations and a certain proclivity towards subjects in which the author believes to have better knowledge. Yet, beyond these considerations lies an additional layer.

London, August 26, 2023

Alessandro Candolini

Part I

THE BASICS

1 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

In comparisong to other areas of physics, quantum mechanics presents unique challenges. On one hand, quantum theory stands as a great reminder of how intuition can be misleading and untrustworthy. Specifically, the intuitions developed from classical physics have proven inadequate in providing the guidance necessary to grasp the complexities of the quantum world. On the other hand, the mathematical machinery required to adeguately formulate quantum mechanics carries substantial complexity since the beginning, which contributes to obfuscate the physical implications, or at least it represents a technical distraction from the main focus.

One might wonder: is this degree of mathematical sophistication required from the start? As it will become clearer soon, the inherent properties of quantum systems — such as Heisenberg's uncertanty principle, probabilistic outcome of quantum measurements, entanglement, and superposition of quantum states — demand alternative mathematical abstractions capable of accommodating such properties. The traditional mathematical tooling empowered by classical physics proven inadeguate to capture such phenomena, naturally driving the search for alternative formalisms.

Remarkably enough, quantum mechanics not only triggered a shift in perspective about physics, but also forced to revisit the *mathematical infrastructure* needed to organise, support, and encode properties and behaviours of quantum systems.

For example, as it will be discussed in depth, incompatibility between physical measurable quantities like position and momentum, which is at the root of uncetainty principle, can be described in terms of *non-commutative* objects, leading to adopt some sort of non-commuting mathematical entity like operators instead of real numbers to describe position and momentum. Similarly, the canonical commutation relations impose further constrains: as they can't be realised in finite-dimensional linear spaces, they force to work with infinite-dimensional spaces.

Is it possible to circumvent the use of Hilbert space? As we proceed towards the final chapters, alternative formulations will be provided, notably the path integral approach to quantum mechanics due to Richard Feynman. This approach has pros and cons, it has turned out to be somehow better suited for relativistic quantum field theories, and provides physical insights, but it still requires mathematical technicisms. More recently, the language of category theory has helped removing some of the noise of the Hilbert

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

space formulation, revealing the core of some of the underlying core concepts (e. g., the arguments of no-cloning theorem) at the price of more advanced mathematical formalisms.

In § 3, the mathematical principles of non-relativistic quantum mechanics are presented systematically using the language of Hilbert spaces. These postulates, crystallized during the 1920s and 1930s, emerged after few decades of intense and captivating endeavors by the pioneering minds of quantum mechanics. This period witnessed substantial advancements in theory and experimentation, intertwined with missteps, erroneous attempts, and an array of intellectual pursuits. Although contemporary students might readily adopt these postulates as a foundation to develop the theory, the historical context often obscures the physical motivations behind this formulation.

In this chapter, a complementary path to introduce quantum mechanics is discussed. This inspiring approach, due to Schwinger and presented in Schwinger [2001, § 1], supplements the content of § 3 by providing the physical intuitition to infer those postulates.

1.2 ALGEBRA OF MEASUREMENTS

REFERENCES FOR § 1

SCHWINGER, JULIAN

2001 *Quantum Mechanics. Symbolism of Atomic Measurements*, Springer-Verlag, Berlin, ISBN: 3-540-41408-8. (Cit. on pp. 3, 4.)

2 LINEAR OPERATORS IN HILBERT SPACES

The operator formulation of non-relativistic quantum mechanics is rooted in the spectral theory of linear self-ajoint operators (bound or unbound) in complex separable Hilbert spaces. This chapter provides the necessary mathematical background in functional analysis to approach non-relativistic quantum mechanics in full generality. Readers already acquainted with these concepts can jump to the next chapter to see how the mathematical theory is applied to quantum mechanics. Familiarity with basic linear algebra (eg, at the level of) and real analysis (e. g., at the level of) is assumed. Some of these notions are formalied in Agda in appendix C.

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 METRIC SPACES

Metric spaces are sets endowed with a notion of "distance" between every pair of elements in the set. Metric spaces provide a natural context where to develop concepts like continuity, convergence, and more.

DEFINITION 2.1 (distance): Let A be a set. A "distance" (or "metric") on A is any application $d\colon A \times A \to \mathbb{R}$, satisfying the following laws:

(a) $d(\varphi, \psi) \geqslant 0, \quad \forall(\varphi, \psi) \in A \times A$ (nonnegativity);

(b) $d(\varphi, \psi) = 0$ *if and only if* $\varphi = \psi$);

(c) $d(\varphi, \psi) = d(\psi, \varphi), \quad \forall(\varphi, \psi) \in A \times A$ (i. e., symmetry);

(d) $d(\varphi, \psi) \leqslant d(\psi, \eta) + d(\eta, \psi), \quad \forall(\varphi, \psi, \eta) \in A \times A \times A$ (the "triangle inequality").

DEFINITION 2.2 (metric space): A "metric space" is an ordered pair (A, d), where A is a set and $d\colon A \times A \to A$ is a distance on A. The elements of A are often referred to as "points" in this context.

Whether to include or not the empty set $\varnothing$ as part of the definition is a matter of convention. Here, definition 2.2 includes the empty case, but it will be uninteresting in what follows.

Given a metric space (A, d), every subset $U \subseteq A$ is a metric space with respect to the "restriction" of d to U.

EXERCISE 2.1 (discrete metric space): Every set A can be turned into a metric space by defining $d\colon A \times A \to \mathbb{R}$ as

$$d(\varphi, \psi) = \begin{cases} 1 & \text{if } \varphi = \psi \\ 0 & \text{otherwise} \end{cases}.$$

Prove that such d satisfies all the properties of a distance.

The next section deals with metric spaces where the distance arises from additional structure (i. e., norm, inner product). Without pretending to extensively develop the theory of metric spaces in this section, it's worth mentioning that notions such as continuity, convergence, and more can be formulated quite generally in the context of metric spaces.

DEFINITION 2.3 (continuity): Let (A, d) and (A', d') be metric spaces and $f\colon A \to A'$ a function. f is "continuous" at a point $\varphi \in A$ (with respect to the given metric) 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,

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall \psi \in A) \left(d(\varphi, \psi) < \delta \Rightarrow d'(f(\varphi), f(\psi)) < \varepsilon \right). \tag{2.1}$$

The notion of continuity can be generalised to functions between *topological spaces*, where there is generally no notion of distance. Equivalent characterisations of continuity in metric spaces will be presented shortly. Before, let's introduce some terminology.

DEFINITION 2.4 (open and closed ball): Let (A, d) be a metric space. Given $\rho > 0$ and $\varphi \in A$, the "open ball" $B_\rho(\varphi)$ and the "closed ball" $\overline{B}_\rho(\varphi)$ of "center" φ and "radius" ρ are the subset of A defined respectively by

$$B_\rho(\varphi) = \{ \psi \in A \mid d(\varphi, \psi) < \rho \}, \tag{2.2}$$

and

$$\overline{B}_\rho(\varphi) = \{ \psi \in A \mid d(\varphi, \psi) < \rho \}. \tag{2.3}$$

i. e., the subset of *all and only* the points of A at a distance respectively strictly less and less or equal to ρ from φ.

The terminology "open" and "closed" will become clear after we introduce the concept of open and closed metric spaces.

DEFINITION 2.5 (neighbourhood): Let (A, d) be a metric space. A "neighbourhood" of $\varphi \in A$ is any subset $U \subseteq A$ such that there exists an open ball entirely contained in U, i. e., there exist $\rho > 0$ such that $B_\rho(\varphi) \subseteq U$.

DEFINITION 2.6 (open subset): Let (A, d) be a metric space. A subset $U \subseteq A$ is "open" if every point of U has a neighborhood contained in U.

The following justifies the previous terminology for open balls.

PROPOSITION 2.1: Every open ball is open.

Proof. Let (A, d) be a metric space and $B_\rho(\varphi)$ an open ball. The goal is to prove that for every $\psi \in B_\rho(\varphi)$, there exists $\rho' > 0$ such that $B_{\rho'}(\psi) \subseteq B_\rho(\varphi)$. Set $\rho' := \rho - d(\varphi, \psi)$. Then, for every $\xi \in B_{\rho'}(\psi)$ the triangle inequality implies $d(\xi, \varphi) \leqslant d(\xi, \psi) + d(\psi, \varphi) < \rho - d(\varphi, \psi) + d(\psi, \varphi) = \rho$. ∎

THEOREM 2.1 (topology of a metric space): Let (A, d) be a metric space. Then

- A and $\varnothing$ are open;
- any finite, countable, or uncountable union of open subsets of A is open;
- Any *finite* intersection of open subsets of A is open.

Proof is left as exercise. Proof that countable intersection don't need to be open is also left as an exercise. Theorem 2.1 defines a structure of "topological space" on A, i.e., a ordered pair of (A, τ) where τ is a set of subsets of A closed under arbitrary unions and finite intersections.

DEFINITION 2.7 (accumulation and isolated point): Let (A, d) be a metric space. A point $\varphi \in A$ is an "accumulation" point of A if *every* neighbourhood of φ contains at least one point of A different from φ itself. Every $\varphi \in A$ that is *not* an accumulation point is a "isolated" point.

DEFINITION 2.8 (limit): Let (A, d) and (A', d') be metric spaces, $D \subseteq A$ a subset, $f\colon D \to A'$ a function on D and φ an accumulation point of D. An element $l \in A'$ is called the "limit" of f as ψ approaches φ, and it's written $\lim_{\psi\to\varphi} f$, if

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall \psi \in D)(\psi \neq \varphi)\left(d(\varphi, \psi) < \delta \Rightarrow d'(l, f(\psi)) < \varepsilon\right). \quad (2.4)$$

Notice: the function f does not need to be defined at φ, however φ needs to be an accumulation point of the domain of f.

Definition 2.8 might not look practically useful: if there's a candidate l, the definition determines unambiguously whether l is or is not the limit, but the identification of a target l needs to happen by other means. It's possible to develop a calculus by a) develop an "alphabet" of known limits for the simplest possible cases of practical relevance, relying solely on definition 2.8; b) develop a number of higher level "combinators" to reduce the calculation of limits of more complex expressions to some combination of the simplest known ones.

2.2 TOPOLOGY OF METRIC SPACES

2.2.1 *Completeness*

DEFINITION 2.9 (Cauchy sequence): Let (A, d) be a metric space. A sequence $(\varphi_n)_{n\in\mathbb{N}} : \mathbb{N} \to A$ is called a Cauchy sequence

PROPOSITION 2.2: Let (A, d) be a metric space. Any convergent sequence $(\varphi_n)_{n\in\mathbb{N}}$ in A is also a Cauchy sequence.

Intuitively, proposition 2.2 follows from the fact that φ_n and φ_m are both "close" to the limit of the sequence, so their distance is somehow bound. To make the argument rigorous, we will rely on the triangle inequality.

Proof. Let $\varphi = \lim_n \varphi_n$. By definition of convergence,

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})\,(n > N \Rightarrow d(\varphi, \varphi_n) < \epsilon)\,.$$

We want to prove that exists $M \in \mathbb{N}$ such that

$$(\forall \varepsilon > 0)(\exists M > 0)\,(n, m > M \Rightarrow d(\varphi_m, \varphi_n) < \epsilon)\,.$$

Take any $M > N$, then

$$d\,(\varphi_m, \varphi_n) \leqslant d\,(\varphi_m, \varphi) + d\,(\varphi, \varphi_n)\,,$$

i.e., $d(\varphi_m, \varphi_n) < \varepsilon + \varepsilon$. ∎

The converse of proposition 2.2 is not true in general, i.e., there exists metric spaces where some Cauchy sequences do not converge.

A typical counter-example works as follow. Let $(\varphi_n)_{n \geqslant 1}$ be a sequence in $\mathbb{R}$ such that for every $n \geqslant 1$ we have: $\varphi_n \in \mathbb{Q}$ and $\sqrt{2} - \frac{1}{n} < \varphi_n < \sqrt{2} + \frac{1}{n}$. (It is always possible to construct such sequence since $\mathbb{Q}$ is dense in $\mathbb{R}$.) It is possible to prove that $\lim_{n \to +\infty} \varphi_n = \sqrt{2}$. Being convergent in $\mathbb{R}$, φ_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 $\sqrt{2} \in \mathbb{R} \backslash \mathbb{Q}$.

Definition 2.10 (completeness): A metric space is "complete" if *every* Cauchy sequency is convergent.

2.2.2 *Compactness*

2.3 BANACH AND HILBER 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 2.11 (norm): Let X be any vector space over $\mathbb{K}$. A "norm" on X is any application $X \to \mathbb{R}$ hereafter denoted by $\|\cdot\|$ satisfying the following properties:

(a) $\|\varphi\| \geqslant 0, \quad \forall \varphi \in X$ (nonnegativity);

(b) $\|\varphi\| = 0$ if and only if $\varphi = 0$ (faithfulness)

(c) $\|\lambda\varphi\| = |\lambda| \, \|\varphi\|, \quad \forall \lambda \in \mathbb{K}$ and $\forall \varphi \in X$ (positive homogeneity);

(d) $\|\varphi + \psi\| \leqslant \|\varphi\| + \|\psi\|, \quad \forall (\varphi, \psi) \in X \times X$ (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 $\|\varphi + (-\varphi)\| \leqslant \|\varphi\| + \|\varphi\|$, and item (b) implies $\|\varphi - \varphi\| = 0$, so $\|\varphi\| \geqslant 0$.

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 2.12 (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 2.2: Let $(X, \|\cdot\|)$ be any normed vector space over $\mathbb{K}$. Let $d \colon X \times X \to \mathbb{R}$ be the function defined by

$$\boxed{d(\varphi, \psi) = \|\varphi - \psi\|,} \tag{2.5}$$

for very pair of vectors $(\varphi, \psi) \in X \times X$. Then, (X, d) is a metric space. The metric in theorem 2.2 is called the "metric induced by the norm" on X.

Proof. Let's show that d in theorem 2.2 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

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

This completes the proof. ∎

The metric induced by a norm fulfilles 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 2.2: Prove items 1 and 2 above.

Theorem 2.2 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 necessary related to the norm; furthermore, theorem 2.2 is not the only one possible distance built from the norm (see **??** 2.3).

Exercise 2.3: Let $(X, \|\cdot\|)$ be a normed vector space. Let $d\colon X \times X \to \mathbb{R}$ 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 2.2, 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:

$$\big|\|\psi\| - \|\varphi\|\big| \leqslant \|\psi - \varphi\|\,, \tag{2.6}$$

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

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

and

$$\|\psi\| = \|\psi - \varphi + \varphi\| \leqslant \|\varphi - \psi\| + \|\psi\|\,.$$

Thus,

$$\|\varphi\| - \|\psi\| \leqslant \|\varphi - \psi\|$$
$$\|\psi\| - \|\varphi\| \leqslant \|\varphi - \psi\| .$$

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 2.3: 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 $(\varphi_k)_{k\in\mathbb{N}}\colon \mathbb{N} \to X$ is a sequence in X convergent to some $\varphi \in X$, i. e.,

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

then

$$\|\varphi\| = \left\| \lim_{k\to+\infty} \varphi_k \right\| = \lim_{k\to+\infty} \|\varphi_k\| . \tag{2.7}$$

DEFINITION 2.13 (equivalence of the norms): Let X be a vector space over $\mathbb{K}$ and $\|\cdot\|_1 : X \to \mathbb{R}$ and $\|\cdot\|_2 : X \to \mathbb{R}$ two norms on X. The two norms are said to be "equivalent" if there exists a pair of strictly positive real numbers λ and μ such that

$$\begin{aligned} \alpha \|\varphi\|_1 &\leqslant \|\varphi\|_2 \\ &\leqslant \beta \|\varphi\|_1 , \end{aligned} \tag{2.8}$$

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.

Let us recall an important fact about Cauchy sequencies.

Banach space

DEFINITION 2.14 (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.

Complex numbers notation

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 $z = \Re z + {}_skmath_imaginary_unit : ni\Im z$ (where $\Re z$ and $\Im z$ are the real and imaginary parts of z, respectively), $z^* = \Re z - {}_skmath_imaginary_unit : ni\Im z$ 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 $|z_1 z_2| = |z_1|\,|z_2|$.

Inner product

DEFINITION 2.15 (inner product): Let X be a vector space over $\mathbb{K}$. A "inner product" on X is any application $X \times X \to \mathbb{K}$, hereafter denoted by $\langle\cdot|\cdot\rangle$, satisfying the following properties:

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

(c) $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*\,, \quad \forall(\varphi,\psi) \in X \times X$;

(d) $\langle\varphi|\varphi\rangle \geqslant 0, \quad \forall\varphi \in X$;

(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*

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

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

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

for all $(\varphi,\psi,\eta) \in X \times X \times X$ and for all $(\lambda,\mu) \in \mathbb{K} \times \mathbb{K}$. 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 *bi*linear); 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 convetion 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 $\langle\varphi|\varphi\rangle \geqslant 0$, since we expect $\langle\varphi|\varphi\rangle \in \mathbb{K}$, and if $\mathbb{K} = \mathbb{C}$ it might seem that the inequlity 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 (φ,ψ) instead of $\langle\varphi|\psi\rangle$; the notation $\langle\varphi|\psi\rangle$ is closer to the one uses 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 ψ 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 "braket"!). Of course, this is just a suggestive naming convention.

Inner product space

DEFINITION 2.16 (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 $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$ is an inner product on X.

The following lemma will be useful later on.

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

$$\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle = 0\,, \tag{2.9}$$

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 $\|\cdot\| : X \to \mathbb{R}$ defined by

$$\boxed{\|\psi\| = \sqrt{\langle\psi|\psi\rangle}, \quad \forall\psi \in X\,.} \tag{2.10}$$

Observe that such $\|\cdot\|$ in eq. () is well-defined, since $\langle\psi|\psi\rangle \geqslant 0$ for every $\psi \in X$. The square root is not ambigous: 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 justify 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.

Cauchy-Schwarz inequality

THEOREM 2.4 (Cauchy-Schwarz inequality): Let $(X, \langle\cdot|\cdot\rangle)$ be an inner product space. The following holds:

$$\boxed{|\langle\varphi|\psi\rangle| \leqslant \|\varphi\|\,\|\psi\|\,, \quad \forall(\varphi, \psi) \in X \times X\,.} \tag{2.11}$$

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

Remark. In theorem 2.4 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 $0 \leqslant 0$, which is satified.

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

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

From item (d) in definition 2.15, the left-hand side of this equation is positive:

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

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

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

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

$$\lambda = \frac{\langle\varphi|\psi\rangle^*}{\langle\psi|\psi\rangle}\,,\quad \psi \neq 0\,, \tag{2.12}$$

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

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

hence

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

from which it follows

$$\langle\varphi|\varphi\rangle\,\langle\psi|\psi\rangle \geqslant |\langle\varphi|\psi\rangle|$$

(using the fact that $\langle\psi|\psi\rangle > 0$). Taking the square root of both sites (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 2.4: For every fixed $(\varphi, \psi) \in X \times X$, consider $f\colon \mathbb{C} \to \mathbb{R}$ defined by

$$\begin{aligned}f(\lambda) &= \|\varphi - \lambda\psi\|^2 \\ &= |\lambda|^2\,\|\psi\|^2 - \lambda\,\langle\varphi|\psi\rangle - \lambda^*\,\langle\varphi|\psi\rangle^* + \|\varphi\|^2\,,\end{aligned} \tag{2.13}$$

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

Theorem 2.5: 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 φ, ψ in X,

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

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

THEOREM 2.6: Let $(X, \langle\cdot|\cdot\rangle)$ be an inner product space over $\mathbb{K}$. $\langle\cdot\rangle\cdot$ 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

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

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

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

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

then

$$\langle\varphi|\psi\rangle = \left\langle \varphi \,\middle|\, \lim_{n\to+\infty} \sum_{k=1}^{n} \psi_k \right\rangle = \lim_{n\to+\infty} \left\langle \varphi \,\middle|\, \sum_{k=1}^{n} \psi_k \right\rangle = \lim_{n\to+\infty} \sum_{k=1}^{n} \langle\varphi|\psi_k\rangle\,. \tag{2.14}$$

EXERCISE 2.5: Justify all steps in eq. ().

Hilbert space

DEFINITION 2.17 (Hilbert space): Let $(X, \langle\cdot|\cdot\rangle)$ be a 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”.

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 iner 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. On 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 a inner product space. In the case $\mathbb{K} = \mathbb{R}$,

$$\langle\varphi|\psi\rangle = \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2\right), \quad \forall(\varphi,\psi) \in X \times X. \tag{2.15}$$

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

$$\begin{aligned}\langle\varphi|\psi\rangle &= \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2\right)\\ &- \frac{\_{s}kmath_{i}maginary_{u}nit : ni}{4}\left(\|\varphi + \_{s}kmath_{i}maginary_{u}nit : ni\psi\|^2\right.\\ &\left.+ \|\varphi - \_{s}kmath_{i}maginary_{u}nit : ni\psi\|^2\right), \quad \forall(\varphi,\psi) \in X \times X.\end{aligned} \tag{2.16}$$

Remark. Equation (2.15) can be recasted in a shorter form:

$$\boxed{\begin{aligned}\langle\varphi|\psi\rangle = \frac{1}{4}\sum_{k=0}^{3}(-\_{s}kmath_{i}maginary_{u}nit : ni)^k \Big\|\varphi\\ + \_{s}kmath_{i}maginary_{u}nit : ni^k\psi\Big\|^2.\end{aligned}} \tag{2.17}$$

Remark. Different signs are possibile in literature if a different convention is chosen for property (b) of definition 2.15, 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,

$$\|\varphi+\psi\|^2 = \|\varphi\|^2 + 2\Re\langle\varphi|\psi\rangle + \|\psi\|^2 \tag{2.18}$$

$$\|\varphi-\psi\|^2 = \|\varphi\|^2 - 2\Re\langle\varphi|\psi\rangle + \|\psi\|^2 \tag{2.19}$$

$$\|\varphi + \_{s}kmath_{i}maginary_{u}nit : ni\psi\| = \|\varphi\|^2 - 2\Im\langle\varphi|\psi\rangle + \|\psi\|^2 \tag{2.20}$$

$$\|\varphi - \_{s}kmath_{i}maginary_{u}nit : ni\psi\| = \|\varphi\|^2 + 2\Im\langle\varphi|\psi\rangle + \|\psi\|^2 \tag{2.21}$$

Subtracting eqs. (2.17) and (2.18) we get

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

subtracting eqs. (2.19) and (2.20) we get

$$\|\varphi + {}_s kmath_i maginary_u nit : ni\psi\| \\ - \|\varphi - {}_s kmath_i maginary_u nit : ni\psi\| = -4\Im\langle\varphi|\psi\rangle\,.$$

Thus

$$\frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2\right) \\ -\frac{{}_s kmath_i maginary_u nit\ : ni}{4}(\|\varphi + {}_s kmath_i maginary_u nit : ni\psi\| \\ - \|\varphi - {}_s kmath_i maginary_u nit : ni\psi\|) = \Re\langle\varphi|\psi\rangle \\ + {}_s kmath_i maginary_u nit \\ : ni\Im\langle\varphi|\psi\rangle \\ = \langle\varphi|\psi\rangle\,.$$

This completes the proof. ∎

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 2.7 (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:

$$\|\varphi+\psi\|^2 + \|\varphi-\psi\|^2 = 2\,\|\varphi\|^2 + \|\psi\|^2, \quad \forall(\varphi,\psi)\in X\times X\,. \qquad (2.22)$$

Theorem 2.7 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 2.7. This is the straightforward part of the proof. By definition of the norm induced by the inner product,

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

for all $(\varphi,\psi)\in X\times X$; summing the two we have theorem 2.7.

Now, let's prove the converse: if $(X, \|\cdot\|)$ is a Banach space whose norm satisfies theorem 2.7, 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 $\langle\cdot|\cdot\rangle : X\times X \to \mathbb{K}$ be the application defined by

$$\langle\varphi|\psi\rangle = \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2\right) \\ -\frac{{}_s kmath_i maginary_u nit : ni}{4}\left(\|\varphi + {}_s kmath_i maginary_u nit : ni\psi\|^2 \\ - \|\varphi - {}_s kmath_i maginary_u nit : ni\psi\|^2\right),$$

[*] 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 2.15. 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

$$2\|\varphi+\psi\|^2 + 2\|\eta\|^2 = \|\varphi+\psi+\eta\|^2 + \|\varphi+\psi-\eta\|^2, \tag{2.23a}$$
$$2\|\varphi-\psi\|^2 + 2\|\eta\|^2 = \|\varphi-\psi+\eta\|^2 + \|\varphi-\psi-\eta\|^2, \tag{2.23b}$$
$$2\|\varphi+\eta\|^2 + 2\|\psi\|^2 = \|\varphi+\psi+\eta\|^2 + \|\varphi-\psi+\eta\|^2, \tag{2.23c}$$
$$2\|\varphi-\eta\|^2 + 2\|\psi\|^2 = \|\varphi+\psi-\eta\|^2 + \|\varphi-\psi-\eta\|^2. \tag{2.23d}$$

Subtracting eqs. (2.23a) and (2.23) we get

$$\begin{aligned} 2\|\varphi+\psi\|^2 - 2\|\varphi-\psi\|^2 = \|\varphi+\psi+\eta\|^2 + \|\varphi+\psi-\eta\|^2 \\ - \|\varphi-\psi+\eta\|^2 - \|\varphi-\psi-\eta\|^2; \end{aligned} \tag{2.24}$$

subtracting eqs. (2.23b) and (2.23c) we get

$$\begin{aligned} 2\|\varphi+\eta\|^2 - 2\|\varphi-\eta\|^2 = \|\varphi+\psi+\eta\|^2 + \|\varphi-\psi+\eta\|^2 \\ - \|\varphi+\psi-\eta\|^2 - \|\varphi-\psi-\eta\|^2; \end{aligned} \tag{2.25}$$

summing eqs. (2.23d) and (2.24) we get

$$\begin{aligned} &2\|\varphi+\psi\|^2 - 2\|\varphi-\psi\|^2 + 2\|\varphi+\eta\|^2 - 2\|\varphi-\eta\|^2 \\ &= \|\varphi+\psi+\eta\|^2 + \|\varphi+\psi-\eta\|^2 - \|\varphi-\psi+\eta\|^2 - \|\varphi-\psi-\eta\|^2 \\ &\quad + \|\varphi+\psi+\eta\|^2 + \|\varphi-\psi+\eta\|^2 - \|\varphi+\psi-\eta\|^2 - \|\varphi-\psi-\eta\|^2 \\ &= 2\|\varphi+\psi+\eta\|^2 - 2\|\varphi-\psi-\eta\|^2. \end{aligned} \tag{2.26}$$

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 $z \in \mathbb{Z}$, $w \in \mathbb{Z} \backslash \{0\}$,

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

yielding

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

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 th 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

$$\langle \varphi | \_s kmath_i maginary_u nit : ni\psi \rangle = \_s kmath_i maginary_u nit : ni \langle \varphi | \psi \rangle,$$

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

Ad (c). This is straightforward:

$$\langle\psi|\varphi\rangle = \frac{1}{4}\Big\{\|\psi+\varphi\|^{2} - \|\psi-\varphi\|^{2}$$

$$-_s kmath_i maginary_u nit:ni \Big(\|\psi+_s kmath_i maginary_u nit:ni\varphi\|^{2} - \|\psi-_s kmath_i magin$$

$$-\|\varphi-\psi\|^{2}\Big)$$

$$+\frac{_s kmath_i maginary_u nit:ni}{4}\Big(\|\varphi+_s kmath_i maginary_u nit:ni\psi\|^{2}\,\|\varphi-_s kmath_i magina$$

Ad (d). Also this is straightforward:

$$\langle\psi|\psi\rangle = \frac{1}{4}\Big\{\|\psi+\psi\|^{2} - \|\psi-\psi\|^{2}$$

$$-_s kmath_i maginary_u nit:ni \Big(\|\psi+_s kmath_i maginary_u nit:ni\psi\|^{2} - \|\psi-_s kmath_i magin$$

$$-_s kmath_i maginary_u nit:ni \Big(|1+_s kmath_i maginary_u nit:ni|^{2}\,\|\psi\|^{2} - |1-_s kmath_i magin$$

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 2.7 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.

2.4 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 2.18 (Convex set): Let X be a vector space over $\mathbb{K}$ and $\Omega \subseteq X$ a non-empty subset of X. Ω is "convex" if for every $(\varphi,\psi) \in \Omega\times\Omega$,

$$\vartheta\varphi + (1-\vartheta)\psi \in \Omega\,, \tag{2.27}$$

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

Remark. Definition 2.18 is a "weighted average" of φ and ψ.

Remark. In order definition 2.18 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.

† Of course, if $\vartheta = 0$ we get ψ and if $\vartheta = 1$ we get φ, so we could have considered equivalently $]0,1[$ instead of $[0,1]$ in definition 2.18.

reason why in definition 2.18 X is a vector space over $\mathbb{R}$ or $\mathbb{C}$. The notion of convexity may be generalized to more general settings if certain propeties 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 2.19 (line segment): Let X be a vector space over $\mathbb{K}$ and consider $(\varphi, \psi) \in X \times X$. The set

Line segment

$$[\varphi, \psi] = \{\vartheta\varphi + (1-\vartheta)\psi \mid \forall\vartheta \in [0, 1]\} \tag{2.28}$$

is called the "(straight) line segment" joining φ and ψ (or with endpoints φ and ψ).

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

Using definition 2.19 we can restate definition 2.18 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. Ω 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 2.6: 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 2.20 (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 $\varphi \leqslant \eta \leqslant \psi$ 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 $\mathbb{R}$. Ω is convex *if and only if* Ω is a real interval.

THEOREM 2.8 (Young inequality): For every $A, B \geqslant 0$ and $0 \leqslant \vartheta \leqslant 1$,

Young inequality

$$\boxed{A^{\vartheta} B^{1-\vartheta} \leqslant \vartheta A + (1-\vartheta) B\,,} \tag{2.29}$$

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

$$\sqrt{AB} \leqslant \frac{A+B}{2}, \quad A, B \geqslant 0\,. \tag{2.30}$$

The general aithmetic-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 rearraning the square terms:

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

yielding

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

If $A, B \geqslant 0$ we can take the square root and find eq. (2.29). 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. (2.29): 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 $4\sqrt{AB} \leqslant 2A + 2B$ 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 2.8 reads

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}. \tag{2.31}$$

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

$$\exp \vartheta\varphi + (1-\vartheta)\psi \leqslant \vartheta \exp \varphi + (1-\vartheta) \exp \psi \tag{2.32}$$

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 2.8. ∎

Another possible proof goes as follows. Consider $f_\vartheta \colon]0, +\infty[\to \mathbb{R}$ 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 $f_\vartheta(x) \leqslant 0$.

2.5 CANONICAL PROTOTYPES OF BANACH AND HILBERT SPACES

2.5.1 *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 $\langle\cdot|\cdot\rangle : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}$ as follows:

$$\boxed{\langle\vec{\varphi}|\vec{\psi}\rangle = \sum_{k=1}^{N} \varphi_k^* \psi_k \,.} \tag{2.33}$$

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

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

Proof. It is straightforward to show that $\langle\cdot|\cdot\rangle$ defined in eq. () satisfies Items (a) to (e) of definition 2.15. ∎

THEOREM 2.10: $(\mathbb{C}^N, \langle\cdot|\cdot\rangle)$, where $\langle\cdot|\cdot\rangle$ is defined by eq. (), is an Hilbert space.

Proof. We need to check completeness. ∎

2.5.2 *The Hilbert space l^2*

Let l^2 be the space of all and only the sequences $(\varphi_k)_{k\geqslant 1}$ 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 $\mathbb{K}$, l^2 is a vector space over $\mathbb{K}$. Define

$$\boxed{\langle\varphi|\psi\rangle = \sum_{k=1}^{+\infty} \varphi_k^* \psi_k \,,} \tag{2.34}$$

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

THEOREM 2.11: $(l^2, \langle\cdot|\cdot\rangle)$, where $\langle\cdot|\cdot\rangle$ is defined by section 2.5.2, 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 2.12: $(l^2, \langle\cdot|\cdot\rangle)$, where $\langle\cdot|\cdot\rangle$ is defined by section 2.5.2, is an Hilbert space.

2.5.3 *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 $\tilde{L}^2(\Omega)$ be the set consisting of all and only the functions defined on Ω and taking values on $\mathbb{K}$, i. e., $\psi\colon \Omega \to \mathbb{K}$, such that their square modulus is Lebesgue-integrable over Ω, 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 techical 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:

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

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,

$$0 \leqslant \int_\Omega |\varphi + \psi|^2 \leqslant \int_\Omega |\psi|^2 + \int_\Omega |\varphi|^2 = 0 ,$$

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 Ω) and thus $\psi - \psi \in M$;

SYMMETRY : for every ψ and φ 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 ψ, φ and η 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 $\eta - \psi \in M$, then also $(\psi - \eta) + (\eta - \varphi) \in M$ by lemma 2.6.

$\blacksquare$

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

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

Proof. There exists η and ξ 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 $\lambda\tilde{\psi} \sim \lambda\psi$. $\blacksquare$

Theorem 2.13: $(L^2(\Omega), +, \cdot)$ 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 2.13 as exercise.

Exercise 2.7: Prove theorem 2.13. 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

$$\boxed{\langle[\varphi]\,|\,[\psi]\rangle = \int_\Omega \varphi^*\psi\,,} \tag{2.35}$$

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

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

Remark. Given $\varphi\colon \Omega \to \mathbb{K}$, the function $\varphi^*\colon \Omega \to \mathbb{K}$ 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 2.14: Let $\varphi\colon \Omega \to \mathbb{K}$ be Lebesgue integrable on Ω and let $\psi\colon \Omega \to \mathbb{K}$ be any function satisfying

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

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

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

$$|zw| \leqslant \frac{1}{2}|z|^2 + \frac{1}{2}|w|^2 ; \tag{2.36}$$

in fact,

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

Thus,

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

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 2.8: 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 2.15: $(L^2(\Omega), \langle\cdot|\cdot\rangle)$, 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 2.15.

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

$$\langle\varphi|\psi + \eta\rangle = \int_\Omega \varphi^* (\psi + \eta) = \int_\Omega \varphi^*\psi + \int_\Omega \varphi^*\eta = \langle\varphi|\psi\rangle + \langle\varphi|\eta\rangle$$
$$\langle\varphi|\lambda\psi\rangle = \int_\Omega \varphi^* (\lambda\psi) = \lambda \int_\Omega \varphi^*\psi = \lambda \langle\varphi|\psi\rangle$$

Furthrmore,

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

The fact that complex conjugation can be moved outside the integration symbol is rigorously justified as follows. Let $\psi\colon \Omega \to \mathbb{C}$ be a complex-valued function defined on Ω, and set $\psi_1, \psi_2\colon \Omega \to \mathbb{R}$ by letting $\psi_1 = \Re\psi$ and $\psi_2 = \Im\psi$. By definition, ψ is Lebesgue integrable on Ω if and only if ψ_1 and ψ_2 are Lebesgue integrable on Ω and in that case we put

$$\int_\Omega \psi = \int_\Omega \psi_1 + _s kmath_i maginary_u nit$$
$$: ni \int_\Omega \psi_2$$

Clearly, if ψ is Lebesgue integrable, ψ^* is Lebesgue integrable and

$$\int_\Omega \psi^* = \int_\Omega \psi_1 - _s kmath_i maginary_u nit$$
$$: ni \int_\Omega \psi_2$$
$$= \left(\int_\Omega \psi\right)^* .$$

Finally, the monotonicity of the integral ensures that

$$\langle\psi|\psi\rangle = \int_\Omega \psi^*\psi = \int_\Omega |\psi|^2 \geqslant 0 ,$$

and $\int_\omega |\psi|^2 = 0$ if and only if ψ belongs to the class of equivalence of the function which is identically zero on Ω, i. e., $\psi = 0$. ∎

THEOREM 2.16 (Riesz-Fischer): $(L^2(\Omega), \langle\cdot|\cdot\rangle)$ is an Hilbert space.

Proof. We need to check completeness. ∎

2.5.4 *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 2.17: 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 $\varphi = (\varphi_n)_{n\geqslant 0}$ and $\psi = (\psi_n)_{n\geqslant 0}$ 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

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

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

2.6 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 extend to export the notion of orthogonality to general Banach spaces, but here we will restrict ourselves to Hilbert spaces.

Orthogonality of two vectors

DEFINITION 2.21: Let $(X, \langle\cdot,\cdot\rangle)$ be any inner product space. For every $(\varphi, \psi) \in X \times X$, φ and ψ are said to be mutually "orthogonal" if

$$\langle\varphi|\psi\rangle = 0\,. \tag{2.37}$$

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 2.22: 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

$$\langle\varphi|\psi\rangle = 0, \quad \forall(\varphi,\psi) \in W \times W \text{ with } \varphi \neq \psi\,. \tag{2.38}$$

If moreover

$$\langle\varphi\rangle = 1, \quad \forall\varphi \in W \tag{2.39}$$

the set W is said to be "orthonormal".

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

THEOREM 2.18 (Pythagorean theorem): Let $(X, \langle\cdot,\cdot\rangle)$.

1. for every orthogonal pair of vectors φ and ψ 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 $\mathbb{N} \to X$ 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.

■

2.7 COMPLETE ORTHONORMAL SETS AND THEIR CHARACTERIZATION

DEFINITION 2.23: 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 $\psi = 0$.

2.8 THE FOURIES SERIES IN THE SPACE OF SQUARE INTEGRABLE FUNCTIONS

2.9 LINEAR OPERATORS: FIRST PROPERTIES

2.10 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. 16.)

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. 18.)

Rockafellar, Tyrrell R.

1970 *Convex Analysis*, Princeton University Press, Princeton, New Jersey, isbn: 978-0-691-01586-6. (Cit. on p. 18.)

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. 18.)

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.)

3 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*

The Hilbert space approach to quantum mechanics is due to von Neumann [1955]. Our presentation is inspired by a number of other sources.

REFERENCES FOR § 3

VON NEUMANN, JOHN

1955 *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton, New Jersey, ISBN: 0-691-08003-8. (Cit. on p. 29.)

4 ONE-DIMENSIONAL SINGLE PARTICLE SPINLESS QUANTUM SYSTEMS

Part II

THE CORE

5 SYMMETRIES

6 ANGULAR MOMENTUM

7

HYDROGEN ATOM

8 PERTURBATION THEORY

9 SCATTERING

Part III

SELECTION OF ADVANCED TOPICS

10

ENTANGLEMENT

11

FEYNMAN'S PATH INTEGRALS

12 SEMICLASSICAL QUANTUM MECHANICS

13 SUPERSYMMETRIC QUANTUM MECHANICS

14 SECOND QUANTIZATION FORMALISM

15 A TASTE OF CATEGORICAL QUANTUM MECHANICS

Part IV

APPENDICES

APPENDIX A

B GLIMPSE INTO CATEGORY THEORY

C

FORMALISING ELEMENTS OF FUNCTIONAL ANALYSIS IN AGDA

BIBLIOGRAPHY

REFERENCES FOR § 1

SCHWINGER, JULIAN

2001 *Quantum Mechanics. Symbolism of Atomic Measurements*, Springer-Verlag, Berlin, ISBN: 3-540-41408-8. (Cit. on pp. 3, 4.)

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. 16.)

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. 18.)

ROCKAFELLAR, TYRRELL R.

1970 *Convex Analysis*, Princeton University Press, Princeton, New Jersey, ISBN: 978-0-691-01586-6. (Cit. on p. 18.)

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. 18.)

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.)

REFERENCES FOR § 3

VON NEUMANN, JOHN

1955 *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton, New Jersey, ISBN: 0-691-08003-8. (Cit. on p. 29.)