Mathematical Foundations of Quantum Physics

Doctorandus Adrian Stan

Zernike Institute for Advanced Materials
– Groningen, The Netherlands –

Han sur Lesse - Winterschool
04-08/12/2006

Contents

1 Clarifications **5**
1.1 What do we learn here? 5
1.2 Purpose 6

2 The Hilbert space **7**
2.1 Some basic definitions and clarifications 8
2.1.1 Topology 8
2.1.2 Vectors in $\mathbf{R}^n$ 10
2.1.3 A brief reminder of complex numbers. 12
2.1.4 Vectors in $\mathbf{C}^n$ 14
2.1.5 The Metric Space 16
2.2 The Hilbert Space 17
2.3 Orthonormal Bases in the Hilbert Space 19
2.4 The $\mathfrak{L}^2$ space 24
2.5 Linear and Antilinear Functionals. The Generalized Scalar Product. . . 26
2.6 Fourier Transforms, ρ space and Temperate Distributions 31
2.6.1 The Fourier transform and the ρ space 31
2.6.2 Temperate distributions ($\mathfrak{D}$) 32
2.6.3 The derivative and the Fourier transform of a distribution . . 33
2.7 The States Space 36
2.8 Summary of the chapter 37

3 Linear Operators and Basic Elements of Spectral Theory. **39**
3.1 Linear Operators 39
3.2 Orthogonal Projectors 45
3.3 Matrix Operators 49
3.4 The Eigenvalue Problem for Operators 52
3.4.1 The Eigenvalue Problem for Bounded Operators 54

3.4.2 The Eigenvalue Problem for Unbounded Operators 57

Chapter 1

Clarifications

1.1 What do we learn here?

I was asked what is the difference between "Quantum Mechanics" and "Quantum Physics" and why do I insist calling this course *Mathematical Foundations of Quantum Physics*?

Quantum physics is the set of quantum theories:

- *quantum mechanics* - a first quantized or semi-classical theory in which particle properties are quantized, but not particle numbers, fields and fundamental interactions.

- *quantum field theory* or QFT - a second or canonically quantized theory in which all aspects of particles, fields and interactions are quantized, with the exception of gravitation. Quantum electrodynamics, quantum chromodynamics and electroweak theory are examples of relativistic fundamental QFTs which taken together form the Standard Model. Solid state physics is a non-fundamental QFT.

- *quantum gravity* - a third quantized theory in which general relativity, *i.e.*, the theory of the gravitational force, is also quantized. In spite of the monumental effort such a theory remains outside the grasp of human knowledge.

A first quantization of a physical system is a semi-classical treatment of quantum mechanics, in which particles or physical objects are treated using quantum wave functions, but the surrounding environment (for example a potential well or a bulk electromagnetic field or gravitational field) is treated classically. First quantization is appropriate for studying a single quantum-mechanical system being controlled by a laboratory apparatus that is itself large enough that classical mechanics is applicable to most of the apparatus. This flavor of quantum mechanics is the subject studied in most undergraduate quantum mechanics courses, and in which the Schrödinger equation and Heisenberg matrix mechanics (together with bra-ket notation) are most simply applied. It may be contrasted with the so called second quantization, which includes quantum-mechanical uncertainty effects in all aspects of an experiment including the controlling fields and boundary conditions. That is to say, the system is no longer isolated but in interaction with the environment [1].

1.2 Purpose

The "mathematical foundations" in this course are mainly the mathematical foundations of the first two "elements" of the "quantum physics set", with a definite accent on the first. The purpose is to give the reader an idea of the underlying mathematical structure of the theory. Some proofs have been omitted because their presence would not do any good at this level – their result being much too intuitive – or because they are too complex for the purpose of this course. If at the end of this course, the reader has acquired a general, intuitive image of how the mathematical objects are interconnected in order to substantiate the Dirac formalism, then the very purpose of this course will be achieved.

The exercises in this course are intentionally easy ones since their role is to show to the reader, from time to time, the underlying grounds – which he already knows from basic quantum mechanics – of the abstract mathematics. Where physical intuition comes to the rescue, the exercises are also missing.

[1] This can be seen in the symmetrization of the wave function.

Chapter 2

The Hilbert space

The rapid development of quantum mechanics asked for the development of an underlying mathematical structure. Although they were moments when, because of the rapid development, a not so rigorous formalism was used, this formalism was always set up later and rigorously proved from the mathematical point of view.

The first synthesis was realized by John von Neumann, by developing the operators theory in Hilbert's space. A Hilbert space is a generalization of the idea of a vector space that is not restricted to finite dimensions. Thus it is an inner product space, which means that it has notions of distance and of angle - especially the notion of orthogonality. Moreover, it satisfies a completeness requirement which ensures that limits exist when expected. Even if this theory is perfectly true, von Neumann's approach does not consider more general spaces - like the distribution spaces - which, even if they are not directly implicated in the interpretation of the theory, they cannot be ignored if one wants to understand the subtile points of the mathematical formulation. This larger frame of quantum mechanics, which combines the Hilbert space with the theory of distributions, was created a bit later, by the russian mathematician Israel Moiseevich Gelfand (b. 1913). He introduced the famous *rigged Hilbert space*, or *the Gelfand triplet* [1].

What is, in fact, the rigged Hilbert space? A rigged Hilbert space is a set which

- has an algebraic structure of a linear space
- it is equipped with a nuclear topology with respect to witch the completion

[1] At the time this course was first given, Israel Moiseevich Gelfand was still alive. He died nearly three years later, at the age of 96, near his home in Highland Park, New Jersey.

gives Φ'

- has a second topology introduced into it by a scalar product, with respect to which this linear space is completed to give a Hilbert space H
- has a third topology, the topology of the dual space Φ' of Φ

In this chapter, we will set up the mathematical foundations of the day-by-day Dirac formalism and, for that, we will go as low as in topology, in order to settle, with a relatively high rigorousity and with the hope the reader will feel challenged to prove what we leave unproven, these mathematical foundations.

2.1 Some basic definitions and clarifications

We start with a rough map of mathematical physics, let's say. We will follow, in class, this map and fill in some of the missing arrows.

2.1.1 Topology

In this section I have defined a few fundamental concepts of topology. Because I will usually explain the terms as they will come in the way, I have inserted most of the definitions in the text or in the footnotes. These, on the other hand, did not fit anywhere, so here they are. We will reefer to them during the lectures.

Let X be any *set* and let $p(X) = \{Y | Y \subset X\}$ be the set of all subsets of X. A subset $\mathfrak{T}$ of $p(X)$ is called *a topology* of X iff (*i.e. "if and only if"*) the following conditions are fulfilled:

- Both the empty set, $\emptyset \in \mathfrak{T}$, and X are elements of $\mathfrak{T}$.
- The union of arbitrary many elements of $\mathfrak{T}$ is an element of $\mathfrak{T}$.
- Any intersection of a finite number of elements of $\mathfrak{T}$ is an element of $\mathfrak{T}$.

If $\mathfrak{T}$ is a topology on X, then X together with $\mathfrak{T}$ (also denoted, sometimes, by the pair $(X, \mathfrak{T})$ is called a *topological space*. A set in $\mathfrak{T}$ is called *open set*.

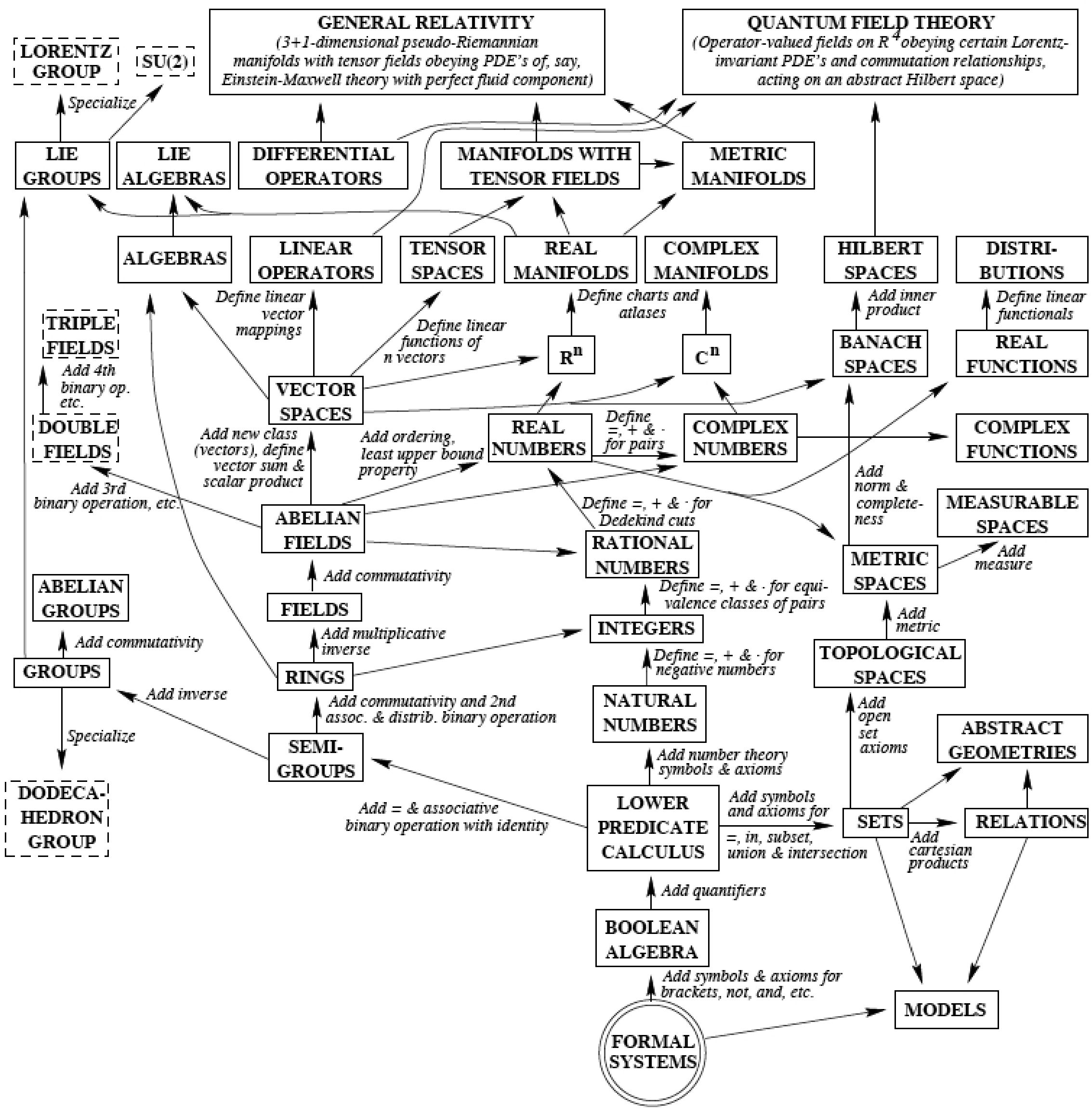
A sequence of points $a_1, a_2, \dots, a_n \dots \in X$ is said to *converge* to $a \in X$ if for every open set O with $a \in O$ there exists an integer $N(O)$ such that $a_n \in O \; \forall \; n > N(O)$

Exercise

Let X be a set and $p(X) = \{Y | Y \subset X\}$ the set of all subsets of X and $Z = \{\phi, X\}$. Prove that:

1) $p(X)$ is a topology of X

2) Z is a topology on X



2.1.2 Vectors in $\mathbf{R}^n$

We will assume the reader is familiar with the space $\mathbf{R}^3$, where all the points in space are represented by ordered triplets of real numbers.

The set of all n-tuples of real numbers, denoted by $\mathbf{R}^n$, is called *n-space*. A particular n-tuple in $\mathbf{R}^n$

$$\mathbf{u} = (a_1, a_2, \ldots, a_n)$$

is called *vector*. The numbers a_i are called *coordinates, components, entries* or *elements* of $\mathbf{u}$. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be equal if they have the same number of components and if the components are equal.

Consider two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$, with $\mathbf{u} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{v} = (b_1, b_2, \ldots, b_n)$. Their *sum*, written $\mathbf{u}+\mathbf{v}$, is the vector is the vector obtained by adding corresponding components from $\mathbf{u}$ and $\mathbf{v}$. That is:

$$\mathbf{u} + \mathbf{v} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

.

The *scalar product* between a vector and a scalar from $\mathbf{R}^n$, is defined like

$$k\mathbf{u} = k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$$

.

The sum of vectors with different number of components is not defined. The *negatives* and the *subtractions*, called *differences*, are defined in $\mathbf{R}^n$ an follows:

$$-\mathbf{u} = (-1)\mathbf{u}; \quad \mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

. For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{R}^n$ and any scalars k, k' in $\mathbf{R}$ we have:

- (i) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$,
- (ii) $(\mathbf{u} + \mathbf{0}) = \mathbf{u}$,
- (iii) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$,
- (iv) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,
- (v) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$,
- (vi) $(k + k')\mathbf{u} = k\mathbf{u} + k'\mathbf{v}$,
- (vii) $(kk')\mathbf{u} = k(k'\mathbf{u})$,

- (viii) $1\mathbf{u} = \mathbf{u}$.

If u and v are vectors in $\mathbf{R}^n$ for which $u = kv$ for some nonzero scalar $k \in \mathbf{R}$, then u is called a *multiple* of v. Also, u is said to be the *same* or *opposite direction* as v, according as $k > 0$ or $k < 0$.

Consider two arbitrary vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$; say:

$$\mathbf{u} = (a_1, a_2, \ldots, a_n) \quad \mathbf{v} = (b_1, b_2, \ldots, b_n)$$

. The *dot product* or *inner product* or *scalar product* of $\mathbf{u}$ and $\mathbf{v}$ is denoted and defined by

$$\mathbf{u} \cdot \mathbf{v} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

That is, $\mathbf{u} \cdot \mathbf{v}$ is obtained by multiplying corresponding components and adding the result products. The vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be *orthogonal* (or *perpendicular* if their dot product is zero, that is $\mathbf{u} \cdot \mathbf{v} = 0$

For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{R}^n$ and a scalar $k \in \mathbf{R}$:

- (i) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$,
- (ii) $(k\mathbf{u}) \cdot \mathbf{v} = k(\mathbf{u} \cdot \mathbf{v})$,[2]
- (iii) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$,
- (iv) $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = 0$

The space $\mathbf{R}^n$ with the above operations of vector addition, scalar multiplication and dot product is usually called *Euclidean n-space.*

The *norm* or *length* of a vector $\mathbf{u} \in \mathbf{R}^n$, denoted by $\|u\|$, is a real number that represents the "size" of the vector, defined to be the nonnegative square root of $\mathbf{u} \cdot \mathbf{u}$. In particular, if $\mathbf{u} = (a_1, a_2, \ldots, a_n)$, then

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}$$

A vector $\mathbf{u}$ is called *unit* vector if $\|u\| = 1$ or, equivalently, if $\mathbf{u} \cdot \mathbf{u} = 1$. For any nonzero vector $\mathbf{v} \in \mathbf{R}^n$, the vector

$$\hat{\mathbf{v}} = \frac{1}{\|v\|}\mathbf{v} \tag{2.1}$$

[2]Note that: $\mathbf{u} \cdot (k\mathbf{v}) = (k\mathbf{v}) \cdot \mathbf{u} = k(\mathbf{v} \cdot \mathbf{u}) = k(\mathbf{u} \cdot \mathbf{v})$. In other words, we can "take k out" also from the second position of the inner product.

is the unique unit vector in the same direction as $\mathbf{v}$. The process of finding $\hat{\mathbf{v}}$ from $\mathbf{v}$ is called *normalizing* $\mathbf{v}$.

The *Schwarz and Minkowski inequalities* are two of the most important ingredients in different branches of mathematics.

The Schwarz inequality: For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$, we have:

$$|u \cdot v| \leq \|u\| \|v\| \tag{2.2}$$

Proof: $\forall\, t \in \mathbf{R}$

$$0 \leq (t\mathbf{u}+\mathbf{v}) \cdot (t\mathbf{u}+\mathbf{v}) = t^2(\mathbf{u}\cdot\mathbf{u}) + 2t(\mathbf{u}+\mathbf{v}) + (\mathbf{v}\cdot\mathbf{v}) = \|\mathbf{u}\|^2 t^2 + 2(\mathbf{u}\cdot\mathbf{v})t + \|\mathbf{v}\|^2.$$

Let $a = \|\mathbf{u}\|^2, b = 2(\mathbf{u}\cdot\mathbf{v})$ and $c = \|v\|^2$. Then, for every value of t, $at^2+bt+c \geq= 0$. This means that the quadric polynomial cannot have two real roots, which implies a discriminant $D = b^2 - 4ac \leq 0$ or equivalently $b^2 \leq 4ac$. Substituting and dividing by 4 we have the inequality proven.

The Minkowski inequality: For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$, we have:

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \tag{2.3}$$

Exercises

1. Let $\mathbf{u} = (2, 4, -5)$ and $\mathbf{v} = (1, -6, 9)$. Compute: $\mathbf{u}+\mathbf{v}$, $7\mathbf{u}$, $-\mathbf{v}$, $3\mathbf{u} - 5\mathbf{v}$ and $(6\mathbf{v} - \mathbf{u}) - 3\mathbf{v}$.

2. Let $\mathbf{u} = (1, -2, 3), \mathbf{v} = (2, 7, 4)$ and $\mathbf{w} = (4, 5, -1)$. Find which two vectors are orthogonal.

3. Suppose $\mathbf{u} = (1, 2, 3, 4)$ and $\mathbf{v} = (6, k, -8, 2)$. Find k so $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

4. Let $\mathbf{u} = (1, -2, -4, 5, 3)$. Find $\|u\|$ and normalize $\mathbf{u}$.

5. Using the Schwarz inequality and the properties of the scalar product, prove the Minkovski inequality.

2.1.3 A brief reminder of complex numbers.

Any complex number can be written in the form

$$z = a + bi \tag{2.4}$$

where $a \equiv \operatorname{Re} z$ and $b \equiv \operatorname{Im} z$ are, respectively, the *real* and *imaginary parts* of z, and $i = \sqrt{-1}$. The mathematical operations are structured as follows: $\forall\, z, w \in$

C where $z = a + bi$, $w = c + di$

$$z + w = (a + bi) + (c + di) = a + c + bi + di = (a + c) + (b + d)i \tag{2.5}$$

$$zw = (a + bi)(c + di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i \tag{2.6}$$

$$-z = -1z \text{ with } w - z = w + (-z) \tag{2.7}$$

The *conjugate* of a complex number z is denoted by $\overline{z}$, and defined by

$$\overline{z} = \overline{a + bi} = a - bi \tag{2.8}$$

The sum and the product of complex numbers can be easily derived by using commutative and distributive laws

$$z + w = (a + bi) + (c + di) = (a + b) + (c + d)i \tag{2.9}$$

$$zw = (ac - bd) + (ad + cb)i \tag{2.10}$$

and also the *negative of a complex number* and the subtraction in **C** by

$$-z = -1z \text{ and} \tag{2.11}$$

$$w - z = w + (-z) \tag{2.12}$$

Complex Conjugate, Absolute Value and Inverse

As we said before, the conjugate of a complex number is denoted and defined by $\overline{z} = \overline{a + bi} = a - bi$. Then

$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2.$$

If we take a complex number $\lambda \in$ **C** , with $\lambda = u + iv$, we can plot this number, and its complex conjugate, in a *complex plane*, as shown in the Fig.2.1.

The *absolute value* of z, denoted by $\|z\|$, is defined to be the nonnegative square root of $z\overline{z}$. Namely,

$$z\overline{z} = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2} \tag{2.13}$$

Note that $\|z\|$ is equal to the norm of the vector $(a, b) \in \mathbf{R}^2$.

Let $z \neq 0$. Then the inverse z^{-1} of z and the *division*, in **C** , of w by z are given, respectively, by

$$z^{-1} = \frac{\overline{z}}{z\overline{z}} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i \quad \text{and} \quad \frac{w}{z} = \frac{w\overline{z}}{z\overline{z}} = wz^{-1}.$$

Exercises

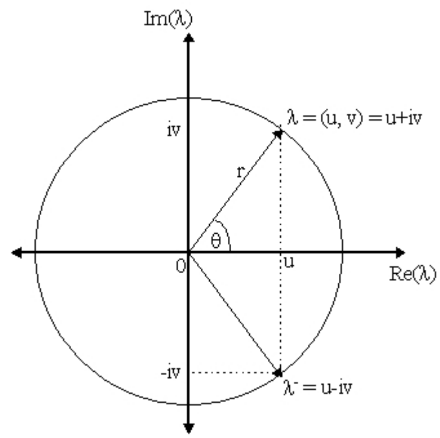


Figure 2.1: The representation of a the complex number, $\lambda = u + iv$, and its complex conjugate, in the complex plane.

1. Let $z = 2 + 3i$ and $w = 5 - 2i$. Calculate: $z + w, zw, \bar{z}, w/z$ and $|z|$.
2. Find the complex conjugate of each of the following: $6 + 4i, 7 - 5i, 4 + i, -3 - i, 6, -3, 4i, -9i$.
3. Find $z\bar{z}$ and $|z|$ when $z = 3 + 4i$.
4. Prove that $\forall\ z, w \in \mathbf{C}$, $\overline{z + w} = \bar{z} + \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$ and $\bar{\bar{z}} = z$.

2.1.4 Vectors in $\mathbf{C}^n$

A set of n-tuples of complex numbers, denoted by $\mathbf{C}^n$, is called *complex* n-*space*. The elements of $\mathbf{C}^n$ are called *vectors* and those of $\mathbf{C}$ are called *scalars*. Vector addition and scalar multiplication, on $\mathbf{C}^n$, are given by

$$[z_1, z_2, \ldots, z_n] + [w_1, w_2, \ldots, w_n] = [z_1 + w_1, z_2 + w_2, \ldots, z_n + w_n] \quad (2.14)$$

$$z[z_1, z_2, \ldots z_n] = [zz_1, zz_2, \ldots, zz_n] \quad (2.15)$$

where $z_i, w_i \in \mathbf{C}$.

The inner product in $\mathbf{C}^n$

Let $\mathbf{u}, \mathbf{v} \in \mathbf{C}^n$ with $\mathbf{u} = [z_1, z_2, \ldots, z_n]$ and $\mathbf{v} = [w_1, w_2, \ldots, w_n]$. The inner product of $\mathbf{u}$ and $\mathbf{v}$ is denoted and defined by

$$\mathbf{u} \cdot \mathbf{v} = z_1\overline{w_1} + z_2\overline{w_2} + \cdots + z_n\overline{w_n} \quad (2.16)$$

When $\mathbf{v}$ is real, this definition reduces to the real case, since in this case $\overline{w_i} = w_i$.

The *norm* of $\mathbf{u}$ is defined by

$$\|u\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{z_1\overline{z_1} + z_2\overline{z_2} + \cdots + z_n\overline{z_n}} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} \quad (2.17)$$

This $\mathbf{C}^n$ space, with the above operations of vector adition, scalar multiplication and dot (inner) product, is called *complex Euclidean n-space*. The properties from $\mathbf{R}^n$ hold with minor modifications. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{C}^n$ and a scalar $k \in \mathbf{C}$:

- (i) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$,
- (ii) $k\mathbf{u} \cdot \mathbf{v} = k(\mathbf{u} \cdot \mathbf{v})$
- (iii) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$,
- (iv) $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = 0$

BUT! There is an interesting and important property of the scalar product in $\mathbf{C}^n$! From (ii) and (iii), we immediately see that:

$$\mathbf{u} \cdot k\mathbf{v} = \overline{kv \cdot u} = \overline{k}(\mathbf{u} \cdot \mathbf{v}) \quad (2.18)$$

This interesting fact allows us to say that the scalar product is **antilinear** in respect to the first term and **linear** in respect to the second term.

The Schwarz's and Minkowski's inequality, are true for $\mathbf{C}^n$ without changes, that is:

The Schwarz inequality: For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{C}^n$, we have:

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\| \quad (2.19)$$

and

The Minkowski inequality: For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{C}^n$, we have:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \quad (2.20)$$

Exercises

1. Find the dot products $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{u}$ when: $\mathbf{u} = (3 - 2i, 4i, 1 + 6i)$ and $\mathbf{v} = (5 + i, 2 - 3i, 7 + 2i)$

2. Prove: For any vectors $\mathbf{u}, \mathbf{v} \in \mathbf{C}^n$ and any scalar $z \in \mathbf{C}$: (i) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$, (ii) $\mathbf{u} \cdot (z\mathbf{v}) = \bar{z}(\mathbf{u} \cdot \mathbf{v})$

2.1.5 The Metric Space

Let X be a set. Then a metric on X is a function $d : X \times X \to R$ which satisfies the following three conditions - also known as the metric space axioms - for all $x, y, z \in X$:

(i) $d(x, y) = 0$ iff $x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, z) \leq d(x, y) + d(y, z)$

A metric space is a pair (X, d), where X is a set and d is a metric on X.

Examples / Exercises

(i) The usual distance in $\mathbf{R}$ is given by $d_1(x, y) = |x - y|$ for x and y in $\mathbf{R}$. This defines a metric on $\mathbf{R}$ which we call *the usual metric on* $\mathbf{R}$.

(ii) Similarly, for any subset A of $\mathbf{R}$ we define the usual metric on A to be the metric $d(x, y) = |x - y|$, for $x, y \in A$.

(iii) $d_\infty = max\{|x_j - y_j|, 1 \leq j \leq n\}$

(iv) The usual metric on $\mathbf{C}$ is given by $d(z, w) = |z - w|$ for $z, w \in \mathbf{C}$. As for the $\mathbf{R}$ above, the same formula is used to define the usual metric on any subset of $\mathbf{C}$.

(v) For $n \in \mathbf{N}$, the usual metric on $\mathbf{R}^n$ is the Euclidean distance, denoted by d_2, which comes from Pythagorass Theorem (or using the Euclidean norm): for $x, y \in \mathbf{R}^n$, with $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$,

$$d_2(x, y) = \sqrt{\sum_{j=1}^{n} |x_j - y_j|^2} \tag{2.21}$$

The same is used to define the usual metric on any subset of $\mathbf{R}^n$.

(vi) (Generalisation of d_1 and d_2) For any real number $p \subset [1, \infty)$, we may define a metric d_p by

$$d_p(x, y) = \left(\sum_{j=1}^{n} |x_j - y_j|^p \right)^{\frac{1}{p}} \tag{2.22}$$

(vii) (The discrete metric) Let V be any set and

$$d(x, y) = \begin{cases} 0, & if\ x = y \\ 1, & otherwise \end{cases} \tag{2.23}$$

2.2 The Hilbert Space

A complex vectorial space, with scalar product and complete under a norm, is called *Hilbert space*. We will denote the Hilbert space with H. A subset of elements from H, with the scalar product from H, is also a Hilbert space and is called *subspace*. The null space, $\{0\} \subset H$ and the whole Hilbert space are two *trivial* spaces. We can construct a subspace of H starting from an arbitrary set $\Phi \subset H$. First we determine the *linear span*, $L(\Phi)$, of the set Φ. This is formed by all the linear combinations (with complex coefficients) of the elements in Φ[3]

$$span(\mathbf{v_1}, \mathbf{v_2}, \dots \mathbf{v_r}) = \{\lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2} + \cdots + \lambda_r \mathbf{v_r} | \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbf{C}\} \tag{2.24}$$

The *closure of the linear span*, $\widetilde{L}(\Phi)$ will be a subspace of H. If $\widetilde{L}(\Phi) = H$ then Φ is said to be *fundamental in* H.

Because we have defined the scalar product, we are able to introduce the orthogonality with all the known properties known from the euclidean spaces. Two elements $\varphi, \varkappa \in H$ are orthogonal $\varphi \perp \varkappa$ if the scalar product of the two vectors is zero

$$< \varphi, \varkappa >= 0 \tag{2.25}$$

Nota bene: *For convenience and for familiarity, until now we have denoted the scalar product by* $\mathbf{a} \cdot \mathbf{b}$*. From now on we will denote it by* $< a, b >$*. This will both allow a natural transition to the bra-ket formalism and a consistency of symbols from this point on.*

Exercise

Let V be the space of nonzero square integrable continuous complex functions in one variable. For every pair of functions, define

$$< f, g >= \int_{-\infty}^{+\infty} f(x)\overline{g}(x)dx$$

Show that with this definition V is an inner product space.

Hint: Prove that $< f, g >= \overline{< g, f >}$, $< f + f', g >=< f, g > + < f', g >$ *and* $< \alpha f, g >= \alpha < f, g >$

---

[3] Actually, the linear span is constructed given a vector space V over a field K, the span of a set S (not necessarily finite) is defined to be the intersection W of all subspaces of V which contain S. When S is a finite set, then W is referred to as the subspace spanned by the vectors in S.

Naturally, an element φ is orthogonal on a subset $\Phi \subset H$ if it is orthogonal on each element from it:

$$\varphi \perp \Phi \text{ iff } < \varphi, v_n >= 0 \ \forall v_n \in \Phi \tag{2.26}$$

Two subsets $\Phi_1, \Phi_2 \in H$ are orthogonal on each other, if every element from Φ_1 is orthogonal on every element from Φ_2

$$\Phi_1 \perp \Phi_2 \text{ iff } < v_{1_n}, v_{2_n} >= 0 \ \forall v_{1_n} \in \Phi_1, v_{2_n} \in \Phi_2 \tag{2.27}$$

If an element φ is orthogonal on a subset $\Phi \subset H$ then, from the linearity of the scalar product, it will be orthogonal also in the linear span $L(\Phi)$. On the other hand, from the continuity of the scalar product, if we have a convergent sequence $\varkappa_n \rightarrow \varkappa$ and an element φ orthogonal on every element $\varkappa_n$ of the sequence, then φ will be orthogonal also on the limit $\varkappa$. This implies that if $\varphi \perp L(\Phi)$ and implicitly on every convergent sequence from $L(\Phi)$, then φ will also be orthogonal on the set of the limits of these sequences, which in fact is the subspace $\widetilde{L}(\Phi)$ (the closure of the linear span). So, the orthogonality of $\varphi \perp \Phi$ implies $\varphi \perp \widetilde{L}(\Phi)$. In particular, if $\varphi \perp \Phi$ and Φ is a fundamental set in H, then $\varphi \perp \widetilde{L}(\Phi) = H$, which means that $\varphi = 0$, because is orthogonal on itself.

The set of all φ elements orthogonal on a subset $\Phi \subset H$, does form a subspace because the fundamental sequences of this set and their limits are orthogonal on Φ. Moreover, from the above considerations, this subspace, denoted by H_1 will be orthogonal on the $H_2 = \widetilde{L}(\Phi)$. the H_1 subspace is called *orthogonal complement* of the set Φ or of the subspace H_2. In general, any subspace X of a product space E has an orthogonal complement $X^\perp$ such that $E = X \oplus X^\perp$.

Given a subspace $H_2 \subset H$ and H_1 his orthogonal complement, then every element $\varphi \in H$ can be uniquely decomposed in

$$\varphi = \varphi_1 + \varphi_2 \quad (\varphi_1 \perp \varphi_2) \tag{2.28}$$

where $\varphi_1 \in H$ and $\varphi_2 \in H_2$. The element φ_1 is called *projection* of φ on the subspace H_1 and φ_2 is called projection of φ on the subspace H_2.

With this decomposition, the norm of φ can be calculated by:

$$\|\varphi\|^2 =< \varphi, \varphi >= \|\varphi_1\|^2 + \|\varphi_2\|^2 \tag{2.29}$$

This is, in fact, a generalization of *the theorem of Pythagoras of Samos*[4].

[4]He a Greek philosopher who lived around 530 BC, mostly in the Greek colony of Crotona in southern Italy. According to tradition he was the first to prove the assertion (theorem) which today bears his name: **If a triangle has sides of length (a, b, c), with sides (a, b) enclosing an angle of 90 degrees, then $a^2 + b^2 = c^2$.**

The relation (2.28) represents a decomposition of the Hilbert space in orthogonal subspaces, denoted by:

$$H = H_1 \oplus H_2 \tag{2.30}$$

We say that the Hilbert space H is the orthogonal sum (or direct sum) of the H_1 and H_2 subspaces. By continuing the decomposition of H in subspaces, in the end we will have a decomposition in subspaces, H_j, orthogonal on each other ($H_i \perp H_j, \ \forall i \neq j$), denoted by

$$H = \sum_i \oplus H_i \tag{2.31}$$

Any element $\varphi \in H$ can be uniquely written as

$$\varphi = \varphi_1 + \varphi_2 + \ldots = \sum_i \varphi_i \tag{2.32}$$

where φ_i is the projection of φ on H_i.

For the finite dimensional case there is a *maximal decomposition* in unidimensional spaces generated by a vector from the orthonormal basis. For the infinite dimensional case, the problem is much more difficult, since even obtaining an orthonormal basis faces the problem of convergence.

2.3 Orthonormal Bases in the Hilbert Space

Let H be a Hilbert space. A *system of elements* is a subset $\{\phi_\alpha\}_{\alpha \in A} \subset H$ indexed after an arbitrary set of indices, A. If $A = \mathbf{N}$, then the system is called *countable.*

A system $\{\phi_\alpha\}_{\alpha \in A}$ is called *orthonormal* if for every $\alpha, \beta \in A$, the orthonormal condition is satisfied

$$< \phi_\alpha, \phi_\beta >= \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases} \tag{2.33}$$

Starting from a countable system, $\{\theta_n\}_{n \in \mathbf{N}}$ of linear independent elements, we can always construct an orthonormal system through an orthonormalization procedure. (This procedure is known from the classical physics or from basic mathematics, as "Gram-Schmidt orthogonalization".) We fist take $\phi_1 = \frac{\theta_1}{\|\theta_1\|}$. We construct $\phi_2 = \lambda_1 \theta_1 + \lambda_2 \theta_2$ with $(\lambda_1, \lambda_2 \in \mathbf{C})$ and we search for λ_1 and λ_2 for which $< \phi_1, \phi_2 >= 0$ and $< \phi_2, \phi_2 >= 1$. We continue with ϕ_3 as a linear combination of the first three elements of the system $\{\theta_n\}$. As we mentioned in the beginning of the paragraph, from a system of linear independent elements we can, in this way, always construct an orthonormal system. Therefore, we will consider only orthonormal systems.

An arbitrary system $\{\phi_\alpha\}_{\alpha\in A}$ is called *complete* if it satisfies one of the two equivalent conditions:

- the system is fundamental in $H(\widetilde{L}\{\phi_\alpha\}) = H$
- the only orthogonal element on $\{\phi_\alpha\}_{\alpha\in A}$ is the zero element $0 \in H$

A Hilbert space, H, in which every orthonormal and complete system is also countable, is called *separable Hilbert space*. Obviously, any orthonormal system belonging to a separable space, is also countable (i.e. finite), but not any countable system is also complete.

Let $\{\phi_n\}_{n\in\mathbf{N}}$ be an orthonormal countable system, not necessary complete, from the separable space H. The orthonormality relation reads

$$< \phi_i, \phi_j >= \delta_{ij}, \quad i, j \in \mathbf{N} \tag{2.34}$$

Consider an element $\psi \in H$. The complex numbers $< \phi_i, \psi >$, $(i \in \mathbf{N})$ are called Fourier coefficients of the element ψ in respect to the considered system. With this elements, we can construct a sequence $\{S_n\}_{n\in\mathbf{N}}$ which has as elements the *partial sums* S_n defined as

$$S_n = \sum_{i=1}^{n} < \phi_i, \psi > \phi_i \tag{2.35}$$

The partial sums sequence is convergent. This can be shown using the above relations in the following way:

$$\|\psi - S_n\| =< \psi - S_n, \psi - S_n >= \|\psi\|^2 - \sum_{i=1}^{n} | < \phi_i, \psi > |^2 \geq 0 \tag{2.36}$$

From this we have the so called *Bessel inequality*:

$$\sum_{n=1}^{n} |\phi_i, \psi|^2 \leq \|\psi\|^2, \quad \forall n \in \mathbf{N} \tag{2.37}$$

For $n \to \infty$ we have the convergence condition for the sequence

$$\sum_{i=1}^{\infty} | < \phi_i, \psi > |^2 \leq \|\psi\|^2 \tag{2.38}$$

From the convergence we see that the sequence $\{S_n\}_{n\in\mathbf{N}}$ is a *fundamental sequence*, because

$$\|S_{n+p} - S_n\|^2 = \sum_{k=n+1}^{n+p} | < \phi_k, \psi > |^2 \to 0, \text{when } \ n \to \infty \tag{2.39}$$

As every fundamental sequence from H is also convergent, the sequence of partial sums is also convergent. The limit of this sequence will be denoted by

$$\psi_s = \sum_{i=1}^{\infty} < \phi_i, \psi > \phi_i \tag{2.40}$$

and this limit will belong to $\widetilde{L}(\{\phi_n\}) \subset H$.

Taking the limit of (2.35) for $n \to \infty$ we see that the partial sums sequence (2.36) is convergent to $\psi(\|\psi - S_n\| \to 0, \text{for } n \to \infty)$ iff the relation

$$\sum_{i=1}^{\infty} | < \phi_i, \psi > |^2 = \|\psi\|^2 \tag{2.41}$$

is satisfied. This relation is called *closure equation* for the element ψ. Now then! Every element $\psi \in H$, for which the closure equation is satisfied, is the limit of the corresponding sequence of partial sums and it can be represented by the *Fourier series*

$$\psi = \sum_{i=1}^{\infty} < \phi_i, \psi > \phi_i \tag{2.42}$$

If for an element the closure equation is not satisfied, then it differs from the limit ψ_s of the partial sums sequence, and it can be written like:

$$\psi = \psi_s + (\psi - \psi_s) \tag{2.43}$$

Using (2.40), after simple manipulations, we obtain:

$$< \psi, \psi_s > = < \psi_s, \psi_s > = \sum_{i=1}^{\infty} | < \phi_i, \psi > |^2 \tag{2.44}$$

which shows that ψ_s and $\psi - \psi_s$ are orthogonal. Because the orthogonal decomposition is unique, ψ_s is the orthogonal projection of ψ on the subspace $\widetilde{L}(\{\phi_n\}) \subset H$ and $\psi - \psi_s$ is the projection on its orthogonal complement. From this we have a very important consequence: if the closure equation is not satisfied for at least one element in H, then $\widetilde{L}(\{\phi_n\})$ has an orthogonal complement $H_1 \neq \{0\}$ and, as a consequence $\{\phi_n\}_{n\in\mathbf{N}}$ is not complete.

If the orthonormal system $\{\phi_n\}_{n\in\mathbf{N}}$ is complete then the only orthogonal element on $\widetilde{L}(\{\phi_n\}) = H$ is the zero element. This means that the closure relation is satisfied for every element $\psi \in H$ because $\psi - \psi_s = 0$. Reciprocal, if the closure equation is satisfied $\forall \ \ \psi \in H$ then we will always have $\psi = \psi_s$ which means that the subspace $\widetilde{L}(\{\phi_n\})$ has as orthogonal complement the element zero so the system $\{\phi_n\}_{n\in\mathbf{N}}$ is complete. In conclusion, the necessary and sufficient condition for an orthonormal system to be complete is for the closure equation (2.41) to be satisfied $\forall \ \ \psi \in H$. Then every element can be written as (2.42).

The Fourier series (2.42) can be interpreted as a decomposition, in an orthogonal sum of unidimensional spaces, of the space H. Denoting by $H_i = \widetilde{L}(\phi_i)$, the unidimensional space generated by ϕ_i, it is easy to motice that the quantity $< \phi_i, \psi >$ is the projection of ψ on H. Because the subspaces $\{H_i\}_{i\in\mathbf{N}}$ are orthogonal on each other, we obtain a maximal decomposition in unidimensional subspaces

$$H = \sum_{i=1}^{\infty} \oplus H_i \tag{2.45}$$

The scalar product of two elements $\psi_1, \psi_2 \in H$ can be written in terms of their Fourier coefficients. From

$$< \phi_i, \phi_j >= \delta_{ij} \quad \text{and} \quad \psi = \sum_{i=1}^{\infty} < \phi_i, \psi > \phi_i$$

after some simple manipulations, we have

$$< \psi_1, \psi_2 >= \sum_{i=1}^{\infty} < \psi_1, \phi_i >< \phi_i, \psi_2 > \tag{2.46}$$

Up to now, we have been investigating the way in which an arbitrary element of a separable Hilbert space can be expanded in a Fourier series in respect to an orthonormal and complete system, $\{\phi_n\}_{n\in\mathbf{N}}$. The inverse problem can also be addressed. We will present here the main lines of this problem, without any proof. The inverse problem is: under what circumstances a countable set of complex numbers $\{\alpha_i\}_{i\in\mathbf{N}}$ can be identified with the set of the Fourier coefficients of an element from a separable Hilbert space? It can be shown (The Riesz- Ficher theorem) that any sequence of (real or complex) numbers, α_i satisfying the condition

$$\sum_{i=1}^{\infty} |\alpha|^2 < \infty \tag{2.47}$$

uniquely determines an element

$$\psi = \sum_{i=1}^{\infty} \alpha_i \phi_i \in H \tag{2.48}$$

such that

$$\alpha_i =< \phi_i, \psi > \quad \text{and} \quad \|\psi\|^2 = \sum_{i=1}^{\infty} |\alpha_i|^2 \tag{2.49}$$

From now on, we will use, interchangeably, the term *base* for *orthonormal and complete system* and the term *components* for *Fourier coefficients*.

Linear isometry

Let H_1 and H_2 be two Hilbert spaces with the scalar products $<,>_1$ and $<,>_2$. If there is a bijective correspondence between them realised by a linear application $f : H_1 \to H_2$

$$\begin{aligned} f(\alpha\varphi + \beta\varkappa) &= \alpha f(\varphi) + \beta f(\varkappa) \\ < \varphi, \varkappa >_1 &= < f(\varphi), f(\varkappa) >_2 \end{aligned} \tag{2.50}$$

then we say that the two spaces are *linear isometric*. If the application is *antilinear*

$$\begin{aligned} f(\alpha\varphi + \beta\varkappa) &= \overline{\alpha} f(\varphi) + \overline{\beta} f(\varkappa) \\ < \varphi, \varkappa >_1 &= < f(\varkappa), f(\varphi) >_2 = \overline{< f(\varphi), f(\varkappa) >_2} \end{aligned} \tag{2.51}$$

the spaces are called *antilinear isometric*. For constructing an isometry it is enough to establish a linear or antilinear correspondence between an orthonormal and complete system from H_1 and one from H_2. For the separable spaces, where these systems are countable, it can be shown that a linear correspondence can always be established $(f(\phi_n^{(1)}) = \phi_n^{(2)}, \quad n \in \mathbf{N})$, so all the separable Hilbert spaces are linear isometric spaces.

The most simple model of separable Hilbert space is the l^2 space. It contains all the vectors of the form

$$\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots) \ \text{with} \ \varphi_n \in \mathbf{C}, \ \sum_{n=1}^{\infty} |\varphi_n|^2 < \infty \tag{2.52}$$

with

$$\begin{aligned} \varphi + \varkappa &= (\varphi_1 + \varkappa_1, \varphi_2 + \varkappa_2, \ldots, \varphi_n + \varkappa_n, \ldots) \\ \alpha\varphi &= (\alpha\varphi_1, \alpha\varphi_2, \ldots, \alpha\varphi_n, \ldots) \end{aligned} \tag{2.53}$$

and the scalar product is a generalization of the scalar product defined on $\mathbf{C}^n$. It is defined

$$< \varphi, \varkappa >= \sum_{i=1}^{\infty} \overline{\varphi_i} \varkappa_i \tag{2.54}$$

In l^2 there is an orthonormal and complete system

$$\begin{aligned} \varphi_1 &= (1, 0, 0, \dots, 0, \dots) \\ \varphi_2 &= (0, 1, 0, \dots, 0, \dots) \\ \text{etc.} & \end{aligned}$$

satisfying the orthogonality relation $< \varphi_i, \varphi_j >= \delta_{i,j}$. The Fourier coefficients of an element $\varphi \in H$ are $< \phi_i, \varphi >= \varphi$. These Fourier coefficients are the vector components of φ.

It can be shown that l^2 is separable and that all the linear isometric and separable Hilbert spaces are linear isometric with l^2.

2.4 The $\mathfrak{L}^2$ space

From the prerequisites to quantum mechanics we know that the wave functions can be physically interpreted if they are square integrable. From the superposition principle, we can intuitively see that the set of all this functions can be viewed as a complex vectorial space. These two observations drive us towards organizing the wave function space as a Hilbert space.

A function $\varphi : \mathbf{R}_x \rightarrow \mathbf{C}$ (the simplest case) is called square integrable if the function $|\varphi(x)|^2$ is locally integrable and if

$$\int_{-\infty}^{+\infty} |\varphi(x)|^2 dx \tag{2.55}$$

is finite. If φ_1 and φ_2 are square integrable, then any combination $\alpha_1\varphi_1 + \alpha_2\varphi_2 \quad (\alpha_1, \alpha_2 \in \mathbf{C})$ will be a square integrable function. So the set of all these functions is organized in a complex vectorial space.

The integral (2.55) might seem a norm but it cannot be a norm since the norm has to be zero only for $\varphi = 0$ and this is not true for φ. Now then! If we want that integral to be zero, then what we have to do is to make it zero *almost everywhere*[5],

[5]In measure theory (a branch of mathematical analysis), one says that a property holds *almost everywhere* if the set of elements for which the property does not hold is a null set, *i.e.* is a set

which means that the function can have non-zero values on a finite, negligible interval. This difficulty can be overcome if instead of the function φ we consider an entire class of functions (called *almost everywhere equivalence class*) formed by all the functions which differ from φ on a negligible set. This class can be represented by any member of it. The space of classes is a complex vectorial space if the operations between classes are defined as:

- the sum of two classes represented by two functions φ_1 and φ_2 is the equivalence class of the function $\varphi_1 + \varphi_2$
- by multiplying a class represented by the function φ with $\alpha \in \mathbf{C}$ we understand the class of the function $\alpha\varphi$

In this space, the set of all almost everywhere null functions forms a class (so, an unique element) which is neutral in respect to addition (0). From these considerations we see that the integral (2.55) is the square of the norm of this space.

This mathematical construction is in perfect agreement with the physical interpretation. The probability density $|\varphi(x)|^2$ cannot be measured point by point. What is in fact measured is the probability of localization on a finite interval which is given by the integral on that interval of the function $|\varphi(x)|^2$. As a conclusion we can say that **all the functions which are equal almost everywhere (from the same equivalence class) describe the same physical state. So, to a particular physical state we do not attach a wave function but an entire corresponding class of equivalence**.

The complex vectorial space of the almost everywhere equivalence classes has a scalar product (defined through representatives)

$$< \varphi, \varkappa >= \int_{-\infty}^{+\infty} \overline{\varphi(x)}\varkappa(x)dx \tag{2.56}$$

It even has a mathematical name and that is $\mathfrak{L}^2(R_x)$. It can be shown that this is a separable Hilbert space. From now on, for simplicity, we will understand that $\varphi \in \mathfrak{L}^2(R_x)$ means the equivalence class and not the function itself.

Because the space is separable, $\exists \, \{\phi_n\}_{n\in\mathbf{N}}$ a countable system, orthonormal and complete. So $\forall \, \psi \in \mathfrak{L}^2(R_x)$ we can write

$$\psi = \sum_{i=1}^{\infty} \phi_i < \phi_i, \psi > \tag{2.57}$$

---

with *measure zero* - a measure is a function that assigns a number, e.g., a "size", "volume", or "probability", to subsets of a given set. The concept has developed in connection with a desire to carry out integration over arbitrary sets rather than on an interval as traditionally done, and is important in mathematical analysis and probability theory.

where the Fourier coefficients are given by the integral

$$< \phi_i, \psi >= \int_{-\infty}^{+\infty} \overline{\phi_i(x)}\psi(x)dx \tag{2.58}$$

Starting from this we can give the proof of one of the most important relations in quantum physics. Let's define the following quantity

$$\sum_{i=1}^{\infty} \phi(x)\overline{\phi(y)} = \delta(x, y) \tag{2.59}$$

and do the following calculation

$$\int_{-\infty}^{+\infty} \delta(x, y)\psi(y)dy = \sum_{i=1}^{\infty} \phi(x) \int_{-\infty}^{+\infty} \overline{\phi_i(x)}\psi(y)dy = \sum_{i=1}^{\infty} \phi(x) < \phi_i, \psi >= \psi(x) \tag{2.60}$$

We see that $\delta(x, y) = \delta(x - y)$ (known as *the Dirac function*) and we obtain the relation

$$\sum_{i=1}^{\infty} \phi(x)\overline{\phi(y)} = \delta(x - y) \tag{2.61}$$

This is true for any orthonormal and complete system and is called *closure equation*.

2.5 Linear and Antilinear Functionals. The Generalized Scalar Product.

The Hilbert space theory is the base of the mathematical structure of quanyum physics: *the hilbertian triad* or *the Gelfand triad/triplet*. This theory is very broad but we will set here the blueprint of it, in accordance with the purpose of this short course.

Let's start with the definition of a functional. Given a Hilbert space, H, and a subset $\Phi \subset H$ - which is a linear space of H - we will call *functional* an application $F : \Phi \to \mathbf{C}$ which assignees to every element $\varphi \in \Phi$ a complex value $F(\varphi)$. If $\forall\ \varphi, \varkappa \in \Phi$ and $\alpha, \beta \in \mathbf{C}$ we have

$$F(\alpha\varphi + \beta\varkappa) = \overline{\alpha}F(\varphi) + \overline{\beta}F(\varkappa) \tag{2.62}$$

we say that the functional, F, is *antilinear*. If $\forall\ \varphi, \varkappa \in \Phi$ and $\alpha, \beta \in \mathbf{C}$ we have

$$F^*(\alpha\varphi + \beta\varkappa) = \alpha F^*(\varphi) + \beta F^*(\varphi) \tag{2.63}$$

we say that the functional, F^*, is *linear*.

The set Φ is the domain of the functional and the set from H for which the functional is zero is called the *kernel* of the functional. If the kernel coincides with the domain than the functional is null.

The above properties of linearity and antilinearity suggest a certain resemblance with the scalar product. This resemblance is not casual and we will use the theory of functionals for defining a generalized scalar product. For this purpose we will introduce a new notation, a notation adequate to the formalism of quantum physics.

A linear functional will be denoted by $\langle F|$ and its values with $\langle F|\varphi\rangle \equiv F^*(\varphi)$. An antilinear functional will be denoted by $|F\rangle$ and $\langle\varphi|F\rangle \equiv F(\varphi)$. These symbols, $\langle|\rangle$, shall not be confused, for now, with the scalar product $<,>$, used for the scalar product in H.

The relations for the functional dependence will read:

$$\langle G|(\varphi) = \langle G|\varphi\rangle \tag{2.64}$$
$$|F\rangle(\varphi) = \langle\varphi|F\rangle \tag{2.65}$$

This notation is somehow ugly or inconvenient, but we will soon replace it with the definitive one. The linear and antilinear properties will be rewritten as:

$$|F\rangle(\alpha\varphi + \beta\varkappa) = \overline{\alpha}\langle\varphi|F\rangle + \overline{\beta}\langle\varkappa|F\rangle \tag{2.66}$$
$$\langle G|(\alpha\varphi + \beta\varkappa) = \alpha\langle G|\varphi\rangle + \beta\langle G|\varkappa\rangle \tag{2.67}$$

To every linear functional $|F\rangle$ we can assign a conjugate functional $\langle F|$ such that

$$\langle\varphi|F\rangle = \overline{\langle F|\varphi\rangle}, \quad \forall\varphi \in \Phi \tag{2.68}$$

In general the domain on which the functionals are defined, Φ, is n̲ot a subspace from H but an arbitrary subset, in which we suppose that we can introduce a topology in order to define the convergence of the sequences in Φ. If $\forall\ \{\varphi_n\}_{n\in\Phi}$ convergent towards φ, in this topology, we have $\langle F|\varphi_n\rangle \to \langle F|\varphi\rangle$, then we say that the functional $\langle F|$ is continuous. In the same time, the functional $\langle F|$ is bounded: $|\langle F|\varphi\rangle \leq C\||\varphi\rangle\|,\ \ \forall\,\varphi \in \Phi$.

The set of all antilinear and continuous functionals, defined on Φ form a vectorial space, denoted by Φ', in respect to addition and multiplication with scalars:

$$(\langle F_1| + \langle F_2|)(\varphi) = \langle\varphi|F_1\rangle + \langle\varphi|F_2\rangle \tag{2.69}$$
$$(\alpha|F\rangle)(\varphi) = \alpha\langle\varphi|F\rangle \tag{2.70}$$

Analogous, for the linear functionals we have a vectorial space $\overline{\Phi}'$, called the *dual* space of Φ'. In it there are defined the operations

$$(|F_1\rangle + |F_2\rangle)(\varphi) = \langle F_1|\varphi\rangle + \langle F_2|\varphi\rangle \tag{2.71}$$

$$(\alpha\langle F|)(\varphi) = \overline{\alpha}\langle F|\varphi\rangle \tag{2.72}$$

The relation (2.68) establishes a bijection between Φ' and $\overline{\Phi'}$, bijection which is antilinear since to $\alpha_1|F_1\rangle + \alpha_2|F_2\rangle \in \Phi'$ it corresponds, by conjugation, the antilinear combination $\overline{\alpha_1}\langle F_1| + \overline{\alpha_2}\langle F_2| \in \overline{\Phi'}$.

The extension of the domain

Let's consider $F : H \to \mathbf{C}$. We will denote by H' the vectorial space of the antilinear functionals (in the H topology) and by $\overline{H}'$ the vectorial space of the linear functionals. $\overline{H}'$ is the dual space of H' and these two are antilinear isometric spaces.

Given the strong topology of H we can define the norms of the (conjugate) functionals

$$\||F\rangle\| = \|\langle F|\| = \sup_{\|\varphi\|\leq 1} |\langle\varphi|F\rangle| \tag{2.73}$$

To every element $\varphi \in H$ we can associate a pair of conjugate functionals $|F_\varphi\rangle$ and $\langle F_\varphi|$ with the values

$$\langle F_\varphi|\varkappa\rangle =< \varphi, \varkappa >, \quad \langle\varkappa|F_\varphi\rangle =< \varkappa, \varphi >, \quad \forall\, \varkappa \in H \tag{2.74}$$

On the other hand, we know, from the Riesz theorem, that to any arbitrary pair of conjugated functionals, continuous on H, we can assign an unique element $\varphi_F \in H$ and we can write the value of the functionals ($\forall\, \varkappa \in H$):

$$\langle F|\varkappa\rangle = < \varphi_F, \varkappa > \tag{2.75}$$

$$\langle\varkappa|F\rangle = < \varkappa, \varphi_F >$$

and have the equality

$$\||F\rangle\| = \|\langle F|\| = \|\varphi_F\|$$

In conclusion, we can establish the following correspondences:

- there is a linear correspondence between H and H'
- there is an antilinear correspondence between H and $\overline{H}'$

If we consider two pairs of conjugate functionals $(|F_1\rangle, \langle F_1|)$ and $(|F_2\rangle, \langle F_2|)$, continuous on H, and the elements $\varphi_{F_1}, \varphi_{F_2} \in H$ corresponding to each pair of functionals (see Riesz theorem), then we can construct the quantity

$$\langle F_1|F_2\rangle \equiv < \varphi_{F_1}, \varphi_{F_2} > \tag{2.76}$$

which has a twofold interpretation:

- the scalar product of the functionals $|F_1\rangle, |F_2\rangle \in H$ denoted by $(|F_1\rangle, |F_2\rangle) = \langle F_1|F_2\rangle$
- the scalar product $(\langle F_1|, \langle F_2|) = \langle F_1|F_2\rangle = \overline{\langle F_2|F_1\rangle}$ of the conjugated functionals from $\overline{H}'$ (We have inverted here the order to ensure the compatibility with the antilinear correspondence between H' and $\overline{H}'$.)

With this insertion of scalar products, the spaces H' and $\overline{H}'$ become Hilbert spaces.

Simplifying the notation: We established before that H' is linearly isometric with H and the dual $\overline{H}'$ is antilinear isometric with H. We can "merge" H and H' by setting $\varphi \equiv |F_\varphi\rangle$, getting rid of the $'$ in H'. We than have only two spaces:

- the H space, with antilinear functionals
- the dual space $\overline{H}$ of linear and continuous functionals on H

Because of the identity between H and H', we will denote the elements $\varphi \in H$ with the symbol for antilinear functionals $|\varphi\rangle$ instead of φ ($|\varphi\rangle \equiv |F_\varphi\rangle$), and the scalar product of two elements $|\varphi\rangle, |\varkappa\rangle \in H$ with $\langle\varphi|\varkappa\rangle \equiv < \varphi, \varkappa >$. The elements from $\overline{H}$ from now on will be denoted by the symbol of the linear functional. For example, the conjugate of the element $|\varphi\rangle \in H$ will be $\langle\varphi| \in \overline{H}$.

Now we go back to the case $\Phi \subset H$. By restraining the continuity domain of the functionals, their space will be enriched by those functionals which are bounded on Φ but unbounded on $H - \Phi$. We then have the following inclusion relations:

$$\begin{aligned} \Phi \subset H \subset \Phi' \\ \overline{\Phi} \subset \overline{H} \subset \overline{\Phi}' \end{aligned} \tag{2.77}$$

where $\overline{\Phi}'$ is the set of linear functionals conjugated with the elements of Φ', which means that $\overline{\Phi}'$ is the dual of Φ'.[6] (If you consult other books or articles - specially of

[6] The dual of a space X, in mathematics, means the set of linear functionals on X, denoted by X'. In this sense $\overline{\Phi}$ is the dual of Φ and Φ' is the dual of $\overline{\Phi}$

mathematics -, pay attention to the definition of the "dual". We use a slightly modified definition which serves our purpose very well and does not affect the consistency of the theory. See also the footnote.)

The triplet of spaces (2.77) are not of much use unless we have defined a topology on Φ and a precise relation between the spaces. This step is rather complicated so we will just blueprint the general frame of the theory.

We say that Φ is *dense* in H if all the elements in H are the limit of a sequence from Φ (in the sense of the topology of H). The triplet $\Phi \subset H \subset \Phi'$, where Φ is dense in H, and H is separable, will be called *hilbertian triad* if it exists at least one orthonormal and complete system of H contained in Φ.

In a Hilbertian triad we can generalize the definition of the scalar product, by giving up the continuity requirement and we can also define a generalized norm. How? We recall that the functionals from Φ' are continuous and bounded on Φ. So if we have a functional $|F\rangle \in \Phi'$ and an element $\varphi \in \Phi$, the value of the functional, $\langle\varphi|F\rangle$, will be finite. This quantity can be interpreted as the scalar product of the elements $|\varphi\rangle \in \Phi \subset \Phi'$ and $|F\rangle \in \Phi'$. From the above hypothesis, there is at least one orthonormal and complete system, denoted now by $\{|\phi_i\rangle\}_{i\in\mathbf{N}} \in \Phi$. This means that the quantities $\langle\phi_i|F\rangle$ are finite ($|F\rangle$ is continuous on Φ). By arbitrary choosing another functional $|G\rangle \in \Phi'$ we obtain again a set of finite numbers $\langle\phi_i|G\rangle$. Considering that $\langle G|\phi_i\rangle = \overline{\langle\phi_i|G\rangle}$ we will define *the generalized scalar product* of the functionals $|G\rangle$ and $|F\rangle$, as a direct generalization of the earlier obtained (2.46):

$$\langle G|F\rangle = \sum_{i=1}^{\infty}\langle G|\phi_i\rangle\langle\phi_i|F\rangle \tag{2.78}$$

This generalized scalar product has necessary properties of linearity and anti-linearity ($\langle F|G\rangle = \overline{\langle G|F\rangle}$).

Starting from (2.78), we can define a *generalized norm* of an element $|F\rangle \in \Phi'$

$$\||F\rangle\|^2 = \langle F|F\rangle = \sum_{i=1}^{\infty}|\langle\phi_i|F\rangle|^2 \tag{2.79}$$

This is an extension of the orthogonality notion to the entire Φ' space. Two elements $|F\rangle, |G\rangle$ are orthogonal if $\langle F|G\rangle = 0$. The discussion about the orthogonal decomposition has no meaning here.

2.6 Fourier Transforms, ρ space and Temperate Distributions

2.6.1 The Fourier transform and the ρ space

We have defined earlier a space called $\mathfrak{L}^2$ as being the vectorial space of the almost everywhere equivalence classes. In order to prove the existence of the Fourier transform for the square integrable functions (recall the wave function) we need to relay on the $\mathfrak{L}^2$ space. The Fourier transform of a function $\varphi : \mathbf{R}_x \to \mathbf{C}$ is a function $\hat{\varphi} : \mathbf{R}_p \to \mathbf{C}$. The two functions are connected through the direct and inverse transformations

$$\hat{\varphi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \varphi(x) e^{-\frac{i}{\hbar}px} dx, \quad \varphi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \hat{\varphi}(x) e^{\frac{i}{\hbar}px} dx \tag{2.80}$$

The transforms give the same result for all almost everywhere equal functions. Unfortunately the elements of $\mathfrak{L}^2$ do not always have Fourier transforms, in the sense of the above relations, since there are square integrable functions for which the above integrals are improper. We can still define the Fourier transform of any element from $\mathfrak{L}^2$ following a specific procedure.

We start from an space $\rho(\mathbf{R}_x)$ of functions $\varphi : \mathbf{R}_x \to \mathbf{C}$ called *test functions*, which are indefinitely derivable (they admit continuous derivatives, $\varphi^{(k)}$, of any order) and they satisfy the condition $(1 + |x|^l)|\varphi^{(k)}(x)| \leq C(l, k, \varphi) < \infty, \quad x \in \mathbf{R}_x, \ \forall\, k, l \in \mathbf{N}$. The test functions decay to zero, when $|x| \to \infty$, faster than any power of $|x|^{-1}$. This will ensure the convergence of the integrals over this functions. A typical example can be $exp(-x^2)$.

The test functions are integrable and implicitly square integrable. Therefore the space $\rho(\mathbf{R}_x)$ is a subset of $\mathfrak{L}^2(\mathbf{R}_x)$ and $\rho(\mathbf{R}_p)$ is a subset of $\mathfrak{L}^2(\mathbf{R}_p)$. Still without a proof we state that $\rho(\mathbf{R}_x)$ is a dense subset of $\mathfrak{L}^2(\mathbf{R}_x)$ and that $\rho(\mathbf{R}_p)$ is a dense subset of $\mathfrak{L}^2(\mathbf{R}_p)$. So, we can now see the test functions and their Fourier transforms as elements from $\mathfrak{L}^2(\mathbf{R}_x)$ and $\mathfrak{L}^2(\mathbf{R}_p)$.

Considering the scalar product we have

$$< \varphi, \varkappa > = \int_{-\infty}^{+\infty} \overline{\varphi}(x) \varkappa(x) dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \overline{\varphi}(x) \left(\int_{-\infty}^{+\infty} \hat{\varkappa}(p) e^{\frac{i}{\hbar}px} dp \right) dx \tag{2.81}$$

Inverting the integration order and considering the equation

$$\overline{\hat{\varphi}}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \overline{\varphi}(x) e^{\frac{i}{\hbar}px} dx \tag{2.82}$$

we obtain

$$< \varphi, \varkappa >= \int_{-\infty}^{+\infty} \overline{\hat{\varphi}}(x)\hat{\varkappa}(p)dp =< \hat{\varphi}, \hat{\varkappa} > \tag{2.83}$$

These considerations allow the extension of the notion of Fourier transforms to the entire space $\mathfrak{L}(\mathbf{R}_x)$. It can be shown that this is also the case for $\mathfrak{L}(\mathbf{R}_p)$ and that the Fourier transform establishes a linear correspondence between $\mathfrak{L}(\mathbf{R}_x)$ and $\mathfrak{L}(\mathbf{R}_p)$.

2.6.2 Temperate distributions ($\mathfrak{D}$)

Continuing with functions spaces we notice that we already have two important terms for constructing a Hilbertian triad *i.e.* the spaces $\rho \subset \mathfrak{L}$. For obtaining the third one we have to construct linear and continuous functionals on ρ – *temperate distributions.*

As in the above section we consider the unidimensional case, and we will call *temperate distribution* a linear or antilinear functional, continuous on $\rho(\mathbf{R}_x)$ and with the codomain $\mathbf{C}$. We will denote by $\langle D|$ the linear distributions and by $|D\rangle$ the antilinear distributions. Obviously, to every linear distribution it corresponds an antilinear one, and vice versa. The study of continuity is a bit more complicated, because we would have to make use of the topology of $\rho(\mathbf{R}_x)$, so we will just say that these functionals are bounded on ρ *i.e.* they cannot have infinite values on ρ).

With $\mathfrak{D}(\mathbf{R}_x)$ we denote the set of all linear distributions and with $\overline{\mathfrak{D}}(\mathbf{R}_x)$ the set of all antilinear distributions – the dual space of $\mathfrak{D}(\mathbf{R}_x)$. If we consider the above unproven statements it follows that $\mathfrak{D}$ and $\overline{\mathfrak{D}}$ are antilinear isometric spaces. Identifying the space of antilinear distributions defined on $\mathfrak{L}^2(\mathbf{R}_x)$ with itself ($\mathfrak{L}' \equiv \mathfrak{L}$) we obtain the following inclusion relation:

$$\rho(\mathbf{R}_x) \subset \mathfrak{L}^2(\mathbf{R}_x) \subset \mathfrak{D}(\mathbf{R}_x) \tag{2.84}$$

ρ is dense in $\mathfrak{L}^2$ so we can show that there is an orthonormal basis $\{\phi_i\}_{n\in\mathbf{N}}$ of $\mathfrak{L}^2$ contained in ρ. So, this triplet is a Hilbertian triad. Before using this, we will give a few examples and we will define the operations with distributions.

The most common distributions are the function type distributions. Using a local integrable function $D(x)$, which satisfies the condition $|D(x)| \leq C|x|^n \;\; \forall\, x \in \mathbf{R}_x$, we can construct the values of the conjugate distributions $|D\rangle$ and $\langle D|$, calculated for an arbitrary test function $|\varphi\rangle \in \rho(\mathbf{R}_x)$

$$\begin{aligned} \langle D|\varphi\rangle &= \int_{-\infty}^{+\infty} \overline{D}(x)\varphi(x)dx \\ \langle \varphi|D\rangle &= \int_{-\infty}^{+\infty} \overline{\varphi}(x)D(x)dx \end{aligned} \tag{2.85}$$

These integrals are convergent due to the presence of the test function.

There are also distributions which do not admit the integral representation (we reefer here to Lebesgue integrability). The typical example are the Dirac distributions, denoted here by $|x\rangle$ and $\langle x|$, and defined by

$$\begin{aligned} \langle x|\varphi\rangle &= \varphi(x) \\ \langle \varphi|x\rangle &= \overline{\varphi}(x) \end{aligned} \tag{2.86}$$

for all $|\varphi\rangle \in \rho(\mathbf{R}_x)$. It can be checked that there is no locally integrable function with which is possible to write the Dirac distribution in the form (2.85). Nevertheless, these distributions are often written down using the *"delta Dirac function"*, $\delta(x-y)$:

$$\begin{aligned} \langle x|\varphi\rangle &= \int_{-\infty}^{+\infty} \delta(x-y)\varphi(y)dy = \varphi(x) \\ \langle \varphi|x\rangle &= \int_{-\infty}^{+\infty} \overline{\varphi}(y)\delta(x-y)dy = \overline{\varphi}(x) \end{aligned} \tag{2.87}$$

These integrals are simple rules of formal calculus and not Lebesgue integrals. This representation is very useful and that's why is very convenient to represent any distribution in its integral form, which can be

- Lebesgue integrals if $D(x)$ is locally integrable
- formal integrals when $D(x)$ cannot represent a function

In the last case $D(x)$ will be still called distributions or *generalized functions*. They must be regarded as symbolic quantities for calculation, having non determined values in the classical sense. Moreover, these quantities will never appear in the final result of calculation.

2.6.3 The derivative and the Fourier transform of a distribution

The derivative of a distribution is defined considering that the derivative of a test function is still a test function. We say that the distribution $\langle D^{(n)}|$ is the n-th derivative of the $\langle D|$ distribution, iff

$$\langle D^{(n)}|\varphi\rangle = (-1)^n\langle D|\varphi^{(n)}\rangle, \ \ \forall|\varphi\rangle \in \rho \tag{2.88}$$

This definition and the equations (2.85) follow from the following formal rule

$$\int_{-\infty}^{+\infty} \overline{D}^{(n)}(x)\varphi(x)dx = (-1)^n \int_{-\infty}^{+\infty} \overline{D}^{(n)}(x)\varphi^{(n)}(x)dx \tag{2.89}$$

which applies regardless if $D(x)$ is a function (derivable or not) or a distribution.

Let's consider a test function $|\varphi\rangle \in \rho(\mathbf{R}_x)$ and its Fourier transform $|\hat{\varphi}\rangle \in \rho(\mathbf{R}_p)$. We call Fourier transform of the distribution $\langle D|$ a new distribution $\langle \hat{D}|$ defined as a linear and continuous functional on $\rho(\mathbf{R}_p)$ satisfying the condition

$$\langle \hat{D}|\hat{\varphi}\rangle \equiv \langle D|\varphi\rangle, \ \ \forall\, |\varphi\rangle \in \rho(\mathbf{R}_x) \tag{2.90}$$

The $|\hat{D}\rangle$ distribution will be represented in integral form

$$\langle \hat{D}|\hat{\varphi}\rangle = \int_{-\infty}^{+\infty} \overline{\hat{D}}(p)\hat{\varphi}(p)dp \tag{2.91}$$

Denoting by $\hat{\mathfrak{D}}(\mathbf{R}_p)$ the space of all Fourier transforms of the distributions $\mathfrak{D}(\mathbf{R}_p)$, we see that between the two spaces there is a linear correspondence. In this way a new hilbertian triad appears:

$$\rho(\hat{\mathbf{R}}_p) \subset \hat{\mathfrak{L}}^2 \subset \hat{\mathfrak{D}}(\mathbf{R}_p) \tag{2.92}$$

with the corresponding triads.

Finally, in the two hilbertian triads, $\rho(\mathbf{R}_x) \subset \mathfrak{L}^2(\mathbf{R}_x) \subset \mathfrak{D}(\mathbf{R}_x)$ and $\rho(\hat{\mathbf{R}}_p) \subset \hat{\mathfrak{L}}^2 \subset \hat{\mathfrak{D}}(\mathbf{R}_p)$, we can introduce the generalized scalar product $\langle G|F\rangle = \sum_{i=1}^{\infty}\langle G|\phi_i\rangle\langle\phi_i|F\rangle$. Let's take the first triad and consider an orthonormal and complete system of $\mathfrak{L}^2(\mathbf{R}_x)$, $\{|\phi_i\rangle\}_{i\in\mathbf{N}} \subset \rho(\mathbf{R}_x)$. Then the scalar product of two distributions $|D_1\rangle, |D_2\rangle \in \mathfrak{D}(\mathbf{R}_x)$ will be

$$\langle D_1|D_2\rangle = \sum_{i=1}^{\infty}\langle D_1|\phi_i\rangle\langle\phi_i|D_2\rangle \tag{2.93}$$

where the quantities $\langle D_1|\phi_i\rangle$ and $\langle\phi_i|D_1\rangle$ are finite ($|\phi_i\rangle \in \rho$), but the result of this sum can be anything. Using this generalized scalar product we can show that $\langle D_1|D_2\rangle = \langle \hat{D}_1|\hat{D}_2\rangle$. From here we have the linear isometry, in generalized sense, of the spaces $\mathfrak{D}(\mathbf{R}_x)$ and $\hat{\mathfrak{D}}(\mathbf{R}_p)$.

Let's write the generalized scalar product for a Dirac distribution:

$$\langle x|y\rangle = \sum_{i=1}^{\infty}\langle x|\phi_i\rangle\langle\phi_i|y\rangle \tag{2.94}$$

Because $|\varphi_i\rangle \in \rho(\mathbf{R}_x)$ we can apply (2.86) and (2.89) and we obtain

$$\langle x|y\rangle = \sum_{i=1}^{\infty} \phi_i(x)\overline{\phi_i}(y) = \delta(x-y) \tag{2.95}$$

If we take the generalized scalar product between a distribution $|D\rangle$ and a Dirac distribution $|x\rangle$, we obtain the formula:

$$\langle D|x\rangle = \sum_{i=1}^{\infty}\langle D|\phi_i\rangle\langle\phi_i|x\rangle = \sum_{i=1}^{\infty}\overline{\phi}_i(x)\langle D|\phi_i\rangle \tag{2.96}$$

which is **meaningless**! But if we formally integrate with a test function $\varphi(x)$ we obtain

$$\int_{-\infty}^{+\infty}\langle D|x\rangle\varphi(x)dx = \sum_{i=1}^{\infty}\langle D|\phi_i\rangle\langle\phi_i|\varphi\rangle \tag{2.97}$$

where considering that $|\varphi\rangle = \sum|\phi\rangle\langle\phi_i|\varphi\rangle$ and that the distribution $|D\rangle$ is antilinear, we have

$$\int_{-\infty}^{+\infty}\langle D|x\rangle\varphi(x)dx = \langle D|(|\varphi\rangle) = \langle D|\varphi\rangle \tag{2.98}$$

which allows us to identify the scalar product $\langle D|\varphi\rangle$ with the distribution $\overline{D}(x)$, obtaining a generalization of (2.86) (valid only for ρ) *i.e.*

$$\begin{aligned}\langle x|D\rangle &= D(x)\\ \langle D|x\rangle &= \overline{D}(x)\end{aligned} \tag{2.99}$$

$\forall\ |D\rangle \in \mathfrak{D}(\mathbf{R}_x)$.

2.7 The States Space

In this section we will gather together the terminology and the definitive rules for calculations within the space associated with the dynamical states of the physical systems.

From the perquisites to quantum mechanics, we know that dynamical states must be described by a normalized wavefunction i.e. by elements from $\mathfrak{L}^2$. In practice though, we may deal with wavefunctions which are not-normalized or distributions which, even if they do not have a direct physical interpretation, they cannot be eliminated from the theory. It results from here that the typical mathematical structure of quantum mechanics is the *rigged Hilbert space* (or *the Gelfand triplet*) $\rho \subset \mathfrak{L}^2 \subset \mathfrak{D}$.

As a general rule, we will accept that for every dynamical state of a physical system we can have a *state vector*, $|\rangle$, called *ket vector* or simply *ket*. For every ket we have a conjugate state vector, denoted by $\langle|$, and called *bra vector* or simply *bra*.

We will denote with the same letter two conjugates state vectors: $|\alpha\rangle$ and its conjugate $\langle\alpha|$. The set of all ket forms the *states space* and the set of all bra the *dual space*.

We will make the hypothesis that the state space has a Gelfand triad structure $\Phi \subset H \subset \Phi'$, and that only the vectors fom H correspond to dynamical states of the physical system. We denote here by Φ' the space of all the ket vectors and by $\overline{\Phi'}$ the space of all the bra vectors. As we have seen, these spaces are antilinear isometric, so that to every arbitrary linear combination

$$|3\rangle = \alpha|1\rangle + \beta|2\rangle \tag{2.100}$$

of ket vectors from Φ' it corresponds, by conjugation, the antilinear combination

$$\langle 3| = \overline{\alpha}\langle 1| + \overline{\beta}\langle 2| \tag{2.101}$$

As showed before, to every pair of vectors $|\alpha\rangle, |\beta\rangle \in \Phi'$ we can associate a complex number $\langle\alpha|\beta\rangle$, called *generalized scalar product*. This product is hermitic ($\langle\alpha|\beta\rangle = \overline{\langle\beta|\alpha\rangle}$ and has known properties of linearity and antilinearity. If $|\alpha\rangle$ and $|\beta\rangle \in H \subset \Phi'$ we find the normal scalar product from H. For simplicity, the generalized scalar product will be called from now on, *scalar product*.

The scalar product $\langle\alpha|\beta\rangle$ represents, in the same time, the value of the linear functional $\langle\alpha|$ calculated for $|\beta\rangle$ and the value of the antilinear functional $|\beta\rangle$ for the vector $\langle\alpha|$ from the dual space. This functional dependence suggests the following rule for formal calculus, introduced by Dirac: *every time when a vector bra meets a ket vector, they come together giving a scalar product.*

This operation is called *scalar multiplication* and is denoted by $(\cdot)$. The scalar product of $|\alpha\rangle$ and $|\beta\rangle$ is obtained by scalar multiplication between the vector $\langle\alpha|$, conjugated by $|\alpha\rangle$, with the vector $|\beta\rangle$:

$$\langle\alpha|\cdot(|\beta\rangle) \equiv (\langle\alpha|)\cdot|\beta\rangle \equiv (\langle\alpha|)\cdot(|\beta\rangle) \equiv \langle\alpha|\beta\rangle \tag{2.102}$$

With this rule we immediately have fulfilled the linearity and antilinearity properties of the dot product. That is:

$$\begin{aligned} \langle 4|3\rangle &= \langle 4|\cdot(\alpha|1\rangle + \beta|2\rangle) = \alpha\langle 4|1\rangle + \beta\langle 4|2\rangle \\ \langle 3|4\rangle &= (\overline{\alpha}\langle 1| + \overline{\beta}\langle 2|)\cdot|4\rangle = \overline{\alpha}\langle 1|4\rangle + \overline{\beta}\langle 4|2\rangle \end{aligned} \tag{2.103}$$

This notation also puts us out of the danger of confusing state vectors with other quantities.

The properties of this generalized scalar product are those from $\mathbf{C}^n$. Only one specification is needed here and that is:

$$\||\alpha\rangle\| = \sqrt{\langle\alpha|\alpha\rangle} \tag{2.104}$$

is the *norm* (and not the generalized norm) of a vector $|\alpha\rangle$. This norm is finite only when $|\alpha\rangle \in H$. If $|\alpha\rangle \in H - \Phi'$ the norm will be infinite.

2.8 Summary of the chapter

It is now time to gather together the main results of this chapter, regarding the Hilbertian triad, $\Phi \subset H \subset \Phi'$. The Hilbert space, H, is a separable space and it has an orthonormal, countable and complete system of vectors $\{\phi_i\}_{i\in\mathbf{N}}$ – a base. Any vector ket $|\alpha\rangle \in H$, can be represented in an unique way through a Fourier series

$$|\alpha\rangle = \sum_{i=1}^{\infty} |\phi_i\rangle\langle\phi_i|\alpha\rangle \tag{2.105}$$

satisfying the completeness equation

$$\langle\alpha|\alpha\rangle = \sum_{i=1}^{\infty}\langle\alpha|\phi_i\rangle\phi_i\alpha < \infty \tag{2.106}$$

The equivalent expansion for $\langle\alpha|$ is

$$\langle\alpha| = \sum_{i=1}^{\infty}\langle\alpha|\phi_i\rangle\langle\phi_i| \tag{2.107}$$

The Fourier coefficients $\langle\phi_i|\alpha\rangle = \overline{\langle\alpha|\phi_i\rangle}$ will be called *components* in the given base. If we consider the base to be orthonormal $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ we can calculate

$$\langle\alpha|\beta\rangle = \left(\sum_{i=1}^{\infty}\langle\alpha|\phi_i\rangle\langle\phi_i|\right)\cdot\left(\sum_{j=1}^{\infty}|\phi_j\rangle\langle\phi_j|\beta\rangle\right) = \tag{2.108}$$
$$= \sum_{i,j=1}^{\infty}\langle\alpha|\phi_i\rangle\langle\phi_i|\phi_j\rangle\langle\phi_j|\beta\rangle = \sum_{i,j=1}^{\infty}\langle\alpha|\phi_i\rangle\langle\phi_i|\beta\rangle \tag{2.109}$$

obtaining (2.46) written in the new notation.

The orthonormal and complete systems from H are not the only systems in respect to which the vectors of this space can be expanded. If we consider the Hilbertian triad $\rho \subset \mathfrak{L}^2 \subset \mathfrak{D}$, in the $\mathfrak{D}$ space there is a system of vectors $\{|x\rangle\}_{x\in\mathbf{R}_x}$ (the Dirac distribution) with the property

$$\langle x|x'\rangle = \delta(x - x'), \quad \forall\, x, x' \in \mathbf{R}_x \tag{2.110}$$

By choosing $|\alpha\rangle \in \mathfrak{L}^2$ we know that it represents a function (in fact an entire equivalence class) which values are determined (up to a discrete set of points) according to $\langle p|D\rangle = \hat{D}(p), \quad \langle D|p\rangle = \overline{\hat{D}}(p)$, by $\langle x|\alpha\rangle$. The scalar product of two vectors $|\alpha\rangle, |\beta\rangle \in \mathfrak{L}^2$ is defined by

$$\langle\alpha|\beta\rangle = \int_{-\infty}^{+\infty}\langle\alpha|x\rangle\langle x|\beta\rangle dx \tag{2.111}$$

Everything is just like we would represent the vectors $|\alpha\rangle$ and $|\beta\rangle$ in the form

$$|\alpha\rangle = \int_{-\infty}^{+\infty}|x\rangle\langle x|\alpha\rangle dx, \quad |\beta\rangle = \int_{-\infty}^{+\infty}|x'\rangle\langle x'|\beta\rangle dx' \tag{2.112}$$

and we would calculate the scalar product using the above equations.

Chapter 3

Linear Operators and Basic Elements of Spectral Theory.

In this chapter we will study the properties of the linear operators acting on the vectorial spaces described in the previous chapter. These operators will play the role of the observable from the classical physics. But the existence of these operators and the beauty of the solid mathematical foundation described (*in* relative *extenso*) earlier, would mean nothing unless we are able to extract the necessary numerical values for the description of the physical world. Keep in mind that we have been talking about abstract infinite dimensional complex spaces, up to now... The numerical values are extracted from the so called *eigenvalue problems of the operators*, and this is the task of the *spectral theory*.

Initially the spectral theory was introduced by David Hilbert, but it was (partially) shaped according to the needs of quantum theory, by John von Neumann. The final form of the spectral theory was given by Gelfand.

3.1 Linear Operators

Te most important quantities in quantum mechanics are the linear operators, because they are associated with observables. We know from the perquisites of quantum mechanics that mean values for physical observables are calculated by applying a certain operator on the wavefunction

$$A\varphi = \varkappa \text{ or } (A\varphi)(x) = \varkappa(x) \tag{3.1}$$

With the notations from Chapter 2 the action of an operator A on a vector ket is

$$A|\varphi\rangle = |\varkappa\rangle \tag{3.2}$$

where φ and $\varkappa$ are, in general, vectors from Φ'. In principle an operator can be defined over the whole Φ' space, but because this space contains vectors of infinite norm controlling the continuity of these operators is a very difficult task. For this reason we will only consider operators defined on the Hilbert space H or on $\Phi \subset H$ of the Gelfand triplet.

The action over 'Kets'

Let's consider operators of the form $A : H \to H_1 \subseteq H$. An operator is *linear* iff

$$A(\alpha|1\rangle + \beta|2\rangle) = \alpha A|1\rangle + \beta A|2\rangle, \text{ with } \alpha, \beta \in \mathbf{C}^n \text{ and } |1\rangle, |2\rangle \in H \tag{3.3}$$

The space H is the *domain* and the space $H_1 \subseteq H$ is the *codomain*.

The *kernel* of the operator is defined as the set of vectors from H which under the action of the operator A correspond to the *null* vector, $\mathbf{0}$; and is denoted by $H_0 = kerr(A)$.

It is an obvious thing that an algebraic operation with operators has as a result an operator with a codomain formed by the intersection of the codomains of the operators used in the algebraic operation.

The algebra

We organize the set of all operators defined on H as *an algebra*, defining:

- The vectorial space operations:

$$A = B + C, \qquad A|A\rangle = B|\varphi\rangle + C|\varphi\rangle \tag{3.4}$$

$$\alpha A, \qquad (\alpha A)|\varphi\rangle = \alpha(A|\phi\rangle \tag{3.5}$$

- Multiplication

$$AB, \quad (AB)|\varphi\rangle = A(B|\varphi\rangle) \tag{3.6}$$

with the properties:

$$(AB)C = A(BC) \tag{3.7}$$

$$A(B + C) = AB + AC \tag{3.8}$$

$$(A + B)C = AC + BC \tag{3.9}$$

$$(\alpha A)B = A(\alpha B) = \alpha AB \tag{3.10}$$

- The *unit operator*, I which for every operator A is
$$IA = AI = A \tag{3.11}$$
- By definition, the *null operator*, O, of the algebra, is the operator obtained by multiplication of $\forall A \in H$ with $0 \in C$:
$$A0 = 0A = O \tag{3.12}$$
and has the property: $\forall A \in H$, $OA = 0$ where $0 \in H$ is a vector.

Commutator, Anticommutator and Inverse

In general $AB \neq BA$. The measure of the noncommutativity of two operators is given by the *commutator*. For any two operators $A, B \in H$ we have:
$$[A, B] = AB - BA \tag{3.13}$$
If $AB = BA$, then the two operators commute. For the commutator, one can use the following properties and relations:
$$[A, B] = -[B, A] \tag{3.14}$$
$$[A, (B + C)] = [A, B] + [A, C] \tag{3.15}$$
$$[A, (\alpha B)] = \alpha[A, B] \tag{3.16}$$
$$[A, BC] = [A, B]C + B[A, C] \tag{3.17}$$
$$[A, [B, C]] + [B, [C, A] + [B, [C, A]] + [C, [A, B]] = 0 \tag{3.18}$$
The *anticommutator* of two operators, A and B, is defined as:
$$\{A, B\} = AB + BA \tag{3.19}$$
The main use of the anticommutator is in defining the product of two operators in the form:
$$AB = \frac{1}{2}[A, B] + \frac{1}{2}\{A, B\} \tag{3.20}$$
A linear operator can be inverted only if it realizes a biunivocal correspondence between the vectors from the domain and those from the codomain. For this to happen it is necessary for the domain and the codomain to have the same dimension and for the nucleus of the operator to be the trivial subspace $\{0\} \in H$. If this is the case, we will say that A^{-1} is the *inverse* of A if it satisfied the relations:
$$A^{-1}A = AA^{-1} = I \tag{3.21}$$
The product of two invertible operators, A and B, is an invertible operator
$$(AB)^{-1} = B^{-1}A^{-1} \tag{3.22}$$

The action over 'Bras'

To every vector ket from H it corresponds a conjugate vector bra from $\overline{H}$. We will define the action of an operator over vectors from $\overline{H}$. Let $A|\varphi\rangle = |\varkappa\rangle$, with $|\varphi\rangle, |\varkappa\rangle \in H$ and a vector $\langle\alpha| \in \overline{H}$. It exists a vector $\langle\beta|$ so that:

$$\langle\alpha|\varkappa\rangle = \langle\beta|\varphi\rangle \tag{3.23}$$

By definition, $\langle\beta|$ is the result of acting with the operator A and we will write:

$$\langle\beta| = \langle\alpha|A \tag{3.24}$$

We have the following rules:

- an operator acts to the right for the ket vectors and to the left for bra vectors
- the scalar product between $\langle\alpha|$ and $A|\varphi\rangle$ is equal with the scalar product between $\langle\alpha|A$ and $\langle\varphi|$:

$$\langle\alpha|A|\varphi\rangle \equiv (\langle\alpha|A) \cdot |\varphi\rangle \equiv \langle\alpha| \cdot (A|\varphi\rangle) \tag{3.25}$$

The Hermitian conjugate of an operator

To every operator A, it corresponds an *adjoint operator* - or a *Hermitian conjugate operator* - denoted by $A^\dagger$ (pronounced "A dagger"), which satisfies the condition:

$$\langle\alpha|A^\dagger|\varphi\rangle = \overline{\langle\beta|A|\alpha\rangle} \quad \text{for } \forall\ |\alpha\rangle, |\beta\rangle \in H \tag{3.26}$$

It is easy to notice that conjugating $|\beta\rangle = A|\alpha\rangle$ we obtain $\langle\beta| = \langle\alpha|A^\dagger$ Properties:

$$(\alpha A + \beta B)^\dagger = \overline{\alpha}A^\dagger + \overline{\beta}B^\dagger \tag{3.27}$$

$$(AB)^\dagger = B^\dagger A^\dagger \tag{3.28}$$

$$(A^\dagger)^\dagger = A \tag{3.29}$$

$$I^\dagger = I \tag{3.30}$$

The operator A is called *hermitic* iff $A =^\dagger$ or, in terms of the scalar product $\langle\alpha|A|\beta\rangle = \langle\beta|A|\alpha\rangle \quad \text{for} \quad \forall\ |\alpha\rangle, |\beta\rangle \in H$. If an operator satisfies the condition $A^\dagger = -A$ than it is said to be *anti-Hermitian*[1]. As a definition, we state that any operator can be written as the sum between a hermitic operator and an antihermitic one:

$$A = \frac{A + A^\dagger}{2} + \frac{A - A^\dagger}{2} \tag{3.31}$$

[1] The diagonal elements of the matrix representation of an anti-Hermitian operator are pure imaginary

Unitary operators

An operator U is called *unitary* if it is the inverse of its adjoint:

$$U^\dagger = U^{-1} \tag{3.32}$$

which implies that

$$U^\dagger U = UU^\dagger = I \tag{3.33}$$

The product of two unitary operators is an unitary operator:

$$(UV)^{-1} = V^{-1}U^{-1} = V^\dagger U^\dagger = (UV)^\dagger \tag{3.34}$$

As linear operators form an associative algebra and thus a *ring*. More generally, in view of the above definitions, an operator A is nilpotent if there is $n \in \mathbf{N}$ such that $A^n = 0$ (the zero function).

Analytical functions of operators

An ordinary analytical function $f(x)$ can be expanded in a Taylor series:

$$f(x) = f(0) + \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n \tag{3.35}$$

In a similar manner, if we replace the powers of the argument with the same powers of an operator A, we obtain - by definition - the corresponding function of an operator:

$$f(A) = f(0)I + \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(0) A^n \tag{3.36}$$

For example, the inverse of the operator $I - A$ is denoted by

$$(I - A)^{-1} = \frac{1}{I - A} \tag{3.37}$$

and by this we understand the geometric series

$$\frac{1}{I - A} = I + A + A^2 + \cdots + A^n + \cdots \tag{3.38}$$

Example

Given $[A, B^n] = n[A, B]B^{n-1}$ show that $[A, F(B)] = [A, B]F'(B)$, where $F'(B)$ is the ordinary derivative of F in respect to B.

Solution Expanding $F(B)$ in a power of series

$$[A, F(B)] = \left[A, \sum_{n=0}^{\infty} b_n B^n\right]$$

and using the equality $[A, B + C] = [A, B] + [A, C]$, we have

$$\left[A, \sum_{n=0}^{\infty} b_n B^n\right] = \sum_{n=0}^{\infty} b_n [A, B]B^{n-1} = [A, B] \sum_{n=0}^{\infty} b_n n B^{n-1}$$

Given a power series expansion $g(x) = \sum a_n x^n$, then $g'(x) = \sum a_n n x^{n-1}$, and so $\sum_{n=1}^{\infty} b_n n B^{n-1} = F'(B)$. Therefore, we have $[A, F(B)] = [A, B]F'(B)$.

A very useful function is the exponential function

$$e^A = I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n \tag{3.39}$$

This is a real function, therefore if A is hermitic then it will represent a hermitic operator.

The product of two exponential functions will be an exponential function but here, the operator at the exponent will not be, in general, the sum of the exponents - as in algebra. By direct calculation we can prove that

$$e^A \cdot e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{4}([[A,B],A]+[[A,B],B])+\cdots} \tag{3.40}$$

If the operators A and B comute, then $e^A \cdot e^B = e^{A+B}$.

We can mention here, also leaving the task of looking up the proof to the reader, another very useful relation:

$$e^A x e^{-A} = x + \frac{1}{1!}[A, x] + \frac{1}{2!}[A, [A, x]] + \cdots \tag{3.41}$$

To every function of operator, $f(A)$ it corresponds a hermitic conjugate function $f(A))^\dagger = \overline{f}(A^\dagger)$ of the form:

$$\overline{f}(A^\dagger) = \overline{f}(0)I + \sum_{n=1}^{\infty} \frac{1}{n!} \overline{f}^{(n)}(0)(A^\dagger)^n \tag{3.42}$$

If the function is real ($\overline{f}^{(n)}(0) = f^{(n)}(0)$) and the operator A is hermitic, then the operator function will be a hermitic operator.

The continuity of linear operators

A linear operator A is continuous on the Hilbert space H if for every convergent series $|\varphi_n\rangle \to |\varphi\rangle$ from H we have $A|\varphi_n\rangle \to A|\varphi\rangle$. The necessary and sufficient condition for a linear operator to be continuous is for it to be *bounded*, that is: $\exists C \in \mathbf{R}^+$ such that

$$\|A|\alpha\rangle\| \leq C\||\varphi\rangle\|, \quad \forall\, |\alpha\rangle \in H \tag{3.43}$$

For every bounded operator a norm can be defined:

$$\|A\| = \underset{\||\alpha\rangle\|\leq 1}{sup} \|A|\alpha\rangle\| < \infty. \tag{3.44}$$

Exercises

1. For A, B and C show that the following identities are valid: $[B, A] = -[A, B]$, $[A+B, C] = [A, C] + [B, C]$ and $[A, BC] = [A, B]C + B[A, C]$

2. Suppose that A, B commute with their commutator *i.e.* $[B, [A, B]] = [A, [A, B]] = 0$. Show that $[A, B^n] = nB^{n-1}[A, B]$ and by using the result that $[A^n, B] = nA^{n-1}[A, B]$.

3. Show that $[A, B]^\dagger = -[A^\dagger, B^\dagger]$

4. If A, B are two Hermitian operators, show that their anticomutator is Hermitian. ($\{A, B\}^\dagger = \{A, B\}$)

3.2 Orthogonal Projectors

In simple words the *spectral theorem* provides conditions under which an operator or a matrix can be diagonalized (that is, represented as a diagonal matrix in some basis). This concept of diagonalization is relatively straightforward for operators on finite-dimensional spaces, but requires some modification for operators on infinite-dimensional spaces. Examples of operators to which the spectral theorem applies are self-adjoint operators or more generally normal operators on Hilbert spaces. The spectral theorem also provides a canonical decomposition, called *the spectral decomposition*, or *eigendecomposition*, of the underlying vector space on which it acts.

A hermitic operator $\Lambda \in H$, which satisfied the condition:

$$\Lambda^2 = \Lambda \tag{3.45}$$

is called *projection operator* or *projector*[2].

[2] Note that the null operator and the unit operator are projectors

If Λ is a projector then $I - \Lambda$ is also a projector. It is trivial to see that

$$\Lambda(I - \Lambda) = \Lambda - \Lambda^2 = 0 \tag{3.46}$$

which means that Λ and $I - \Lambda$ are *orthogonal*.

The role of this orthogonal projectors is to decompose the Hilbert space in two orthogonal subspaces. Any vector $|\varphi\rangle \in H$ can be uniquely decomposed in a sum of two orthogonal vectors

$$|\varphi\rangle = \Lambda|\varphi\rangle + (I - \Lambda)|\varphi\rangle \tag{3.47}$$

which represent the projections of this vector on the two orthogonal subspaces.

In the same way, starting from an orthogonal decomposition $H = H_1 \oplus H_2$, one can define an operator Λ such that $\Lambda|\varphi\rangle = |\varphi_1\rangle \in H_1$. In this case, $I - \Lambda$ will project the vector on H_2 - the orthogonal complement OF H_1. The subspace H_2 is the kernel of the Λ projector, and the space H_2 is the kernel of the $I - \Lambda$ projector.

Let's consider an arbitrary decomposition of the Hilbert space, H:

$$H = \sum_{i=1}^{N} \oplus H_i \tag{3.48}$$

in respect to which every vector $|\varphi\rangle \in H$ can be uniquely written as

$$|\varphi\rangle = |\varphi_1\rangle + |\varphi_2\rangle + \cdots + |\varphi_N\rangle. \tag{3.49}$$

For every subspace H_i the Λ_i projectors are defined as

$$\Lambda_i|\varphi\rangle = |\varphi_i\rangle \;\in\; H \tag{3.50}$$

If we successively apply two operators, we get

$$\Lambda_i\Lambda_j|\varphi\rangle = \begin{cases} |\varphi_i\rangle & i = j \\ 0 & i \neq j \end{cases} \tag{3.51}$$

which means that we can have the orthogonality relation:

$$\Lambda_i\Lambda_j = \Lambda_j\Lambda_i = \Lambda_i\delta_{ij} \tag{3.52}$$

The set of all projectors, $\{\Lambda_i\}$ $(i = 1, 2, \ldots)$ satisfying the orthogonality relation (3.52) forms a *system of orthogonal projectors* corresponding to the decomposition (3.48). If $\Lambda_i \in H_i, \Lambda_j \in H_j$, where H_i and H_j are orthogonal subspaces, then the sum $\Lambda_i + \Lambda_j \in H_i \oplus H_j$.

An orthogonal system of projectors is said to be *complete* when the sum of all the projectors equals the projector for the entire space H - the unit operator, I. That is

$$I = \Lambda_1 + \Lambda_2 + \cdots + \Lambda_N \tag{3.53}$$

This is also called *unit decomposition*.

Taking into consideration the fact that the projectors are hermitic, it is easy to show, by rewriting the equations above, that any projection operator on H is also a projection operator on $\overline{H}$ [3].

If the projector makes a projection on an unidimensional space, than it is called *elementary projector*. An unidimensional subspace H_1 contains vectors like: $\alpha|\varphi\rangle$ (with $\langle 1|1\rangle = 1,\ \alpha = \mathbf{C}$). Such an elementary projector can be represented as:

$$\Lambda_1 = |1\rangle\langle 1| \tag{3.54}$$

Let's consider a separable Hilbert space, H, with a countable system of vectors, $|n\rangle_{n\in\mathbf{N}}$, orthonormal and complete. Every normed vector of the basis, $|n\rangle$, determines an unidimensional space H_n with an elementary projector $\Lambda = |n\rangle\langle n|$. If the basis is orthogonal (as supposed earlier), than we have

$$\langle n|m\rangle = \delta_{n,m} \tag{3.55}$$

$$\Lambda_m\Lambda_n = |n\rangle\langle n|m\rangle\langle m| = \langle n|\delta_{m,n} = \Lambda_n\delta_{m,n} \tag{3.56}$$

This shows that the countable set of projectors $\{\Lambda_n\}_{n\in\mathbf{N}}$ does form a system of orthogonal projectors. The only orthogonal vector on this system is the vector $0 \in H$, which has as orthogonal complement the entire Hilbert space, H. So, the sum of all the orthogonal projectors Λ will be the projector on all the Hilbert space H, namely, the unit operator. Hence, we can write the unit decomposition:

$$I = \sum_{n=1}^{\infty} \Lambda_n = \sum_{n=1}^{\infty} |n\rangle\langle n| \tag{3.57}$$

This relation represents the necessary and sufficient condition for the orthonormal system $|n\rangle_{n\in\mathbf{N}}$ to be complete. It is called *closure relation*.

Exercises

1. Is the sum of two projectors, P_1, P_2 a projector? Prove it.
2. Determine if the product of two projection operators is still a projector.
3. $A = |a\rangle\langle a| - i|a\rangle\langle b| + i|b\rangle\langle a| - |b\rangle\langle b|$ Is A a projection operator? (*Hint: test, again, the conditions:* $A = A^\dagger$ *and* $A = A^2$)

[3]When multiplied or added, only the projectors from an orthogonal system have as a result another projector.

Generalized basis

The generalized basis are those basis with a discrete and countable part, $|n\rangle_{n\in\mathbf{N}} \subset H$, and a continuous part, $|\lambda\rangle_{n\in\mathbf{D}} \subset \Phi' - H$, formed by vectors of *infinite norm* For the discret part, we can construct an orthogonal system of of elementary projectors which, in this case, is <u>not</u> complete. The sum of this elementary projectors will be the projector

$$\Lambda_d = \sum_{n=1}^{\infty} |n\rangle\langle n| \tag{3.58}$$

on the subspace $H_d \subset H$ in which the discrete basis was introduced. We can show that considering the generalized orthonormality relations ths H_d space is orthogonal on the space H_c of the vectors

$$|\varphi\rangle = \int_D |\lambda\rangle\langle\lambda||\varphi\rangle d\lambda \in H \tag{3.59}$$

for which the continuous basis $\{|\lambda\rangle\}_{\lambda\in D}$ was given.The projector Λ_c on the space H_c can be represented as an integral:

$$\Lambda_c = \int_D |\lambda\rangle d\lambda\langle\lambda|. \tag{3.60}$$

It can be directly verified, using the generalized orthonormality relation $\langle\lambda|\lambda'\rangle = \delta(\lambda - \lambda')$, that if $\varphi \in H_c$ then $\Lambda_c|\varphi\rangle = |\varphi\rangle$.

The quantity $|\lambda\rangle\langle\lambda|$, in the equation (3.60), is not a projector! It is a *projector density*, since the square of it has no sense:

$$(|\lambda\rangle\langle\lambda|)(|\lambda'\rangle\langle\lambda'|) = |\lambda\rangle\langle\lambda|\lambda'\rangle\langle\lambda'| = |\lambda\rangle\langle\lambda'|\delta(\lambda - \lambda') \tag{3.61}$$

Performing an integration on the domain $\Delta \subset H$, we obtain the projectors in H_c:

$$\Lambda(D) = \int_\Delta |\lambda\rangle d\lambda\langle\lambda|. \tag{3.62}$$

Let $\Delta_1, \Delta_2 \subset D$ $(\subset R)$ two intervals which define the operators $\Lambda(\Delta_1)$ and $\Lambda(\Delta_2)$. Using (3.61) we have:

$$\Lambda(\Delta_1)\Lambda(\Delta_2) = \int_{\Delta_1} d\lambda \int_{\Delta_2} d\lambda'|\lambda\rangle\langle\lambda|\delta(\lambda - \lambda') \tag{3.63}$$

Using the propertied of the delta function ($\delta(x) = \overline{\delta(x)}$) we obtain:

$$\Lambda(\Delta_1)\Lambda(\Delta_2) = \Lambda(D_1 \cap D_2) \tag{3.64}$$

This means that two projectors defined like in (3.62) are orthogonal iff their definition intervals are disjoint[4]. So, we can construct an arbitrary system of projectors by choosing an arbitrary partition of disjoint intervals. It can be shown that the space H_c and all the spaces obtained by applying tha projectors $\Lambda(\Delta)$ are spaces with countable basis and hence infinitely dimensional. (This proof is not necessary here.)

Summarizing, for the generalized basis in the Hilbert space the necessary and sufficient condition for being complete can be written as the following unitary decomposition:

$$I = \Lambda_d + \Lambda_c = \sum_{i=1}^{\infty} |n\rangle\langle n| + \int_D |\lambda\rangle d\lambda\langle\lambda| \tag{3.65}$$

The typical example of generalized basis (which have only continuous part) are the basis of distributions $\{|x\rangle\}_{x\in\mathbf{R}_x}$ and $\{|p\rangle\}_{p\in\mathbf{R}_p}$ introduced for the Hilbert space $L^2(\mathbf{R}_x)$. For this basis we have

$$I = \int_{-\infty}^{\infty} dx|x\rangle\langle x|, \qquad I = \int_{-\infty}^{\infty} dp|p\rangle\langle p| \tag{3.66}$$

From now on, we will be able to write any operatorial relation in components.

Exercise

Show that the expansion of a ket, $|\Psi\rangle$, in a continuous orthonormal basis $|\alpha\rangle$ is unique.

3.3 Matrix Operators

For finite dimensional spaces, when we have a base, to each operator will correspond a matrix. This kind of correspondence is also to be established for infinite dimensional spaces.

Let's consider the bounded operator $A : H \to H$, a countable basis $\{|n\rangle\}_{n\in\mathbf{N}}\ inH$ and the orthonormality condition $\langle n|n'\rangle = \delta_{n,n'}$. Using the unitary decomposition of the A operator we have:

$$A = IAI = \sum_{i,j} |i\rangle \underbrace{\langle i|A|j\rangle}_{elements}\langle j| \tag{3.67}$$

The *quantities* $\langle i|A|j\rangle$ are nothing more than the *matrix elements* of the operator A. These matrix elements are *complex numbers* and the first index represents the

[4]Two sets are said to be **disjoint** if they have no element in common. For example: $\{1, 2, 3\}$ and $\{4, 5, 6\}$ are disjoint sets.

columns and the second index represents the rows. This is called *matrix of the operator* A in the base $\{|n\rangle\}_{n\in\mathbf{N}}$ and is the natural generalization of a matrix from the finite dimensional spaces.

Properties of the matrix operators

- the matrix of the sum of two operators is the sum of the corresponding matrices calculated element by element
- the matrix of the operator αA is obtained by multiplying every element of the matrix with the scalar α
- the matrix of the product AB is given by $\langle i|AB|j\rangle = \langle i|AIB|j\rangle = \sum_{k=1}^{\infty}\langle i|A|k\rangle\langle k|B|j\rangle$
- the matrix of the $A^\dagger$ operator is the hermitic conjugate of the matrix of the operator A - *i.e.* transposed and complex conjugated:

$$\langle i|A^\dagger|j\rangle = \overline{\langle j|A|i\rangle}$$

In a given base the actions of the operators can be written in matrix form. For example, the equation $|\varphi\rangle = A\varkappa$ can be written as

$$\langle i||\varphi\rangle = \langle i|AI|\varkappa\rangle = \sum_{n=1}^{\infty}\langle i|A|k\rangle\langle k|\varphi\rangle \tag{3.68}$$

or, in other "words" (with $i \in \mathbf{N}$)

$$\begin{pmatrix} \langle 1|\varphi\rangle \\ \langle 2|\varphi\rangle \\ \langle 3|\varphi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & \langle 1|A|3\rangle & \cdots \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & \langle 2|A|3\rangle & \cdots \\ \langle 3|A|1\rangle & \langle 3|A|2\rangle & \langle 3|A|3\rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \langle 1|\varkappa\rangle \\ \langle 2|\varkappa\rangle \\ \langle 3|\varkappa\rangle \\ \vdots \end{pmatrix} \tag{3.69}$$

where the operator A is represented by its matrix.

Let's consider, now, an arbitrary orthogonal decomposition of the Hilbert space $H = \sum_{i=1}^{\infty}\oplus H_i$ and the corresponding system of orthogonal projectors $\{\Lambda_i\}_{i=1,\ldots,N}$. Using the closure relation (3.57) we can write down the following decomposition:

$$A = AIA = \left(\sum_{i=1}^{N}\Lambda_i\right) A \left(\sum_{j=1}^{N}\Lambda_j\right) = \left(\sum_{i,j=1}^{N}\Lambda_i A \Lambda_i\right) \tag{3.70}$$

in which operators of the form $A_{i,j} = \Lambda_i A \Lambda_j$ appear. If $i = j$ than A_{ij} is called the projection of the operator A on the subspace H_i. If $i \neq j$ then the operators A_{ij} are nilpotent:

$$(A_{ij})^2 = \Lambda_i A \Lambda_j \, \Lambda_i A \Lambda_j = 0 \tag{3.71}$$

Important! In general, an operator A_{ij} acts on the subspace H_j and has the codomain in H_i. The matrices of these operators, in a base formed by the union of the basis of the subspaces $H_1, H_2, \ldots, H_n$ (in this order), represent subtables of the operator A, with a number of columns equal to $dimH_i$ and a number of rows equal to $dimH_j$. These subtables are called *blocks*. The matrix of the operator is decomposable in blocks:

$$\begin{vmatrix} \vdots & A_{11} & \vdots & A_{12} & \vdots & \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots \\ \vdots & A_{21} & \vdots & A_{22} & \vdots & \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots \\ \vdots & & \vdots & & \vdots & \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots \end{vmatrix} \tag{3.72}$$

and every block represents the matrix of an operator A_{ij}.

A subspace H_k is said to be an *invariant subspace* of the operator A if $\forall |k_i\rangle \in H_k$ we have $A|k_i\rangle = |k_j\rangle$ where $|k_i\rangle, |k_j\rangle \in H_k$. In other words, acting with the operator A on any vector from H_k produces another vector from H_k. This means that all the A_{ik} operators are zero for $i \neq k$ and so the matrix of the operator A will have all the blocks on the k column zero; the only nonzero block will be the block corresponding to the A_{kk} operator. This matrix is called *reducible matrix*.

If the orthogonal complement of H_k is also an invariant subspace, it will be impossible to obtain the vectors from H_k by acting with the operator on the vectors from its orthogonal complement. In this case, also the operators $A_k j$ with $j \neq k$ will be zero, which is equivalent with the proposition: the operator A commutes with the projector Λ_k. Now, the matrix will have also the blocks from the k row zero; the only nonzero block will be the block corresponding to the A_{kk} operator. This matrix is called *completely reducible*.

From $\langle i|A^\dagger|j\rangle = \overline{\langle j|A|i\rangle}$ it follows that if the operator is hermitic and its matrix reducible, then the matrix is also completely reducible. So, if the operator is hermitic and it admits an invariant subspace, then also its orthogonal complement will be an invariant subspace. In general a hermitic operator, A, can admit more invariant subspaces $H_1, H_2, \ldots, H_N$ which will form an orthogonal, complete system. In a base formed by the union of the basis of these spaces, the operators' matrix will have the form:

$$\left|\begin{array}{ccccccccc} \vdots & A_{11} & \vdots & 0 & \vdots & 0\ldots & \vdots & 0 & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & 0 & \vdots & A_{22} & \vdots & 0\ldots & \vdots & 0 & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & 0 & \vdots & 0 & \vdots & \ddots & \vdots & & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & 0 & \vdots & 0 & \vdots & & \vdots & A_{NN} & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \end{array}\right|$$

If the projectors on the invariant subspaces are denoted by $\Lambda_1, \Lambda_2, \ldots, \Lambda_N$ we see they will commute with the operator A so in the decomposition

$$A = AIA = \left(\sum_{i=1}^{N} \Lambda_i\right) A \left(\sum_{j=1}^{N} \Lambda_j\right) = \left(\sum_{i,j=1}^{N} \Lambda_i A \Lambda_i\right)$$

only invariant subspaces will appear:

$$A = \sum_{i=1}^{N} \Lambda_i A \Lambda_i = \sum_{i=1}^{N} \Lambda_i A = \sum_{i=1}^{N} A \Lambda_i \tag{3.73}$$

Wen the matrix elements of an operator A in the base $\{|n\rangle_{n\in\mathbf{N}}$ have the form

$$\langle i|A|j\rangle = a_i \delta_{ij} \tag{3.74}$$

the matrix is called *diagonal*. That is: the operator can be *diagonalized* in this basis. In the following, we will see that the basis in which an operator becomes diagonal is its basis of eigenvectors.

3.4 The Eigenvalue Problem for Operators

Matrix diagonalization is the process of taking a square matrix, $n \times n$, and converting it into a special type of matrix - a so-called *diagonal matrix* - that shares the same fundamental properties of the underlying matrix. Matrix diagonalization is equivalent to transforming the underlying system of equations into a special set of coordinate axes in which the matrix takes this canonical form. Diagonalizing a matrix is also equivalent to finding the matrix's *eigenvalues*, which turn out to be precisely the entries of the diagonalized matrix. Similarly, the *eigenvectors* make up the new set of axes corresponding to the diagonal matrix.

The remarkable relationship between a diagonalized matrix, eigenvalues, and eigenvectors follows from the mathematical identity (*the eigen decomposition*) that a square matrix A can be decomposed into the very special form

$$A = PDP^{-1} \tag{3.75}$$

where P is a matrix composed of the eigenvectors of A, D is the diagonal matrix constructed from the corresponding eigenvalues, and P^{-1} is the *matrix inverse* of P. According to the eigen decomposition theorem [5], an initial matrix equation

$$A\mathbf{X} = \mathbf{Y} \tag{3.76}$$

can always be written

$$PDP^{-1}\mathbf{X} = \mathbf{Y} \tag{3.77}$$

(at least as long as P is a square matrix), and premultiplying both sides by P^{-1} gives

$$DP^{-1}\mathbf{X} = P^{-1}\mathbf{Y} \tag{3.78}$$

Since the same linear transformation P^{-1} is being applied to both $\mathbf{X}$ and $\mathbf{Y}$, solving the original system is equivalent to solving the transformed system

$$D\mathbf{X}' = \mathbf{Y}' \tag{3.79}$$

where $\mathbf{X}' \equiv P^{-1}\mathbf{X}$ and $\mathbf{Y}' \equiv P^{-1}\mathbf{Y}$. This provides a way to canonicalize a system into the simplest possible form, reduce the number of parameters from $n \times n$ for an arbitrary matrix to n for a diagonal matrix, and obtain the characteristic properties of the initial matrix.

The eigenvalue problems from quantum mechanics are a generalization to the infinite dimensional case of the matrix diagonalization procedure, described above. The eigenvalue problem for the operators defined on the Hilbert space depends very much upon the nature of the operators. We can distinguish two types of operators: *bounded operators* and *unbounded operators*.

[5] Let P be a matrix of eigenvectors of a given square matrix A and D be a diagonal matrix with the corresponding eigenvalues on the diagonal. Then, as long as P is a square matrix, A can be written as an eigen decomposition $A = PDP^{-1}$ where D is a diagonal matrix. Furthermore, if A is *symmetric*, then the columns of P are orthogonal vectors.

3.4.1 The Eigenvalue Problem for Bounded Operators

The bounded operators are the operators for which the domain and the codomain is the entire Hilbert space. Such an operator is hermitic if $\langle\alpha|A|\beta\rangle = \overline{\langle\beta|A|\alpha\rangle}, \quad \forall\, |\alpha\rangle, |\beta\rangle \in H$. Considering a hermitic operator A, the equation

$$A|a\rangle = a|a\rangle, \qquad |a\rangle \in H \tag{3.80}$$

is called *the eigenvalue equation in* H. To it, it corresponds a conjugate, equivalent equation in the dual space $\overline{H}$

$$\langle a|A = \overline{a}\langle a| \tag{3.81}$$

These equations admit a trivial solution $|a\rangle = 0, \quad \forall\, a \in \mathbf{C}$. If the solutions are not trivial ($|a\rangle \neq 0$) we say: **the vector** $|a\rangle$ **is the eigenvector corresponding to the eigenvalue** a. More eigenvectors, $|a,1\rangle, |a,2\rangle, \ldots$ can correspond to an arbitrary eigenvalue, a. Every linear combination of eigenvectors is a new eigenvector so every vector from the subspace $H_a = \widetilde{L}(\{|a,1\rangle, |a,2\rangle, \ldots\})$ will be an eigenvector. This subspace will be called *eigensubspace* corresponding to the eigenvalue a. Every eigensubspace is an invariant subspace. If the eigensubspace is unidimensional then the eigenvalue is called *nondegenerate*. Otherwise it is called *degenerate* and degeneration degree, $g(a)$ is given by the dimension of the eigensubspace:

$$g(a) = dim\, H_a \tag{3.82}$$

If the eigenvalue is zero, the eigensubspace H_0 contains all the $|\alpha\rangle$ eigenvectors for which $A|a\rangle = 0$. Obviously, $H_0 = kerr\, A$. If $H_0 \neq \{0\}$ then the operator will not be invertible.

The set of all eigenvalues of an operator, A, forms the *spectrum* of that operator, denoted by $S(A)$. This spectrum is called *simple spectrum* if all the eigenvalues are nondegenerated.

The eigenvalues of a hermitic operator

Let's consider two nontrivial eigenvectors $|a_1\rangle$ and $|a_2\rangle$ corresponding to the eigenvalues a_1 and a_2, which satisfy the equations:

$$A|a_1\rangle = a_1|a_1\rangle \qquad A|a_2\rangle = a_2|a_2\rangle$$

Multiplying the first with $|a_2\rangle$ and the second with $|a_1\rangle$ we obtain:

$$0 = \langle a_2|A|a_1\rangle - \overline{\langle a_1|A|a_2\rangle} = (a_1 - \overline{a}_2)\langle a_2|a_1\rangle \tag{3.83}$$

- If $|a_1\rangle = |a_2\rangle$ (and implicitly $a_1 = a_2$) then, because $|a_1\rangle$ is a nontrivial eigenvector, we will have $\langle a_1|a_1\rangle \neq 0$ which implies that $a_1 = \overline{a}_1$. Hence, *all the eigenvalues of a hermitic operator are real numbers.*

- If $a_1 \neq a_2$ then the above equation will hold only if $\langle a_1|a_2\rangle = 0$. This means that two eigenvectors, corresponding to two different eigenvalues, are orthogonal. Hence, the corresponding eigenspaces H_{a_1} and H_{a_2} are orthogonal.

To every eigenvalue from the spectrum it corresponds an eigensubspace. These subspaces will be orthogonal on each other. It can be proven that the direct sum of these eigensubspaces coincides with the whole Hilbert space, H:

$$H = \sum_{a \in S(A)} \oplus H_a \tag{3.84}$$

which means that the system of eigensubspaces is complete.

A separable Hilbert space admits a countable basis and hence a maximal countable orthogonal decomposition in unidimensional orthogonal spaces. Because the eigensubspaces can orthogonally decompose the Hilbert space in orthogonal spaces, follows that the set of these eigensubspaces is countable (finite or countable). Obviously the corresponding spectrum will be countable *i.e. a discrete spectrum.*

Because the system of eigensubspaces is a complete one, the set of all projectors $\{\Lambda_{a_n}\}_{a_n \in S(A)}$ - where Λ_{a_n} is the projector of the subspace H_{a_n} - is complete. Such, we have:

$$I = \sum_{a_n \in S(A)} \Lambda_{a_n} \tag{3.85}$$

With this system of orthogonal projectors we can define the *spectral decomposition*

$$A = \sum_{a_n \in S(A)} a_n \Lambda_{a_n} \tag{3.86}$$

In this sum, each term represents a projection of the operator A on an eigensubspace.

If the operator has a simple spectrum, the eigensubspaces are unidimensional. We can take from every subspace a normed vector and build a countable system (obviously!), orthonormal and complete, of eigenvectors, $\{|a_n\rangle\}_{a_n \in S(A)}$, and the projection operators will be elementary: $\Lambda_{a_n} = |a_n\rangle\langle a_n|$. The above spectral decompositions will be written, in this case:

$$I = \sum_{a_n \in S(A)} |a_n\rangle\langle a_n|, \quad A = \sum_{a_n \in S(A)} |a_n\rangle\langle a_n|a_n \tag{3.87}$$

and the orthonormal relations will be:

$$\langle a_n | a'_n \rangle = \delta_{n,n'} \tag{3.88}$$

If the spectrum is not simple, the eigensubspace H_{a_n} has a dimension $g(a_n) > 1$. We can choose an arbitrary orthonormal basis from H_{a_n}, $\{|a_n, r_n\rangle\}_{r_n=1,2,\ldots,g(a_n)}$, for which the orthonormal condition is written

$$\langle a_n, r_n | a'_n, r_n \rangle = \delta_{n,n'} \delta_{r_n, r'_n} \tag{3.89}$$

Considering the union of all these basis, we will obtain an orthonormed and complete system in H, with the above orthonormal condition. In this base the Λ_{a_n} projectors are expanded as

$$\Lambda_{a_n} = \sum_{r_n=1}^{g(A)} \langle a_n, r_n | a'_n, r_n \rangle \tag{3.90}$$

and, as a consequence, the spectral decompositions are

$$I = \sum_{a_n \in S(A)} \sum_{r_n=1}^{g(A)} |a_n, r_n\rangle\langle a'_n, r_n|, \qquad A = \sum_{a_n \in S(A)} a_n \sum_{r_n=1}^{g(A)} |a_n, r_n\rangle\langle a'_n, r_n| \tag{3.91}$$

The matrix elements of the operator A in its basis of eigenvectors, are nonzero only if they are diagonal

$$\langle a'_n, r_n | A | a_n, r_n \rangle = a_n \delta_{n,n'} \delta_{r_n, r'_n} \tag{3.92}$$

So, like in the finite-dimensonal case, the basis in which the matrix of an operator is diagonal, is the basis of the eigenvectors of this operator.

We can now consider the inverse problem: Taking an arbitrary countable basis $\{|n\rangle\}_{n \in \mathbf{N}} \subset H$ we can always construct a hermitic operator, A, diagonal in this basis. For this it is enough to have a finite series of natural numbers $\{a_n\}_{n \in \mathbf{N}}$ which will represent the spectrum $S(A)$ of the operator A. In this case the operator can be written as

$$A = \sum_{n=1}^{\infty} |n\rangle a_n \langle n| \tag{3.93}$$

Obviously, every vector from the base is an eigenvector of A, corresponding to the eigenvalue a_n.

This method of construction (think a bit to it!), shows clearly that an operator is fully determined by its spectrum. From the spectrums' structure it must be visible if the operator is or not bounded. We can impose the bounding condition, by firstly calculating:

$$\|A|\varphi\rangle\|^2 = \langle\varphi|A^2|\varphi\rangle = \langle\varphi|\left(\sum_{n=1}^{\infty}|n\rangle a_n\langle n|\right)|\varphi\rangle = \sum_{n=1}^{\infty}a_n^2|\langle n|\varphi\rangle|^2 \tag{3.94}$$

and considering that $\|A\| = \sup_{\||\alpha\rangle\|\leq 1}\|A|\alpha\rangle\| < \infty$ we obtain

$$\sum_{n=1}^{\infty}a_n^2|\langle n|\varphi\rangle|^2 \leq C^2\||\varphi\rangle\|^2 \equiv C^2\sum_{n=1}^{\infty}|\langle n|\varphi\rangle|^2. \tag{3.95}$$

This relation should be satisfied $\forall\,|\alpha\rangle \in H$. This happens only if the sequence of the eigenvalues is inferiorly and superiorly bounded ($|a_n| < C$) or, in other words, *the spectrum is bounded.*

3.4.2 The Eigenvalue Problem for Unbounded Operators

From the usual problems of quantum mechanics, we see that the operators attached to the observables have always an unbounded spectrum. This mean that we have to work, in general, with unbounded operators. The difficulties arise right from the beginning, when we have to define the domain on which these operators are defined. We choose to illustrate these difficulties on a special class of operators with an unbounded spectrum, $S(A) = \{a_n\}_{n_{\mathbf{N}}}$, defined by (3.93). This class is particular because we have chosen the operators with discrete spectrum which admit a complete system of eigenvectors in H. We will see that there are operators which neither admit a discrete spectrum nor eigenvectors in H. For the following discussion the operators defined by (3.93) are enough.

Let's take a an operator A with an eigenvalue system of the form $a_n = n^k$ where $k \in \mathbf{N}$. The domain will be defined only by those operators $|\varphi\rangle\ inH$ which satisfy the condition (3.95). This fact ensures that the operator is bounded and also ensures the continuity. In other words, the definition domain will contain only vectors for which the sequence $\sum_n n^{2k}|\langle n|\varphi\rangle|^2$ is convergent. These vectors must have Fourier coefficients $\langle n|\varphi\rangle$ which go to zero faster than $1/n^k$, when $n \to \infty$. Their set, denoted by ϵ_k, can be considered to be the definition domain of the operator A. We can do this only if we give up doing algebraic operations, because the spectrum of the A^2 operator will diverge like the sequence n^{2k}, being defined only in $\Phi_{2k} \subset \Phi_k$. It is obvious now, that the definition domain of the unbounded operators must be taken at the intersection of all the Φ_k sets. This forms the inclusion relation $\Phi_{k+1} \subset \Phi_k$.

To each Φ_k se t it corresponds a set of antilinear functionals Φ'_k. These sets of antilinear functionals are also ordered by the above inclusion relation, $\Phi_{k+1} \subset \Phi_k$, because as the definition domain becomes smaller, the functionals' space becomes richer. In this way we get the sequence:

$$\ldots \Phi_k \subset \Phi_{k-1} \subset \ldots \Phi_1 \subset H\Phi'_1 \ldots \Phi'_{k-1} \subset \Phi'_k \subset \ldots$$

According to the above logic, the definition domain of the unbounded operators will be the set:

$$\Phi = \bigcap_{k-1}^{\infty} \Phi_k$$

So, this set contains all the vectors which have Fourier coefficients exponentially decaying to zero when $n \to \infty$. These vectors are not just a few, they are quite many so the set is pretty large. It can be even shown that this set is dense in H.

On the set Φ we can define antilinear and continuous functionals which will belong to the set

$$\Phi' = \bigcup_{k=1}^{\infty} \Phi'_k$$

We obtain a hilbertian triad (or a Gelfand triad) $\Phi \subset H \subset \Phi'$ in which Φ is the definition domain for the unbounded operators. It can be shown (but that's beyond the purpose of this course) that by taking $H = \mathfrak{L}^2(\mathbf{R}^n_x)$, this construction method leads to the known Gelfand triad: $\rho(\mathbf{R}^n_x \subset \mathfrak{L}^2(\mathbf{R}^n_x) \subset \mathfrak{D}^2(\mathbf{R}^n_x)$.

In the quantum mechanics we will use only those unbounded operators which are defined on Φ and which through their action: i) to the right on any vector from Φ will produce a vector which is also from Φ and ii) to the left on any vector from $\overline{\Phi}$ will produce a vector which is also from $\overline{\Phi}$. Mathematically, we denote this by $A : \Phi \to \Phi$.

In practice we need to extend the action of these operators on the whole Φ' space, space which contains both vectors of finite norm and infinite norm. This is done through a procedure which resembles very much with the definition of derivatives in the case of distributions, starting from the test functions from ρ. As we have shown in the previous chapter, in spite the fact that the derivation is defined, in the sense given by mathematical analysis, only for those functions in ρ which are infinitely derivable, it can be extended in the weak sense of the distributions, to all the distributions from$\mathfrak{D}$. We will proceed in the same way for the unbounded operators defined on Φ.

Let's consider an arbitrary vector $\langle\varphi| \in \overline{\Phi}$. The operator A acts to the left on this vector, producing $\langle\varkappa| = \langle\varphi|A$. We further consider a vector $|F\rangle \in \Phi'$. We know that the scalar product $\langle\varkappa|F\rangle$ is finite, because the space Φ' is the space of the antilinear and continuous functionals defined on Φ. By *definition* we will say that the vector $|G\rangle \in \Phi'$ is the result of the action of the A operator on the vector $|F\rangle$ and we will denote this by

$$|G\rangle = A|F\rangle \tag{3.96}$$

if the equality

$$\langle\varphi|G\rangle = \langle\varkappa|F\rangle \tag{3.97}$$

with $\langle\varkappa| = \langle\varphi|A$, is valid $\forall\,\langle\varphi| \in \Phi'$. Analogous we can define the action at the left on the vectors from $\overline{\Phi'}$. For this we will take $|\varphi\rangle \in \Phi$ and $|\varkappa\rangle = A|\varphi\rangle \in \Phi$. The vector $\langle G|$ is the result of acting with A on $\langle F|$

$$\langle G| = \langle F|A \tag{3.98}$$

if the equality

$$\langle G|\varphi\rangle = \langle F|\varkappa\rangle \tag{3.99}$$

is satisfied, $\forall\,|\varphi\rangle \in \Phi$.

Like in the case of the operations defined for distributions, the relations (3.96) and (3.98) must be considered to be formal relations from which we, in fact, understand that the equalities (3.97) and (3.97) are satisfied vor every vector from $\overline{\Phi}$ and Φ.

An unbounded operator A is called *symmetric* if $\forall\,|\varkappa\rangle, |\varphi\rangle \in \Phi$ we have

$$\langle\varkappa|A|\varphi\rangle = \overline{\langle\varphi|A|\varkappa\rangle}. \tag{3.100}$$

If a symmetric operator is given, this satisfies the condition

$$\langle F|A|G\rangle = \overline{\langle G|A|F\rangle} \tag{3.101}$$

for all vectors $|F\rangle, |G\rangle \in \Phi$ for which this condition is finite, we will say that the operator is hermitic and we will write $A = A^{\dagger}$.

The *spectral analysis of unbounded operators*, is a much more complicated subject for the purpose and the length of this lectures, which is why we do not discuss it here.