Mathematical background

Rocco Martinazzo

October 6, 2015

Contents

1 Linear spaces: basic definitions

Let $\mathcal{E}$ be a set whose elements will be called *vectors*. $\mathcal{E}$ is a linear space on the field $\mathbb{K}$ (usually $\mathbb{R}$ or $\mathbb{C}$), if it is closed with respect two operations, namely the sum $(+)$ between elements of $\mathcal{E}$

$$\underline{x}, \underline{y} \in \mathcal{E} \Longrightarrow \underline{x} + \underline{y} \in \mathcal{E}$$

and the product (no symbols, or at most $\cdot$) between the elements of $\mathcal{E}$ and the elements of $\mathbb{K}$

$$\underline{x} \in \mathcal{E}, \lambda \in \mathbb{K} \Longrightarrow \lambda \underline{x} \in \mathcal{E}$$

such that ($\lambda, \mu \in \mathbb{K}$ and $\underline{x}, \underline{y} \in \mathcal{E}$)

$$\lambda(\underline{x} + \underline{y}) = \lambda \underline{x} + \lambda \underline{y}$$

$$(\lambda + \mu)\underline{x} = \lambda \underline{x} + \mu \underline{x}$$

$$\lambda(\mu \underline{x}) = (\lambda \mu)\underline{x}$$

It follows $\lambda \underline{0} = \underline{0}$ for any $\lambda \in \mathbb{K}$ and $0\underline{x} = 0$ for any $\underline{x} \in \mathcal{E}$. Note that the field $\mathbb{K}$ is itself a linear space. From now on, no special symbols will be used to identify vectors, *i.e.* x will be everywhere used in place of $\underline{x}$.

1.1 Linear independence

Let $\{x_i\}_{i=1}^n \subset \mathcal{E}$ be a set of vectors. The vectors can be combined to give other vectors,

$$y = \sum_{i=1}^{N} \lambda_i x_i$$

for given $\lambda_i \in \mathbb{K}$. We call the r.h.s of this expression a linear combination of the vectors x_i. The vectors $\{x_i\}_{i=1}^n$ are said to be linearly independent (of each other) if

$$\sum_{i=1}^{N} \lambda_i x_i = 0 \Longrightarrow \lambda_i = 0 \ \forall i = 1, n$$

It is clear that any subset of a linearly independent set is linearly independent. If n is the maximum number of linearly indepedent vectors, the space $\mathcal{E}$ is said to have dimension n, and we write $n = dim\mathcal{E}$; such n can be inifinite.

If $\mathcal{E}$ is a $n-$dimensional space any set of n linearly independent vectors is called a *basis* of $\mathcal{E}$. It follows from the definition that:

- every vector of $\mathcal{E}$ can be represented as a linear combination of basis vector,

- the coefficients of this linear combination are uniquely determined by the basis vectors themselves.

For the first result, just notice that $\forall x_0 \in \mathcal{E}$ the set $\{x_0\} \bigcup \{x_i\}_{i=1}^n$ is linearly depedent, since $\{x_i\}_{i=1}^n$ is maximal. Then, there exist non-vanishing $\lambda_i \in \mathbb{K}$ such that $\sum_{i=0}^n \lambda_i x_i = 0$. If $\lambda_0 = 0$, it follows $\lambda_i = 0$ for $i = 1, n$ since $\{x_i\}_{i=1}^n$ is a basis, *i.e.* $\lambda_i = 0$ for any i, in contradiction with the hypothesis. If $\lambda_0 \neq 0$ we get the desired result, $x = \sum_{i=1}^n c_i x_i$ where $c_i = -\lambda_i/\lambda_0$. For the second result, suppose $x = \sum_{i=1}^n c_i x_i = \sum_{i=1}^n c'_i x_i$; it follows $0 = \sum_{i=1}^n (c_i - c'_i) x_i$, *i.e.* $c_i = c'_i$ for $i = 1, n$. Basis vectors are usually denoted as e_i, in place of x_i.

1.2 Subspace

Let be $V \subset \mathcal{E}$. V is a subspace if it is closed with respect to the sum and the product of $\mathcal{E}$,

$$x, y \in V \implies x + y \in V$$

$$x \in V, \lambda \in \mathbb{K} \implies \lambda x \in V$$

A necessary condition for V to be a subspace is that $0 \in V$.

1.3 Dual space

Let f be a linear *functional*, *i.e.* a map from $\mathcal{E}$ to $\mathbb{K}$

$$f : \mathcal{E} \longrightarrow \mathbb{K}$$

$$\mathcal{E} \ni x \rightsquigarrow f(x) \in \mathbb{K}$$

which is “compatible” with the linear structure in $\mathcal{E}$,

$$f(x + y) = f(x) + f(y)$$

$$f(\lambda x) = \lambda f(x)$$

Note that on the r.h.s. the sum and product are those defined in the field $\mathbb{K}$, whereas on the l.h.s. they are those of $\mathcal{E}$.

Let $\mathcal{E}^*$ be the set of all linear functionals above with two operations, a sum $(f + g)$ and a product for elements of $\mathbb{K}$ (λf), defined pointwise

$$(f + g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda^* f(x)$$

i.e. $\forall x \in \mathcal{E}$. With these operations $\mathcal{E}^*$ becomes a linear space, closely related to $\mathcal{E}$, called the *dual space* of $\mathcal{E}$. To emphasize that $f \in \mathcal{E}^*$ is a *linear* functional we write

$$f(x) =< f, x >$$

With this notation $< f + g, x >=< f, x > + < g, x >$, $< \lambda f, x >= \lambda^* < f, x >$ and so on.

Let $\mathcal{E}$ be an $n-$dimensional vector space, and $\{e_i\}_{i=1}^n$ one of its basis. A linear functional is known if and only if its "action" on the basis elements is known, since $\forall x \in \mathcal{E}$, we may write $x = \sum_{i=1}^n c_i e_i$ and

$$< f, x >= \sum_{i=1}^{n} c_i < f, e_i >$$

i.e. it suffices to know the values $< f, e_i >$ for $i = 1, n$.

Given the basis $\{e_i\}_{i=1}^n$ let's then define the following functionals f_i $(i = 1, n)$

$$f_i(e_j) = \delta_{ij}$$

It follows that if $\mathcal{E} \ni x = \sum_{i=1}^n c_i e_i$, the coefficients c_i are simply given by $c_i = f_i(x)$, however we choose $x \in \mathcal{E}$. The set $\{f_i\}_{i=1}^n$ is linearly independent, for $\sum_{i=1}^n \lambda_i f_i = 0$ means $\sum_{i=1}^n \lambda_i f_i(x) = 0$ $\forall x \in \mathcal{E}$ and we may select $x = e_j$ to get $\lambda_j = 0$ $(j = 1, n)$. It also follows that the set is maximal, since if $\{f_0\} \bigcup \{f_i\}_{i=1}^n$ were linearly indepedent there would not exist a non-vanishing set of coefficients λ_i such that $\sum_{i=0}^n \lambda_i f_i = 0$. However, for $\lambda_0 = 1$ and $\lambda_i = -f_0(e_i)$ we have, $\forall x \in \mathcal{E}$,

$$f_0(x) - \sum_{i=1}^{n} f_0(e_i) f_i(x) = f_0(x) - f_0(\sum_{i=1}^{n} c_i e_i) = 0$$

since for $c_i = f_i(x)$ we have $\sum_{i=1}^n c_i e_i = x$.

2 Scalar product

Given a vector space $\mathcal{E}$, an application u from $\mathcal{E} \times \mathcal{E}$ to $\mathbb{K}$ (what is called a *2-form*)

$$u : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{K}$$

$$\mathcal{E} \times \mathcal{E} \ni x, y \rightsquigarrow u(x, y) \in \mathbb{K}$$

is said to be *sequilinear* if it is linear in the second variable and antilinear in the first, namely

$$u(\lambda x, y) = \lambda^* u(x, y)$$

$$u(x+y,z) = u(x,z) + u(y,z)$$

$$u(x,\lambda y) = \lambda u(x,y)$$

$$u(z,x+y) = u(z,x) + u(z,y)$$

The application u is a *scalar* (or *inner*) *product* if the following properties hold

$$u(x,y) = u(y,x)^*$$

$$u(x,x) \geq 0,\ u(x,x) = 0 \Longrightarrow x = 0$$

The first represents the *hermitian symmetry* and guarantees that $u(x,x)$ is a real number; with the second property, one asks that the scalar product satisfies *positivity* and *non-degeneracy.* For such u the following notation is widespread

$$(x|y) = u(x,y)$$

and in the following we adopt such convention.

From its very definition, it follows the *Schwarz inequality*,

$$|(x|y)|^2 \leq (x|x)(y|y)$$

Proof. Let us consider the special case $(x|x) = (y|y) = 1$. For any $\lambda \in \mathbb{K}$ it holds

$$\begin{aligned} 0 \ \leq \ & (x-\lambda y|x-\lambda y) = (x|x) + |\lambda|^2(y|y) - \lambda^*(y|x) - \lambda(x|y) \\ & = 1 + |\lambda|^2 - \lambda^*(y|x) - \lambda(x|y) \end{aligned}$$

Choosing $\lambda = (y|x)$ we get the desired result. In general, x and y do not satisfy $(x|x) = (y|y) = 1$; however, if $(x|x)(y|y) > 0$, the vectors $x' = x/\sqrt{(x|x)}$ and $y' = y/\sqrt{(y|y)}$ are such that $(x'|x') = (y'|y') = 1$ and we get $|(x'|y')|^2 \leq 1$ *i.e.* $|(x|y)|^2 \leq (x|x)(y|y)$. Of course, if $(x|x) = 0$, it follows $x = 0$ and the above inequality holds again.

Note that the above inequality becomes an equality *if and only if* the vectors x and y are linearly dependent, *i.e.* $x = \lambda y$.

2.1 Topology and limits

Once a scalar product is given in the vector space $\mathcal{E}$, a *norm* naturally follows. Remember that a norm (usually denoted as $\|\cdot\|$) is a functional on $\mathcal{E}$ with the properties

$$\|x\| \geq 0, \|x\| = 0 \Longrightarrow x = 0$$

$$\|\lambda x\| = |\lambda|\,\|x\|$$

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

It is easily verified that $\|x\|^2 := (x|x)$ does define a norm; for the third condition -the so-called *triangle inequality*, for obvious reasons- , just notice that

$$\begin{aligned}(x+y|x+y) &= (x|x)+(y|y)+2\mathrm{Re}(x|y) \leq (x|x)+(y|y)+2|(x|y)| \\ &\leq (x|x)+(y|y)+2(x|x)^{1/2}(y|y)^{1/2} = (\sqrt{(x|x)}+\sqrt{(y|y)})^2\end{aligned}$$

where in the last step we have used the Schwarz inequality above. The norm itself induces a *distance* in $\mathcal{E}$, *i.e.* an application d from $\mathcal{E}\times\mathcal{E}$ to $\mathbb{R}$ with the properties

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

$$d(x,y) \geq 0,\ d(x,y) = 0 \Longrightarrow x = y$$

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

Indeed, it is easily verified that $d(x,y) = \|x-y\|$ is a proper definition of distance, one which is translationally invariant. Note, however, that a distance can be defined without the help of an underlying linear structure; in such cases one talks about *metric spaces*. Linear spaces with a norm are called normed space or, apart from some subtle distinction to be detailed below, *Banach spaces*. With the same token, linear spaces with a scalar product are called *Hilbert spaces*.

A distance, such as the one induced by the norm above, allows one to define a very simple topology (a *metric topology*), *i.e.* to introduce the concept of convergent sequences and their limits. This is extremely important in inifinite-dimensional spaces, in that it is the only way to give a meaning to linear combinations of infinite number of vectors. Remember that, with the distance above, $\{x_n\}_{n\in\mathbb{N}}$ is said to converge to x - and we correspondingly write $x = \lim_n x_n$ or $x_n \to x$ - if

$$\forall\epsilon > 0\, \exists N_\epsilon \text{ such that}\, n > N_\epsilon \Longrightarrow \|x - x_n\| < \epsilon$$

Notice that if $\{x_n\}_{n\in\mathbb{N}}$ is a convergent sequence, then

$$\forall\epsilon > 0\, \exists N_\epsilon \text{ such that}\, n, m > N_\epsilon \Longrightarrow \|x_m - x_n\| < \epsilon$$

i.e. the sequence "gets denser" as n increases. Indeed, for a given ϵ let be N_ϵ such that $\|x - x_n\| < \epsilon/2$; then for $n, m > N_\epsilon$

$$\|x_m - x_n\| = \|(x_m - x) + (x - x_n)\| < \|x_m - x\| + \|x - x_n\| < \epsilon$$

A sequence which gets denser in the sense above is called a *Cauchy sequence*, and the condition above is known as Cauchy criterion. It is much easier to check as it does *not* require any prior knowledge of the supposed limit x. We have thus shown that any convergent sequence is a Cauchy sequence; if the converse is true, *i.e.* that *any Cauchy sequence is convergent*, we say that the space is *complete*. The name is very suggestive, in that if the space is "complete" (without "holes") any sequence which gets denser must converge to some point in the space. The names Banach ($\mathcal{B}$) / Hilbert ($\mathcal{H}$) spaces given above actually refer to normed / inner-product spaces which are complete in the metric topology induced by the norm.

We say that $M \subset \mathcal{E}$ is *closed*, and write $M = \bar{M}$, if *any convergent* sequence $\{x_n\}_{n\in\mathbb{N}} \subset M$ has limit in M, *i.e.* if

$$\{x_n\}_{n\in\mathbb{N}} \subset M,\, \lim_n x_n = x \Longrightarrow x \in M$$

This concept basically means that a closed subset contains "its own limits". It follows from the definition that a closed subset of a complete space is complete. Indeed, let be $\{x_n\}_{n\in\mathbb{N}} \subset M \subset \mathcal{E}$ a Cauchy sequence; it converges since $\mathcal{E}$ is complete and its limit $x \in M$, since M is closed. In general, for any given M one can realize the *closure* of M, written as $\bar{M}$, by adding the limits of (convergent!) sequences $\{x_n\}_{n\in\mathbb{N}} \subset M$. In this context, $M \subset \mathcal{E}$ is said to be *dense* in $\mathcal{E}$ if $\bar{M} = \mathcal{E}$.

Once a topology is introduced, the linear space $\mathcal{E}$ above becomes a *topological vector space*[1]. Along with convergent sequences, the notion of continuos maps can be introduced for applications between such kind of spaces. For instance, let us consider two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ and a map $f : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$. f is said to be continuos (in x) if

$$\forall \epsilon > 0\, \exists \delta_\epsilon \text{ such that } \left\|x - x'\right\|_1 < \delta_\epsilon \Longrightarrow \left\|f(x) - f(x')\right\|_2 < \epsilon$$

Equivalently, this $\epsilon - \delta$ definition means that for *any* convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ with limit x the sequence $f(x_n)$ is convergent and its limit is $f(x)$. Note that the first norm is that defined in $\mathcal{H}_1$ whereas the second one is that of $\mathcal{H}_2$. It is not hard to show that:

- the sum and the product by scalars are continuos operations in $\mathcal{H}$ (luckily!);
- the scalar product is a continuos operation from $\mathcal{H}$ to $\mathbb{K}$

For the second property, note that for a fixed y and a convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ with limit x we have for $n \to \infty$

$$|(y|x) - (y|x_n)| = |(y|x - x_n)| \le \|y\|\, \|x - x_n\| \to 0$$

that is, $\lim_n (y|x_n) = (y|x) = (y|\lim_n x_n)$.

For a *linear* map $f : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ the following sets are important. The first is a subset of $\mathcal{H}_1$, the *kernel*,

$$\mathsf{Ker} f = \{x \in \mathcal{H}_1 | f(x) = 0\}$$

and the second is a subset of $\mathcal{H}_2$, the *image*,

$$\mathsf{Im} f = \{y \in \mathcal{H}_2 |\, y = f(x), x \in \mathcal{H}_1\}$$

It is not difficult to verify that both sets are linear, and that if f is continuos $\mathsf{Ker} f$ is closed, *i.e.* it is a true subspace of $\mathcal{H}_1$. Notice, however, that the continuity of f does not guarantee any property on $\mathsf{Im} f$.

[1] From now on, if $\mathcal{E}$ is such a topological space the notion of *subspace* will be used to mean any *closed,* linear subset of $\mathcal{E}$.

2.2 Orthogonality (geometry)

Let $\mathcal{H}$ be a Hilbert space. $x, y \in \mathcal{H}$ are said to be orthogonal to each other (one often uses the symbol $x \perp y$) if

$$(x|y) = 0$$

This simple definition actually entails a very important geometric result, the famous "Pitagora's theorem"

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

which holds for any two orthogonal vectors x, y.

Let $V \subset \mathcal{H}$ be an arbitrary subset. We call the orthogonal complement of V (denoted as $V^\perp$) the set

$$V^\perp = \{x \in \mathcal{H} | \, (x|y) = 0 \, \forall y \in V\}$$

It is easily verified that

- $V^\perp$ is a *linear* space, no matter whether V is linear or not.
- $V^\perp \bigcap V = \{0\} \bigcap V$, *i.e.* the only element in common between V and $V^\perp$ is 0 (if $0 \in V$).
- $\{0\}^\perp = \mathcal{H}$ and $\mathcal{H}^\perp = \{0\}$
- $V^\perp$ is *closed*, *i.e.* $V^\perp$ is a *subspace*.
- $V \subset W \Longrightarrow W^\perp \subset V^\perp$
- $V \subset V^{\perp\perp}$

For the fourth property, *e.g.,* let $\{x_n\}_{n \in \mathbb{N}} \subset V^\perp$ be convergent, $x_n \to x$; it follows that, for any $y \in V$, $0 = \lim_n (x_n|y) = (x|y)$, that is $x \in V^\perp$.

Finally, let $V \subset \mathcal{H}$ be an arbitrary subset, as before, and let $\overline{V}$ be the *closed linear space spanned* by the elements of V (sometime denoted as *span*$\{V\}$). Then, an important result is

$$V^{\perp\perp} = \overline{V}$$

This follows from the last property above, once it is shown that if V is a *subspace* (a condition which we can now write as $V = \overline{V}$) $V^{\perp\perp} \subset V$. In practice,

$$V \, \text{subspace} \implies V = V^{\perp\perp}$$

This result is not really trivial, and won't be proved here. Note, however, that V is *necessarily* closed if it has to coincide with an orthogonal complement of some set. This result is useful for the following

Projection theorem. Let $V \subset \mathcal{H}$ a subspace and $V^\perp$ its orthogonal complement. Then, for every $x \in \mathcal{H}$ there exists one and only one pair of vectors $x_V \in V$ and $x'_V \in V^\perp$ such that

$$x = x_V + x'_V$$

or, in other words, $\mathcal{H} = V \oplus V^\perp$ where $\oplus$ stands for the (inner) direct sum[2].

Proof. Let $V \neq \{0\}$, otherwise $\{0\}^\perp = \mathcal{H}$ and the proof is trivial. Let be $S = V \oplus V^\perp \subset \mathcal{H}$. S is closed since it is the direct sum of two subspace (*i.e. closed* set) and the sum is continuos; then S is a subspace. From $S \supset V, V^\perp$ it follows $S^\perp \subset V^\perp, V^{\perp\perp} = V$, *i.e.* $S^\perp \subset V \bigcap V^\perp = \{0\}$, or also $S = S^{\perp\perp} \supset \{0\}^\perp = \mathcal{H}$.

Finally, the following results are often very useful in applications. Let $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ be an orthonormal set[3], *i.e.* such that $(u_n|u_m) = \delta_{nm}$. For any c-sequence $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ which is square-summable (*i.e.* such that $\sum_{n\in\mathbb{N}} |c_n|^2 < \infty$) the vector $x = \sum_{n\in\mathbb{N}} c_n u_n$ is well defined[4] and

$$c_n = (u_n|x)$$

and

$$||x||^2 = \sum_{n\in\mathbb{N}} |c|^2$$

$\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ defines a subspace of elements which can be obtained from the (closed) set of the linear combination of its elements, *i.e.* the space $V = \{u_n\}^{\perp\perp}_{n\in\mathbb{N}}$. Applying the projection theorem above to the space V and noticing that for any $x \in \mathcal{H}$ $||x||^2 \geq ||x_V||^2$, we arrive at

$$\sum_{n\in\mathbb{N}} |(u_n|x)|^2 \leq ||x||^2$$

We say that the set is *complete* if $\{u_n\}^\perp_{n\in\mathbb{N}} = \{0\}$. For any such set it easily follows that: (i) any vector x can be written in the form $x = \sum_n u_n(u_n|x)$ (hint: $\{u_n\}^{\perp\perp}_{n\in\mathbb{N}} = \mathcal{H}$); (ii) $||x||^2 = \sum_n |(u_n|x)|^2$.

---

[2] Given V and V' two linear spaces on the same field the direct sum $V \oplus V'$is the cartesian product $V \times V'$ equipped with a linear structure "induced" by the linear structures of V and V'. In other words, $V \oplus V'$ is the set of pairs (x, y) with an "exterior" sum defined to be $(x_1, y_1)+(x_2, y_2) := (x_1+x_2, y_1+y_2)$ and $\lambda(x, y) := (\lambda x, \lambda y)$. Note that the zero vector is given by $(0, 0)$. If the spaces V and V' are subspaces of the same vector space $\mathcal{E}$ *and* $V \bigcap V' = \{0\}$, then $V \oplus V'$ can be *identified* with the ordinary sum as defined in $\mathcal{E}$. For, consider the *linear* map $t : V \oplus V' \to \mathcal{E}$ defined by $(x, x') \rightsquigarrow t(x, x') = x + x'$ with the property $t((x, x') + (y, y')) = t(x, x') + t(y, y') = x + x' + y + y'$. The map is iniettive, since $t(x, x') = x + x' = 0$ for $x \in V$ and $x' \in V'$ implies $x = -x' \in V, V'$, *i.e.* $x = x' = 0$ or, equivalently, (x, x') must be the zero of $V \oplus V'$. The linear space $\mathsf{Im}t$ is actually $\mathsf{span}\{V \bigcup V'\}$, and one identifies the inner sum as $t \circ \oplus$.

[3] We have assumed that the orthornormal set is *numerable.* This requires an additional property on $\mathcal{H}$ which has to be *serapable.* $\mathcal{H}$ is said to be serapable if there exists a *numerable dense* set $A \subset \mathcal{H}$, *i.e.* $\mathsf{card}A = \aleph_0$ and $\bar{A} = \mathcal{H}$. It can be shown, then, that $\mathcal{H}$ is separable if and only if any orthonormal set is numerable.

[4] The sequence of partial sums $s_n = \sum_{k\leq n} c_k u_k$ is a Cauchy sequence, since for $m \geq n$ $||s_n - s_m||^2 \equiv \sum_{k=n+1}^{m} |c_k|^2 = |S_n - S_m|$ where S_n is the *converging* sequence of complex number $S_n = \sum_{k\leq n} |c_k|^2$.

2.3 Dual space

The presence of a topology in the Hilbert space suggests to re-define the dual space in such a way to include only those functional which are continuos in the above topology. We say then $f \in \mathcal{H}^*$ if $f : \mathcal{H} \to \mathbb{K}$ is linear and

$$f(x_n) \to f(x) \,\text{whenever}\; x_n \to x$$

In other words, with the help of linearity,

$$\forall \epsilon > 0 \,\exists \delta_\epsilon \text{ such that } \|x\| < \delta_\epsilon \Longrightarrow |f(x)| < \epsilon$$

Thus, continuity has to be checked just for $x = 0$ and if holds it does for any $x \in \mathcal{H}$. Note that this condition is also equivalent to

$$\exists M > 0 \,\text{such that } |f(x)| < M\,\|x\| \;\forall x \in \mathcal{H}$$

The proof is simple: if the first holds, consider $\epsilon > 0$ and $x' = x\delta_\epsilon/2\,\|x\|$, x arbitrary; since $\|x'\| = \delta_\epsilon/2 < \delta_\epsilon$ we get $|f(x)| < (2\epsilon/\delta_\epsilon)\,\|x\|$, *i.e.* we have found $M = 2\epsilon/\delta_\epsilon > 0$ for the second condition. Conversely, if the second condition holds, consider $\epsilon > 0$; then, for $\|x\| < \epsilon/M := \delta_\epsilon$ we get $|f(x)| < \epsilon$.

The second condition is useful for introducing a *norm* in $\mathcal{H}^*$,

$$\|f\|_\infty = \inf\{M > 0, |f(x)| < M\,\|x\|\; x \in \mathcal{H}\}$$

such that we can write $|f(x)| \leq \|f\|_\infty \|x\|$ or, equivalently, $| < f, x > | \leq \|f\|_\infty \|x\|$. That $\|\cdot\|_\infty$ here defined is actually a norm is not hard to show .

As an important example of a *continuos* linear functional, let be $y \in \mathcal{H}$ and consider $f_y : \mathcal{H} \to \mathbb{K}$ defined as

$$< f_y, x >:= (y|x)$$

Such f_y is indeed (linear and) continous, since $| < f_y, x > | = |(y|x)| \leq \|y\|\,\|x\|$, and $\|f\|_\infty \leq \|y\|$; actually, we can be more specific, since from $| < f_y, y > | = |(y|y)| = \|y\|^2 \leq ||f_y||||y||$ it follows $||f||_\infty \geq ||y||$, *i.e.* $||f_y||_\infty \equiv ||y||$. This functional can be defined for any given $y \in \mathcal{H}$, that is we can consider the map $T : \mathcal{H} \to \mathcal{H}^*$ defined by

$$y \rightsquigarrow Ty = f_y$$

or equivalently

$$< Ty, x >= (y|x)\,\forall x \in \mathcal{H}$$

This map is linear and continuos, since

$$< T(y_1 + y_2), x >= (y_1 + y_2|x) =< Ty_1, x > + < Ty_2, x >$$

$$< T(\lambda y), x >= (\lambda y|x) = \lambda^*(y|x) =< \lambda Ty, x >$$

and *isometric*, $||Ty||_\infty = ||y||$. This implies that T is *invertible,* since for $Ty = 0$ necessarily $y = 0$, and thus there exists a map $T^{-1} : \mathcal{H}^* \supset \mathsf{Im}T \to \mathcal{H}$ such that $T^{-1}Ty \equiv y$ for any $y \in \mathcal{H}$. A very important result is the following

Riesz theorem. T defined above is *suriettive,* $\mathsf{Im}T = \mathcal{H}^*$, *i.e.* $\forall f \in \mathcal{H}^*$ there exist $y_f \in \mathcal{H}$ such that

$$< f, x >= (y_f|x) \, \forall x \in \mathcal{H} \quad \text{and} \quad ||f||_\infty = ||y_f||$$

In other words T is an isometric isomorphism between *Banach* spaces. Thanks to it, we can equip the space $\mathcal{H}^*$ with a scalar product,

$$(\alpha|\beta)_* := (T^{-1}\alpha|T^{-1}\beta)$$

(where on the r.h.s. the scalar product is that defined on $\mathcal{H}$) and in this way T becomes an isomorphism between *Hilbert* spaces. $\mathcal{H}^*$ can thus be *identified* with $\mathcal{H}$.

Before leaving this section, it is worth noticing that with the above definition of the dual space one can introduce a topology in $\mathcal{H}$ which *differs* from the usual norm topology[5]. This is the *weak* topology and can be defined as follows. We say that the $\{x_n\} \subset \mathcal{H}$ converges to x if for *any* $f \in \mathcal{H}^* < f, x_n > \to < f, x >$, and we write $x = w - \lim_n x_n$. Clearly, if $\{x_n\} \subset \mathcal{H}$ converges "properly" (*i.e.* in norm) to x it also weakly converges to the same limit, since in that case $| < f, x - x_n > | \leq ||f||_\infty ||x - x_n|| \to 0$. The converse is not true: for instance, consider a sequence of orthonormal vectors, $\{e_n\}_{n \in \mathbb{N}}$, where $(e_n|e_m) = \delta_{nm}$; this sequence is not converging in norm, since $||e_n - e_m|| = \sqrt{2}\delta_{nm}$, but $\sum_n |(x|u_n)|^2 \leq ||x||^2$ implies $|(x|u_n)| \to 0$ for any x, *i.e.* $w - \lim_n e_n = 0$. The topology is weaker than the usual one since it contains *less* (more) open (closed) sets.

3 Tensor product

There is a further important structure between linear spaces, one that allows us to build vectors of a linear space by "multiplying" vectors of other spaces. This kind of multiplication, called *tensor product* and denoted as $\otimes$, is required to behave similarly to an ordinary product

$$x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$$
$$(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$$
$$x \otimes (\lambda y) = (\lambda x) \otimes y = \lambda(x \otimes y)$$

i.e. to be a bilinear (in general, *multilinear*) map t from $\mathcal{E}_1 \times \mathcal{E}_2$ to some other linear space ($t : \mathcal{E}_1 \times \mathcal{E}_2 \to \mathcal{E}$), whose action is simply denoted as $t(x, y) = x \otimes y$.

A general, though rather abstract formulation, is the following[6]. Let us consider three vector spaces $\mathcal{E}_i (i = 1, 3)$ and let F be a multilinear function from $\mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$ to $\mathbb{K}$

$$F : \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3 \to \mathbb{K}$$

[5] Here the scalar product is irrelevant, and the result holds for any Banach space.

[6] The use of three vector spaces allows sufficient generality and remove "singularities" which may appear in the case of two vector spaces.

$$(x_1, x_2, x_2) \rightsquigarrow F(x_1, x_2, x_3)$$

that is a function which is *linear* in each "position". Clearly, $F(0, x_2, x_3) = F(x_1, 0, x_3) = F(x_1, x_2, 0) = 0$.

Then, under quite general conditions, it can be proved that there exists a linear space $\mathcal{E}$ and a multilinear map from $\mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$ to $\mathcal{E}$,

$$t : \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3 \to \mathcal{E}$$

$$(x_1, x_2, x_2) \rightsquigarrow t(x_1, x_2, x_3)$$

such that *any* multilinear function F above can be written as

$$F = f \circ t$$

where f is a linear functional on $\mathcal{E}$, determined only by F. In other words, it is possible to "decompose" any multilinear function as composition of two maps: a *multilinear* map to some vector space $\mathcal{E}$ and a *linear* functional on this vector space. The above map t, called tensor product, and the corresponding space $\mathcal{E}$ are "universal" in the sense that if another pair $(t', \mathcal{E}')$ is found with the same properties then there exists an isomorphism between the two spaces $\mathcal{E}$ and $\mathcal{E}'$.

As already said before, $t(x_1, x_2, x_3)$ is usually denoted as $x_1 \otimes x_2 \otimes x_3$. These elements $x_1 \otimes x_2 \otimes x_3 \in \mathsf{Im} t \subset \mathcal{E}$ are called *decomposable* or simply *product elements.* Any linear combination ($(x_i, x_j, x_k) \in \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$)

$$\sum_{i,j,k} c_{ijk} x_i \otimes x_j \otimes x_k$$

may or not belong to $\mathsf{Im} t$, but nevertheless they span the whole linear space needed for the above relation to hold. The space $\mathsf{span}\{\mathsf{Im} t\}$, which is in general a subspace of $\mathcal{E}$, is called the tensor product of the spaces $\mathcal{E}_i$, written as

$$\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3 = \mathsf{span}\{\mathsf{Im} t\}$$

We can identify such space with the space $\mathcal{E}$ above[7] and the map t with $\otimes$.

An important result is

Theorem If $\{e_{1,i}\}_{i=1}^{n_1} \subset \mathcal{E}_1$, $\{e_{2,j}\}_{j=1}^{n_2} \subset \mathcal{E}_2$ and $\{e_{3,k}\}_{k=1}^{n_3} \subset \mathcal{E}_3$ are three linearly independent sets, the set

$$\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3 \supset \{e_{1,i} \otimes e_{2,j} \otimes e_{3,k}\}_{i=1,n_1;,j=1,n_2;k=1,n_3}$$

is linearly independent.

[7] In general, one starts with some "large" space $\mathcal{F}$ which allows the "transfer" of the multilinear functions, $F = f \circ t$. Once this is found, the subspace $\mathsf{span}\{\mathsf{Im} t\} \subset \mathcal{F}$ is the linear space of the theorem above.

Proof. Let us consider the linear combination

$$\sum_{i,j,k} c_{i,j,k} e_{1,i} \otimes e_{2,j} \otimes e_{3,k} = u$$

The condition $u = 0$ implies $f(u) = 0$ for any linear functional f on $\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3$. Let us take, then, three functionals (on $\mathcal{E}_1$,$\mathcal{E}_2$ and $\mathcal{E}_3$, respectively) satisfying

$$\alpha(e_{1,m}) = \delta_{im}\, \beta(e_{2,m}) = \delta_{jm} \text{ and} \gamma(e_{3,m}) = \delta_{km}$$

for some indeces (i, j, k). *They are well defined since the sets* $\{e_{i,j}\}_{j=1}^{n_i} \subset \mathcal{E}_i$ *are linearly independent*[8]. The function

$$F(x_1, x_2, x_3) = \alpha(x_1)\beta(x_2)\gamma(x_3)$$

is a multilinear applications of $\mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$ to $\mathbb{K}$. According to the result above, there exits a linear functional f on $\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3$ such that

$$f(x_1 \otimes x_2 \otimes x_3) = F(x_1, x_2, x_3) = \alpha(x_1)\beta(x_2)\gamma(x_3)$$

It is not hard to check that $f(u) = c_{i,j,k}$, since

$$\begin{aligned} f(u) &= \sum_{i',j',k'} c_{i',j',k'} f(e_{1,i'} \otimes e_{2,j'} \otimes e_{3,k'}) = \\ &= \sum_{i',j',k'} c_{i',j',k'} \alpha(e_{1,i'})\beta(e_{2,j'})\gamma(e_{3,k'}) = \sum_{i',j',k'} c_{i',j',k'} \delta_{ii'}\delta_{jj'}\delta_{kk'} \end{aligned}$$

Therefore $u = 0$ implies $c_{i,j,k} = 0$, and this holds for any arbitrary choice of indices (i, j, k).

It follows that if the spaces $\mathcal{E}_1$,$\mathcal{E}_2$ and $\mathcal{E}_3$ have dimension n_1, n_2 and n_3, respectively, $\dim(\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3) = n_1 n_2 n_3$, and the theorem above shows how to build a basis in the tensor product space, given a basis in each defining space.

If $\mathcal{E}_1$,$\mathcal{E}_2$ and $\mathcal{E}_3$ have some structure the tensor product $\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3$ can be equipped with the same structure. For example, let us consider three Hilbert spaces $\mathcal{H}_1$,$\mathcal{H}_2$ and $\mathcal{H}_3$; the space $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ can be equipped with a scalar product, defined to be

$$(x_1 \otimes x_2 \otimes x_3 | y_1 \otimes y_2 \otimes y_3) := (x_1|y_1)(x_2|y_2)(x_3|y_3)$$

where on the r.h.s. we make use of the scalar product defined in each space $\mathcal{H}_i$. This formula actually defines a functional on $\mathsf{Im}\otimes \times \mathsf{Im}\otimes$; however, it can be linearly extended to the whole tensor product space, namely if $u, v \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ then

$$u = \sum_i c_i x_{1i} \otimes x_{2i} \otimes x_{3i}$$

[8]They are also *uniquely* defined if the sets are *basis* of the spaces. In general, a basis can always be built which has the above vectors among its basis elements; in that case, the functionals above can take arbitrary values on the remaining basis elements, but this does not matter for the proof below.

$$v = \sum_i d_i y_{1i} \otimes y_{2i} \otimes y_{3i}$$

and, therefore, we may set

$$(u|v) = \sum_{i,j} c_i^* d_j (x_{1i} \otimes x_{2i} \otimes x_{3i} | y_{1j} \otimes y_{2j} \otimes y_{3j})$$

More subtle questions concern topological properties. For that purpose, for instance, $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ is best defined as the closure of $\mathsf{span}\{\mathsf{Im}\otimes\}$, in such a way that it always includes its limits.

4 Operators: basic definitions

Let A be a *linear* map between two vector spaces, $A : \mathcal{E} \to \mathcal{F}$, $x \rightsquigarrow A(x)$. To emphasize that A is linear, and distinguish it from more general maps, one usually writes $A(x) = Ax$. A is called an *operator* from $\mathcal{E}$ to $\mathcal{F}$. For any such A, there are two important spaces related to it

$$\mathcal{E} \supset \mathsf{Ker} A = \{x \in \mathcal{E} |\, Ax = 0\}$$

$$\mathcal{F} \supset \mathsf{Im} A = \{y \in \mathcal{F} |\, y = Ax,\, x \in \mathcal{E}\}$$

It is easy to verify that both of them are *linear*. In the following we suppose that A is defined on the whole space $\mathcal{E}$ (something which is not always true in practical applications). If $\mathcal{E}$,$\mathcal{F}$ are topological vector spaces (Banach spaces in the following), A is said to be continous if, for *any* convergent sequence $x_n \to x$, we have $Ax_n \to Ax$. Equivalently, thanks to the linearity,

$$\forall \epsilon > 0\, \exists \delta_\epsilon \text{ such that } \|x\| < \delta_\epsilon \Longrightarrow \|A(x)\| < \epsilon$$

or also

$$\exists M > 0 \,\text{such that }\, \|Ax\| < M\,\|x\| \;\forall x \in \mathcal{H}$$

The proof is analogous to the one already given for functionals, and allows one to introduce an operator *norm* $\|\cdot\|_\infty$, such that

$$||Ax|| \leq ||A||_\infty ||x||$$

namely

$$||A||_\infty = \inf\{M \in \mathbb{R} |\; ||Ax|| \leq M||x||, \forall x \in \mathcal{E}\}$$

It easily follows that if A is continous (or, equivalently, *bound*) $\mathsf{Ker}A$ is closed, *i.e.* it is a subspace of $\mathcal{E}$. $\mathsf{Im}A$ is more difficult to characterize and continutiy of A does not really help in this context: the properties of any sequence $\{y_n\} \subset \mathsf{Im}A$ can be in no way related to any property of the (possibly many) sequences $\{x_n\} \subset \mathcal{E}$ of elements such that $y_n = Ax_n$.

$\mathsf{Ker}A$ is important for the existence of the inverse operator. Remember that, in general, a map $f : \mathcal{H}_1 \to \mathcal{H}_2$ is *invertible* if and only if it is *iniettive*,

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

and its inverse f^{-1} will be a map $f^{-1} : \mathcal{H}_2 \supset \mathsf{Im}f \to \mathsf{H}_1$ where $\mathsf{Im}f = \{y \in \mathcal{H}_2 | y = f(x),\, x \in \mathcal{H}_1\}$ is the *image* of f. For a linear map $A : \mathcal{E} \to \mathcal{F}$ this condition can be re-written as

$$Ax = 0 \Rightarrow x = 0$$

that is, $A^{-1} : \mathcal{F} \supset \mathsf{Im}A \to \mathcal{E}$ exists if and only if the kernel is trivial, $\mathsf{Ker}A = \{0\}$. If this is case, then, A preserves linear independence, for if $\{x_i\} \subset \mathcal{E}$ is linearly independent it follows

$$0 = \sum_i c_i(Ax_i) = A(\sum_i c_i x_i) \Rightarrow \sum_i c_i x_i = 0$$

and the latter implies $c_i = 0$. This means that, in this case, $\mathsf{Im}A$ is isomorphic to $\mathcal{E}$ and, in particular, that $\mathsf{Im}A$ and $\mathcal{E}$ have the same dimensions. In general, if $\mathsf{Ker}A \neq \{0\}$ we can always split $\mathcal{E}$ as $\mathcal{E} = \mathsf{Ker}A \oplus V'$ where V' is the supplementary space[9], and then it is not hard to see that the *restriction* of A to V', what we call $A' : V' \to \mathcal{F}$, satisfies

$$\mathsf{Ker}A' = \{0\},\ \mathsf{Im}A' = \mathsf{Im}A$$

threby estabilishing an isomorphism between V' and $\mathsf{Im}A$. For finite-dimensional vector spaces, it follows

$$dim\mathcal{E} = n(A) + r(A)$$

where the "nucleus" $n(A)$ and the "rank" $r(A)$ are the dimensions of $\mathsf{Ker}A$ and $\mathsf{Im}A$, respectively.

When it comes to topological properties, the following result is useful: A is invertible *and* A^{-1} is continuos if and only if

$$\exists \mu > 0 \text{ such that } ||Ax|| \geq \mu ||x||$$

Indeed, if A^{-1} exists and is continuos, there exists $M > 0$ such that $||A^{-1}y|| < M||y||$, which for $y = Ax$ gives the condition above with $M = \mu^{-1}$. Conversely, from the condition above, A is invertible ($||Ax|| = 0 \implies ||x|| = 0$) and the inverse is bound, $||A^{-1}y|| \leq ||y||/\mu$.

Notice that continuity of A is irrelevant for this result; *if* A is continuos and satisfies the condition above then it is easily verified that $\mathsf{Im}A$ is *closed*: if $\{y_n\} \subset \mathsf{Im}A$ converges to y then $x_n = A^{-1}y_n$ is convergent because of the continuity of A^{-1}, and $A \lim_n x_n = \lim_n Ax_n = \lim_n y_n \in \mathsf{Im}A$, thanks to the continuity of A.

[9]For Hilbert spaces V' can be identified with $(\mathsf{Ker}A)^\perp$. In general, for every vector x one can define $V(x) = \{x' \in \mathcal{E} | x' = x + y,\, y \in \mathsf{Ker}A\}$. The set of all sets $V(x)$ is a linear space with sum $V(x) + V(x') := V(x + x')$ and product $\lambda V(x) := V(\lambda x)$ (notice that $V(x) = 0$ if and only if $x \in \mathsf{Ker}A$). The newly defined vector space is isomorphic to the above V'.

4.1 Product, commutators

We now focus on the cases where $A : \mathcal{E} \to \mathcal{E}$, *i.e.* on those operators working *within* the space $\mathcal{E}$. For such operators, not only the sum and the product can be defined to give operators of the same kind,

$$(A + B)x := Ax + Bx \ \ \forall x \in \mathcal{E}$$

$$(\lambda A)x := \lambda(Ax) \ \ \forall x \in \mathcal{E}$$

but also the (composition) product

$$(AB)x := A(Bx) \ \ \forall x \in \mathcal{E}$$

has the same property. For bound operators it also follows that the operators defined above are bound too, since for example

$$\|A(Bx)\| \leq \|A\|_\infty \|Bx\| \leq \|A\|_\infty \|B\|_\infty \|x\|$$

(this also implies $\|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty$). Therefore, bound operators form an *algebra*, usually denoted as $L(\mathcal{E})$. $L(\mathcal{E})$ is *non-commutative* since, in general, $AB \neq BA$, but is a topological algebra thanks to the presence of a norm. This allows one to define limits and series of operators $A \in L(\mathcal{E})$ such as

$$(1 - A)^{-1} = \sum_{n=0}^{\infty} A^n \ \text{ and } e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

and possibly to show that the latter is always convergent whereas the first converges provided $\|A\| < 1$. Usual algebraic rules hold for operators, though care has to be taken to handle non-commutavity. Notice, for example, that for non-commuting operators

$$e^{A+B} \neq e^A e^B$$

as may be guessed from the fact that when expanding the l.h.s. the operators A and B appear in all possible orders, whereas on the right the powers of A always precede those of B.

In this context the following property is interesting

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}$$

Here the crucial *commutator* $[A, B] := AB - BA$ is supposed to satisfy $[A, [A, B]] = [B, [A, B]] = 0$. The result can be proved as follows. Let us first show that when $[A, [A, B]] = 0$

$$e^{\lambda A} B e^{-\lambda A} = B + \lambda [A, B]$$

Indeed, let $F(\lambda) = e^{\lambda A} B e^{-\lambda A}$ and consider the derivative[10] $F'(\lambda) = e^{\lambda A} A B e^{-\lambda A} - e^{\lambda A} B A e^{-\lambda A} = e^{\lambda A}[A, B] e^{-\lambda A} = [A, B]$, which does no longer depend on λ (in the last step we have used $[A, [A, B]] = 0$) . This shows that

$$F(\lambda) = F(0) + \lambda F'(\lambda) = B + \lambda[A, B]$$

Let us now consider $G(\lambda) = e^{\lambda A} e^{\lambda B}$ and its derivative, $G'(\lambda) = AG(\lambda) + G(\lambda)B = (A + F(\lambda))G(\lambda) = (A + B + \lambda[A, B])G(\lambda)$. Upon integration (which is straightforward since for $H(\lambda) = A + B + \lambda[A, B]$ we have $[H(\lambda), H(\lambda')] = 0$)

$$G(\lambda) = G(0) e^{\lambda(A+B) + \frac{\lambda^2}{2}[A,B]} = e^{\lambda(A+B) + \frac{\lambda^2}{2}[A,B]}$$

Setting $\lambda = 1$ we get the final result

$$e^{A+B} = G(1) e^{-\frac{1}{2}[A,B]} = e^A e^B e^{-\frac{1}{2}[A,B]}$$

We see from the above example that the commutator $[A, B]$ plays a crucial role in the theory. It satisfies the following properties

$$[A, B] = -[B, A]$$

$$[A, \lambda B] = \lambda[A, B]$$

$$[A, B + C] = [A, B] + [A, C]$$

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$$

Therefore, it is a kind of (antisymmetric) product satisfying the special *Jacobi identity* written in the last equation. With this product $L(\mathcal{E})$ is a so-called *Lie algebra*[11].

4.2 Transpose operator

Let $\alpha \in \mathcal{E}^*$ be a continuous functional on $\mathcal{E}$, $A \in L(\mathcal{E})$ a bound operator and consider

$$\mathcal{E} \ni x \rightsquigarrow < \alpha, Ax > \in \mathbb{K}$$

This defines a linear, continuous map from $\mathcal{E}$ to $\mathbb{K}$, *i.e.* a functional $\beta = \alpha \circ A \in \mathcal{E}^*$, since

$$| < \alpha, Ax > | \leq ||\alpha||_\infty ||Ax|| \leq ||\alpha||_\infty ||A||_\infty \, ||x||$$

holds for any $x \in \mathcal{E}$ (notice further that it also follows $||\beta||_\infty \leq ||\alpha||_\infty ||A||_\infty$).

[10] Derivatives with respect to a parameter can be defined as usual, $A'(t) = lim_{h\to 0}(A(t+h) - A(t))/h$ and satisfy “standard” rules as long as they do not require commutativity.

[11] In general, a Lie algebra is a vector space equipped with a special product satisfying the rules given in the text. An example is $\mathbb{R}^3$ with the “vector product” . When starting from an ordinary algebra, the commutator defined above always provides a Lie product.

This is true for any $\alpha \in \mathcal{E}^*$ and thus we can consider the map *in* $\mathcal{E}^*$

$$\mathcal{E}^* \to \mathcal{E}^*$$

$$\alpha \rightsquigarrow \beta = \alpha \circ A$$

It is easily verified that such map is linear, *i.e.* it defines an operator A^t which is called the *transpose* of A, namely $A^t\alpha := \alpha \circ A$, and which by definition satisfies

$$< A^t\alpha, x >=< \alpha, Ax >$$

for any $x \in \mathcal{E}$. The map is also *continuous* (or equivalently $A^t \in L(\mathcal{E}^*)$),

$$||A^t\alpha|| \leq ||\alpha||_\infty ||A||_\infty$$

and obviously $||A^t||_\infty \leq ||A||_\infty$; more precisely, one can show that $||A^t||_\infty = ||A||_\infty$ as a consequence of the so-called "Hahn-Banach theorem".

5 Operator algebra in Hilbert spaces

When the vector space is a Hilbert space, the operator $L(\mathcal{H})$ has intriguing properties that come from the mentioned Riesz isomorphism between $\mathcal{H}$ and $\mathcal{H}^*$. This allows us to define an operator $A^\dagger$ which is just the Riesz image of the above transpose operator, *i.e.* a new operator on $\mathcal{H}$ closely related to A, which is known as the *adjont* of A. The presence of a *conjugation* operation (†) in $L(\mathcal{H})$ makes this algebra very attractive from many points of view. In the following, we shall focus on such structure[12].

5.1 Adjoint operator

Let then $\mathcal{E} = \mathcal{H}$ be a Hilbert space and consider A and A^t the operators defined above and $T : \mathcal{H} \to \mathcal{H}^*$ the Riesz isomorphism. The composition map

$$A^\dagger = T^{-1} \circ A^t \circ T$$

is linear and continuos, and thus defines a linear operator $A^\dagger$ in $\mathcal{H}$ (the adjoint of A) "closing" the diagram

$$\begin{array}{ccc} \mathcal{H} & \overset{T}{\longrightarrow} & \mathcal{H}^* \\ A^\dagger \ \Downarrow & & \downarrow \ A^t \\ \mathcal{H} & \overset{T^{-1}}{\longleftarrow} & \mathcal{H}^* \end{array}$$

[12]Notice that in this case, thanks to the presence of an inner product, one can formulate the results in terms of expressions involving scalar products between arbitrary vectors. For instance, A and B are equal to each other if and only if $(x|Ay) = (x|By)$ for any $x, y \in \mathcal{H}$, as it follows from $(A - B)y \in \mathcal{H}^\perp = \{0\}$ for any $y \in \mathcal{H}$. This is, however, a rather cumbersome way of proceeding and we try to avoid it as much as possble.

In other words, let y_α and y_β be the Riesz images of an arbitrary functional α and of $\beta = A^t\alpha$ (that is, such that $(y_\alpha|x) =< \alpha, x >$ and $(y_\beta|x) =< \beta, x >$ hold for any $x \in \mathcal{H}$) and consider the map $A^\dagger : y_\alpha \rightsquigarrow y_\beta$. This *defines* a linear, bound operator $A^\dagger$ satisfying

$$(A^\dagger y|x) = (y|Ax)$$

for any $x, y \in \mathcal{H}$, and

$$||A^\dagger||_\infty \leq ||A||_\infty$$

Note that the following properties hold

$$(A^\dagger)^\dagger = A$$

$$(A + B)^\dagger = A^\dagger + B^\dagger$$
$$(\lambda A)^\dagger = \lambda^* A^\dagger$$
$$(AB)^\dagger = B^\dagger A^\dagger$$
$$\mathsf{ker} A^\dagger = \mathsf{Im} A^\perp$$

as can be easily verified. For the second last property, for instance,

$$((AB)^\dagger y|x) = (y|ABx) \equiv (A^\dagger y|Bx) \equiv (B^\dagger A^\dagger y|x)$$

Though trivial, they have important consequences, for instance from the first it follows $||A||_\infty = ||A^\dagger||_\infty$. These properties are characteristic of a *conjugation* operation: the adjoint is for operators what the conjugation is for complex numbers. Correspondingly, the $L(\mathcal{H})$ algebra becomes a C^*-algebra.

Given $A \in L(\mathcal{H})$ we can always write

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

or, defining the *real* (ReA) and *imaginary* (ImA) parts of A (for reasons which will become soon clear) as

$$\mathrm{Re}A = \frac{A + A^\dagger}{2} \quad \mathrm{Im}A = \frac{A - A^\dagger}{2i}$$

we can write

$$A = \mathrm{Re}A + i\mathrm{Im}A$$

This expression is very suggestive of the complex number representation with the help of real and imaginary parts, since it is easily verified that $(\mathrm{Re}A)^\dagger = \mathrm{Re}A$ and $(\mathrm{Im}A)^\dagger = \mathrm{Im}A$, *i.e.*

$$A^\dagger = \mathrm{Re}A - i\mathrm{Im}A$$

The similarity can be made stronger by considering the *expectation values* of an operator, defined as the values taken by $(x|A|x)$ when x lies on the unit sphere $||x|| = 1$, or equivalently

$$< A >_x := (x|A|x)_{|||x||=1} \equiv \frac{(x|Ax)}{(x|x)}$$

Indeed, it readily follows

$$< \mathrm{Re}A >_x = \frac{1}{2}(x|(A + A^\dagger)x) = \frac{1}{2}(x|Ax) + \frac{1}{2}(x|Ax)^* \equiv \mathrm{Re}(x|A|x)$$

$$< \mathrm{Im}A >_x = \frac{1}{2i}(x|(A - A^\dagger)x) = \frac{1}{2i}(x|Ax) - \frac{1}{2i}(x|Ax)^* \equiv \mathrm{Im}(x|A|x)$$

5.2 Self-adjoint operators

According to the above results, operators such that $A = A^\dagger$ have a special role in the theory, as they play the same role that real numbers play in the complex plane. These operators are called *self-adjoint* operators and thus satisfy

$$< A >_x = \mathrm{Re}(x|A|x) \in \mathbb{R}$$

since $A \equiv \mathrm{Re}A$. The set of these operators form a *real* linear space in $L(\mathcal{H})$, since $\alpha A + \beta B$ is self-adjoint if A, B are self-adjoint operators and $\alpha, \beta \in \mathbb{R}$. However, this *cannot* be an algebra, since for $A = A^\dagger$ and $B = B^\dagger$ it follows at most

$$(AB)^\dagger = BA$$

which is generally different from AB; the only algebra(s) that can be built are those of *commuting* self-adjoint operators or, in the language of quantum mechanics, of *compatible observables*.

The analogy between self-adjoint operators and real number can be pushed forward by defining inequality expressions such as

$$A \leq B$$

if

$$(x|Ax) \leq (x|Bx)$$

holds for any $x \in \mathcal{H}$. In particular, if $A > 0$ ($A \geq 0$) the operator is said to be *positive (semipositive) defined*[13]. Examples of positive operators are $A^\dagger A$ and $AA^\dagger$, as can be easily checked by a direct calculation. The converse is also true, namely that if A is positive defined there exists an operator S such that $A = S^\dagger S$, but we need to consider the spectral theory of self-adjoint operators to prove this result; as we shall see, this representation is not unique and one can even choose $S = S^\dagger$.

[13]Obviously, for positive operators the condition $(x|Ax) > 0$ is meant to hold for any $x \neq 0$.

With this result we can prove the "generalized Schwartz inequality" for any semipositive defined operator[14], *i.e.*

$$|(x|Ay)|^2 \leq (x|A|x)(y|A|y)$$

Proof. The proof is analogous to that of the Schwartz inequality. Let us consider the special case $(x|Ax) = (y|Ay) = 1$; if either $(x|Ax)$ or $(y|Ay)$ are null, $Ax = 0$ $(Ay = 0)$ follows with the help of the representation $A = S^\dagger S$, and the above inequality is satisfied. For any $\lambda \in \mathbb{K}$ it holds

$$\begin{aligned} 0 &\leq (x - \lambda Ay|x - \lambda Ay) = (x|Ax) + |\lambda|^2(y|Ay) - \lambda^*(y|Ax) - \lambda(x|Ay) \\ &= 1 + |\lambda|^2 - \lambda^*(y|Ax) - \lambda(x|Ay) \end{aligned}$$

Choosing $\lambda = (y|Ax)$ we get the desired result. In general, x and y do not satisfy $(x|Ax) = (y|Ay) = 1$; however, if $(x|Ax)(y|Ay) > 0$, the vectors $x' = x/\sqrt{(x|Ax)}$ and $y' = y/\sqrt{(y|Ay)}$ are such that $(x'|Ax') = (y'|Ay') = 1$ and we get $|(x'|Ay')|^2 \leq 1$ *i.e.* $|(x|Ay)|^2 \leq (x|Ax)(y|Ay)$.

5.3 Projectors

We recall the projection theorem, which states that for any given *subspace* $V \in \mathcal{H}$ we can write $\mathcal{H} = V \oplus V^\perp$, that is, any vector $x \in \mathcal{H}$ is uniquely decomposed into its components in V and $V^\perp$, $x = x_V + x_{V^\perp}$, where x_V, $x_{V^\perp}$ are unique. This theorem allows us to define the following operator

$$P_V : \mathcal{H} \to \mathcal{H}$$

$$x \rightsquigarrow x_V$$

which *projects* any given vector onto the subspace V. The *projector* P_V is a linear, bound map, since

$$||P_V x|| = ||x_V|| \leq ||x||$$

(from which it also follows $||P_V||_\infty \leq 1$). It is a self-ajoint and *idempotent* operator,

$$P_V = P_V^\dagger \;\; P_V = P_V^2$$

as can be easily checked by direct computation that

$$(x|P_V y) = (x|y_V) = (x_V|y_V) = (x_V|y) \equiv (P_V x|y)$$

$$P_V x_V \equiv x_V \Rightarrow P_V P_V x = P_V x$$

[14]The unsatisfied reader may add the condition $(x|Ax)(y|Ay) \neq 0$ and later on consider the case where either $(x|Ax)$ or $(y|Ay)$ is null. The case of a *positive* operator does not cause, of course, any problem.

hold $\forall x, y \in \mathcal{H}$. Furthermore, for any $x \in V$ $P_V x = x$ holds and thus $||x|| = ||P_V x|| \leq ||P_V||_\infty ||x||$, *i.e.* necessarily $||P_V||_\infty = 1$.

In addition, it holds $\mathsf{Ker} P_V = V^\perp$ and $\mathsf{Im} P_V = V$. For the first notice that if $x \in V^\perp$, it follows $x \equiv x_{V^\perp}$, *i.e.* $x_V = 0$ and thus $x \in \mathsf{Ker} P_V$, and, on the other hand, if $x \in \mathsf{Ker} P_V$, by definition we have $P_V x = x_V = 0$ and thus $x \equiv x_{V^\perp} \in V^\perp$. Analogously, for $\mathsf{Im} P_V = V$, if $x \in \mathsf{Im} P_V$ $x \in V$ by constuction and, on the other hand, $x \in V$ implies $x = P_V x \in \mathsf{Im} P_V$. Notice that in this case both $\mathsf{Ker} P_V$ and $\mathsf{Im} P_V$ are *closed*: the first is a general result already proved for arbitrary bound operators, whereas the second can be checked explicitly by noticing that if $\{y_n\} \subset \mathsf{Im} P_V$ is convergent to y, after applying P_V to the sequence we obtain $P_V \lim_n y_n = P_V y \equiv \lim_n P_V y_n = \lim_n y_n = y \in \mathsf{Im} P_V$, where in the last equations we have used continuity and idempotency.

In general, we can state the following imporant result: P is a projector if and only if $P = P^\dagger = P^2$, and in that case $\mathsf{Im} P$ is the subspace onto which it projects.

Proof. We have already shown that if P is a projector on V it satisfies $P = P^\dagger = P^2$ and $\mathsf{Im} P = V$. Conversely, consider the following decomposition which holds for any vector

$$x = Px + (1 - P)x$$

Here $x' = Px$ and $x" = (1 - P)x$ are orthogonal to each other ($(Px|(1 - P)x) = (x|P(1 - P)x) = (x|(P - P^2)x) = 0$) and are legitimate condidates for the projections we are looking for. P is bound, since from the above decomposition

$$||x||^2 = ||Px||^2 + ||(1 - P)x||^2 \geq ||Px||^2$$

and $\mathsf{Im} P$ is a subspace, since (as before) if $\{y_n\} \subset \mathsf{Im} P$ converges to y, when applying P to the sequence and using $P^2 = P$, we get

$$Py = P \lim_n y_n = \lim_n P y_n = \lim_n y_n = y$$

i.e. $y \in \mathsf{Im} P$. Thus, applying the projector theorem with $V = \mathsf{Im} P$, namely $x = x_V + x_{V^\perp}$, we conclude -from the above decomposition- that $x_V \equiv Px$ and $x_{V^\perp} \equiv (1 - P)x$.

Note also that a projector P is positive defined ($P = P^\dagger P$) and for x such that $||x|| = 1$ we have

$$(x|Px) = (x|P^2 x) = (Px|Px) = ||Px||^2 \leq ||x|| = 1$$

which suggests to write

$$0 \leq P \leq 1$$

with reference to the expectation values $< P >= (x|Px)/(x|x)$ for $x \neq 0$.

The following properties characterize completely the algebra of projectors. Let P_A and P_B be two projectors, then

- $P_A + P_B$ is a projector if and only if $P_A P_B = P_B P_A = 0$

Proof. $P_A + P_B$ is self-adjoint and $(P_A + P_B)^2 = P_A + P_B + P_A P_B + P_B P_A$. For this to be a projector, then, we need

$$P_A P_B + P_B P_A = 0$$

Multiplying by P_A on the left, we obtain $P_A P_B + P_A P_B P_A = 0$ and, using $P_A P_B = -P_B P_A$, we arrive at $P_A P_B - P_B P_A = 0$, *i.e.* $P_A P_B = P_B P_A = 0$. On the other hand, it is obvious that if this condition is satisfied $P_A + P_B$ is idempotent. Notice that the condition $P_A P_B = 0$ means that P_A and P_B project onto orthogonal subspaces, since

$$(P_A x | P_B y) = (x | P_A P_B y) = 0$$

and, in that case, $P_A + P_B$ projects onto the direct sum of these subspaces.

- $P_A P_B$ is a projector if and only if $[P_A, P_B] = 0$

Proof. Sufficiency is obvious. For the necessity notice that self-adjointness requires $P_A P_B = (P_A P_B)^\dagger = P_B P_A$. Notice that if $V_A = \mathsf{Im} P_A$ and $V_B = \mathsf{Im} P_B$ are the subspaces onto which P_A and P_B project, then an operator $P_A P_B$ satisfying the condition above projects onto $V_A \bigcap V_B$. Indeed,

$$\mathsf{Im} P_A \supset \mathsf{Im} P_A P_B = \mathsf{Im} P_B P_A \subset \mathsf{Im} P_B$$

and, on the other hand, if $x \in V_A \bigcap V_B$ $x = P_A x = P_B x$ and thus, upon multiplying by P_A, it follows $x = P_A x = P_A P_B x \in \mathsf{Im} P_A P_B$.

- $P_A - P_B$ is a projector if and only if $P_A P_B = P_B P_A = P_B$

Proof. $P_A - P_B$ is self-adjoint and $(P_A - P_B)^2 = P_A + P_B - P_A P_B - P_B P_A$. Sufficiency is therefore obvious. For the necessity notice that from

$$P_A + P_B - P_A P_B - P_B P_A = P_A - P_B$$

it follows

$$2P_B = P_A P_B + P_B P_A$$

Left multiplication by P_B gives

$$2P_B = P_B P_A P_B + P_B P_A$$

whereas right multiplication gives

$$2P_B = P_A P_B + P_B P_A P_B$$

Hence, $P_A P_B = P_B P_A$ and, using this result in the first equation, $P_B = P_A P_B$. Notice that, if the above condition is satisfied, the space onto which $P_A - P_B$ projects is $V_A \bigcap V_B^\perp$ and

$$V_B \subset V_A$$

where, as before, $V_A = \mathsf{Im} P_A$ and $V_B = \mathsf{Im} P_B$. Indeed, let $x \in V_B$, from $x = P_B x$ it follows $P_A x = P_A P_B x = P_B x = x \in V_A$. On the other hand, to prove $\mathsf{Im}(P_A - P_B) \subset V_A \bigcap V_B^\perp$, let $x \in \mathsf{Im}(P_A - P_B)$, *i.e.* $x = (P_A - P_B)x$. It follows

$$P_B x = (P_B P_A - P_B)x = 0$$

and

$$(1 - P_A)x = (1 - P_A)(P_A - P_B)x = (P_A - P_B - P_A^2 + P_A P_B)x = 0$$

or, equivalently, $x \in V_B^\perp$ *and* $x \in V_A$. Conversely, to prove $\mathsf{Im}(P_A - P_B) \supset V_A \bigcap V_B^\perp$, let $x \in V_A \bigcap V_B^\perp$, that is $(1 - P_B)x = x$ (or, equivalently, $P_B x = 0$) and $x = P_A x$. It follows

$$(1 - (P_A - P_B))x = (1 + P_B)x - P_A x = x - x \equiv 0$$

which implies

$$x = (P_A - P_B)x + (1 - (P_A - P_B))x \equiv (P_A - P_B)x \in \mathsf{Im}(P_A - P_B)$$

Finally, the following property is interesting in applications. P is a projector if and only if there exists an orthonormal set $\{u_n\} \subset \mathcal{H}$ such that

$$Px = \sum_n u_n(u_n|x) \quad \forall x \in \mathcal{H}$$

Indeed, if such a set exists it is easy to show that the above relation defines a map P satisfying $P = P^\dagger = P^2$; on the other hand, if P is a projector $\mathsf{Im}P$ is an Hilbert space and it is enough to introduce an orthonormal system in such space and verify that the above relation holds for any $x \in \mathcal{H}$.

5.4 Unitary operators

A *unitary* operator is an operator $U : \mathcal{H} \to \mathcal{H}$ with the following properties:

- $(Ux|Ux') = (x|x')$ for any $x, x' \in \mathcal{H}$
- U is suriettive, *i.e.* $\mathsf{Im}U = \mathcal{H}$

Notice that the first condition is *equivalent* to the *isometric* property

$$||Ux|| = ||x|| \;\; \forall x \in \mathcal{H}$$

which follows from the above condition by setting $x' = x$. Indeed, if U is isometric,

$$||U(x + \lambda x')||^2 = ||(x + \lambda x')||^2 \;\; \forall x, x' \in \mathcal{H}, \forall \lambda \in \mathbb{K}$$

and upon expanding the square and evaluating for $\lambda = 1, i$ we get the desired result. It then easily follows that if U is unitary (isometric is enough) U is invertible, and the inverse is a map $U^{-1} : \mathcal{H} \supset \mathsf{Im}U \rightarrow \mathcal{H}$. Notice further that $||Ux|| \leq ||x||$ implies $||U||_\infty \leq 1$ and $||x|| = ||Ux|| \leq ||U||_\infty ||x||$ imples $||U||_\infty \geq 1$, *i.e.* $||U||_\infty = 1$.

Equivalently, we can say that U is unitary if and only if

$$U^\dagger U = 1 = UU^\dagger$$

The first is the isometric property of above

$$(x|x') = (Ux|Ux') = (x|U^\dagger Ux) \;\; \forall x, x' \in \mathcal{H}$$

and shows that $U^{-1} \equiv U^\dagger$, since for any $y \in \mathsf{Im}U$ there exists $x = U^{-1}y$ such that $y = Ux$ and $U^\dagger y = U^\dagger Ux = x = U^{-1}y$ holds. The second equality above, on the other hand, is the suriettivity of U. Indeed, if U is suriettive $\forall y \in \mathcal{H}$ there exists $x \in \mathcal{H}$ such that $y = Ux$ but we have seen above that $x = U^{-1}y \equiv U^\dagger y$, *i.e.* $UU^\dagger y = y$ holds for any $y \in \mathcal{H}$; on the other hand, if $U(U^\dagger x) = x$ holds it follows $x \in \mathsf{Im}U$.

There are no interesting "algebraic" properties of unitary operators *but* the following fundamental one: the product of two unitary operators U, V is a unitary operator, as can easily verified by applying the conditions above. This means that the unitary operators form a group in $L(\mathcal{H})$, usually denoted $U(\mathcal{H})$, which is the group of *automorphisms* of $\mathcal{H}$.

Notice that in finite-dimensional vector spaces the above equalities $U^\dagger U = UU^\dagger = 1$ are equivalent to each other since the isometric condition implies, in particular, that $\mathsf{Im}U$ and $\mathcal{H}$ have the same dimensions, *i.e.* that U is also suriettive. This is no longer true in *infinite* dimensional vector spaces and it is not hard to define an isometric map ($U^\dagger U = 1$) which is not suriettive. For instance, let $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ be a complete orthonormal set; a map can be defined by fixing its values on the elements of this set and we define

$$Uu_n = u_{n+1} \;\; n \geq 0$$

Clearly, this map is isometric but not suriettive since $u_0 \notin \mathsf{Im}U$. In this context, it is also worth noticing that U is unitary if and only if it transforms complete orthonormal sets $\{u_n\}_{n\in\mathbb{N}}$ in *complete* orthonormal sets $\{Uu_n\}_{n\in\mathbb{N}}$; the proof is simple and will not be given here.

Finally, of interest are also the *antiunitary* operators. Analogously to the unitary operators above, U is said to be anitunitary if it is *antilinear* and

- $(Ux|Ux') = (x|x')$ for any $x, x' \in \mathcal{H}$
- U is suriettive, *i.e.* $\mathsf{Im}U = \mathcal{H}$

Hence, the true difference is in the antilinearity property, namely

$$U(x+y) = Ux + Uy$$

$$U(\lambda x) = \lambda^* Ux$$

for any $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{K}$. Equivalently, it is not hard to show that U is antiunitary if and only if $U^\dagger U = UU^\dagger = 1$. Here, the *adjoint of an antilinear operator* satisfies

$$(U^\dagger x|x') = (x|Ux')^*$$

Notice that the conjugation is needed on the r.h.s. in order to leave the above expression linear in *both* x and x'.

5.5 Remarks on common operators

Up to now we have considered operators $A : \mathcal{E} \to \mathcal{F}$ defined in the *whole* space $\mathcal{E}$ which are continuos (bound), and showed that the corresponding set is a normed topological linear space; the set even became an albegra when $\mathcal{F} = \mathcal{E}$ $(L(\mathcal{E}))$. Apart from continuity, the most important situation which is often found in applications is that in which A is *not* defined in the whole space, rather on some *linear* subset D_A of $\mathcal{E}$. This requires some extension (and weakening) of the properties mentioned so far.

According to above, a linear operator A is properly defined if a *domain* $D_A \subset \mathcal{E}$ is given, it is a linear space, and for each $x \in D_A$ there exists a linear map $x \rightsquigarrow Ax \in \mathcal{F}$. It is clear that traditional pointwise operations between operators have to be re-defined, e.g.

$$(A+B)x := Ax + Bx \quad \forall x \in D_A \bigcap D_B$$

and this *prevents* the set of operators of this kind, call it $\mathcal{O}(\mathcal{E} \to \mathcal{F})$, to be a linear space. Indeed, the zero operator O is defined with $D_O = \mathcal{E}$ but for an arbitrary operator A no opposite operator $-A$ exists such that $A - A = 0$, unless $D_A = \mathcal{E}$. In addition one defines operator extensions and restrictions: B is said to be an *extension* of A (in symbols $B \supset A$) if $D_B \supset D_A$ and $Bx = Ax$ for any $x \in D_A$; B is said to be a *restriction* of A to the linear space $D \subset \mathcal{E}$ (in symbols $B = A_{|D}$) if $D \subset D_A$ and, for any $x \in D$, $Bx = Ax$.

It turns out that the interesting properties of the operators seen so far are related to the combination of a *closed* domain ($D_A = \mathcal{E}$ so far) and continuity. Generalizations require to consider, for any operator A, the *operator graph* $\mathcal{G}_A$ which is a subset of $\mathcal{E} \oplus \mathcal{F}$ defined in this way

$$\mathcal{G}_A = \{(x, y) \in \mathcal{E} \oplus \mathcal{F} | x \in D_A, y = Ax\}$$

This set has the following two important properties:

1. $\mathcal{G}_A$ is linear

2. $(0, y) \in \mathcal{G}_A \Rightarrow y = 0$

They can be easily verified: for the first notice that if (x, y) and (x', y') are elements of $\mathcal{G}_A$, $y = Ax$, $y' = Ax'$ and their sum $(x, y) + (x', y') \equiv (x + x', A(x + x'))$; for the second, just notice that for $x = 0$, $y = Ax = 0$.

Conversely, given a set $\mathcal{G} \subset \mathcal{E} \oplus \mathcal{F}$ with the above two properties one can *define* an operator. This can be realized with the help of the "natural" projections π_E and π_F

$$\pi_E : \mathcal{E} \oplus \mathcal{F} \to \mathcal{E}$$

$$(x, y) \rightsquigarrow \pi_E(x, y) := x$$

and similarly for π_F. Indeed, an operator A can be defined by noticing that condition 2 means that for any $\boldsymbol{x} = (x, y) \in \mathcal{G}$

$$(x, y_1) = (x, y_2) \Rightarrow y_1 = y_2$$

i.e. $\pi_F(\boldsymbol{x}) = \pi_F(\boldsymbol{x}')$ if $\pi_E(\boldsymbol{x}) = \pi_E(\boldsymbol{x}')$. Thus

$$D_A = \{x \in \mathcal{E} | x = \pi_E(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{G}\}$$

$$Ax = \pi_F(\boldsymbol{x}), \quad x \in D_A$$

is a properly defined *map* since for any $x \in D_A$ ($x = \pi_E(\boldsymbol{x})$) there exists a unique $\pi_F(\boldsymbol{x})$. It is also easily verified that the map is linear.

Then, important classes of operators are provided by the *closed* operators and the *closable* operators. A is said to be closed if its graph is closed (and one writes $A = \bar{A}$) or, in other words,

$$\text{if } \mathcal{G}_A \supset \{(x_n, y_n)\}_{n\in\mathbb{N}} \to (x, y), \text{ i.e.} \begin{cases} x_n \to x \\ \\ Ax_n \to y \end{cases} \Rightarrow (x, y) \in \mathcal{G}_A, \text{ i.e.} \begin{cases} x \in D_A \\ \\ y = Ax \end{cases}$$

A is closable if the closure of $\mathcal{G}_A$ ($\bar{\mathcal{G}}_A$) is an operator graph, *i.e.* it satisfies conditions 1 and 2 of above. Notice that condition 1 is trivial but condition 2 reads as

$$\text{if } \mathcal{G}_A \supset \{(x_n, y_n)\}_{n\in\mathbb{N}} \to (0, y), \text{ i.e.} \begin{cases} x_n \to 0 \\ \\ Ax_n \to y \end{cases} \Rightarrow y = 0$$

If A is closable the operator defined by the graph $\bar{\mathcal{G}}_A$ is called the *closure* of A, and denoted $\bar{A}$. It is the "smallest" closed extension of A, $\bar{A} \supset A$, and of course $\mathcal{G}_{\bar{A}} = \bar{\mathcal{G}}_A$.

One can then define continuity with reference to D_A in the usual way, and introduce a "norm"

$$||A||_\infty = \mathsf{Inf}\{M > 0 | ||Ax|| \leq M||x||, \, x \in D_A\}$$

which however cannot be a true norm since $\mathcal{O}$ is not a linear space. This property is closely related to the topological properties of the domain *and* of the graph, e.g. if D_A is closed ($D_A = \bar{D}_A$) and $\mathcal{G}_A$ is closed ($A = \bar{A}$) then A is continuous.

For operators *in* a Hilbert space $A : \mathcal{H} \to \mathcal{H}$ the graph $\mathcal{G}_A$ is a subset of $\mathcal{H} \oplus \mathcal{H}$. For bound operators the adjoint $A^\dagger$ is well defined and the graph of $A^\dagger$, $\mathcal{G}_{A^\dagger}$, is related to that of A by the relation

$$(x'|y) = (x'|Ax) = (A^\dagger x'|x) = (y'|x)$$

where $(x', y') \in \mathcal{G}_{A^\dagger}$ and $(x, y) \in \mathcal{G}_A$. In other words, by introducing the (unitary) operator

$$V : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}$$
$$(x, y) \rightsquigarrow V(x, y) := (-y, x)$$

the above condition can be rewritten as

$$0 = (x'|y) - (y'|x) = \big((x', y')|V(x, y)\big)_\oplus$$

where $(.|.)_\oplus$ is the scalar product naturally defined in $\mathcal{H} \oplus \mathcal{H}$. Thus, $\mathcal{G}_{A^\dagger} \equiv (V(\mathcal{G}_A))^\perp$, where $V(\mathcal{G}_A)$ is the "image of $\mathcal{G}_A$ through the map V", *i.e.* $V(\mathcal{G}_A) = \mathsf{Im}_{\boldsymbol{x} \in \mathcal{G}_A} V$.

Then, in general, one says that A is *adjointable* if $(V(\mathcal{G}_A))^\perp$ is an operator graph (*i.e.*, again, it satisfies properties 1 and 2 above) and its adjoint is the operator defined by such graph. Notice that $A^\dagger$ defined in this way, when possible, is a closed operator ($A^\dagger = \bar{A}^\dagger$) since its graph is an orthogonal complement of some subset, *i.e.* a closed space. In addition,

$$\begin{gathered}(0, y) \in (V(\mathcal{G}_A))^\perp \\ \Updownarrow \\ ((0, y)|V(x', y'))_\oplus = ((0, y)|(-y', x'))_\oplus = (y|x') = 0 \; \forall x' \in D_A, y' = Ax \\ \Updownarrow \\ y \in D_A^\perp\end{gathered}$$

and thus, for the validity of condition 2, it follows that $D_A^\perp = \{0\}$ or, equivalently, $\bar{D}_A = D_A^{\perp\perp} = \mathcal{H}$. In other words, a necessary condition for defining the adjoint is that the operator domain has to be dense in $\mathcal{H}$, *i.e.* $\bar{D}_A = \mathcal{H}$. Notice that if this property holds, A can be closed to an operator $\bar{A}$ which is defined on the whole space $\mathcal{H}$ and $A^\dagger := \bar{A}^\dagger$ can be defined. Thus, A **is adjointable if and only if its domain is dense in** $\mathcal{H}$. The domain of the adjoint, on the other hand, reads as

$$D_{A^\dagger} = \{y \in \mathcal{H} | \exists y' \in \mathcal{H}, \, (y|Ax) = (y'|x) \, \forall x \in D_A\}$$

One can then show that the main properties of the adjoint hold provided some conditions are given. For instance, $A^\dagger$ is closed but not necessarly adjointable. For A to be adjointable

$$(V(\mathcal{G}_{A^\dagger}))^\perp = V(V(\mathcal{G}_A)^\perp)^\perp = V^2(\mathcal{G}_A)^{\perp\perp} = \bar{\mathcal{G}}_A$$

(we have used $V^2 = -1$) has to be the graph of an operator, *i.e.* A has to be closable, and in that case $A^{\dagger\dagger} = \bar{A}$. Notice that A closed *and* D_A closed implies A *continuos*, so if A is closed and adjointable D_A cannot be $\mathcal{H}$ unless A is continuous.

The interest, therefore, is in operators densely defined (adjointable) and closed (or closable), and in particular, in those operators for which $A = A^\dagger$ which are closed (self-adjoint operators) or at most $A \subset A^{\dagger\dagger} = A^\dagger$ which can be closed to give a self-adjoint operator (essentially self-adjoint operators).

6 Spectral theory of operators

In the following we focus on *complex* Hilbert spaces $\mathcal{H}$ and on many interesting properties of the operators in $\mathcal{H}$ that follow from the simple question[15]: does the operator $(\alpha - A)^{-1}$ exist for a given $\alpha \in \mathbb{C}$?

If the answer is *positive,* $G_A(\alpha) = (\alpha - A)^{-1}$ is called the resolvent of A, since it solves the problem $(\alpha - A)x = y$ for the unknown x, provided $y \in \mathsf{Im}(\alpha - A)$.

If the answer is *negative*, we have found a *special* value of A (and corresponding *special* vectors $x \in \mathsf{ker}(\alpha - A)$) which characterizes A.

6.1 Eigenvalues and eigenvectors

Let $A \in L(\mathcal{H})$ be a bound operator, $\alpha \in \mathbb{C}$ and consider the equation

$$Ax = \alpha x$$

If this equation is satisfied for $x \neq 0$ then x is called an *eigenvector* of A and α is the corresponding *eigenvalue.* By definition, eigenvectors are the "natural" vectors of an operator, *i.e.* those for which the action of A is just a contraction/expansion. For a given eigenvalue α, the set of vectors satisfying the above equation, namely, $V_\alpha^A = \{x \in \mathcal{H} | Ax = \alpha x\}$ (or simply V_α if no possibility of confusion arises) is linear and closed, *i.e.* it is a subspace, called the $\alpha-$eigenspace (in fact, $V_\alpha \equiv \mathsf{ker}(\alpha - A)$). Here, closedness is a consequence of continuity, since for any converging sequence in V_α, *i.e.* $\{x_n\}_{n\in\mathbb{N}} \subset V_\alpha, x_n \to x$, it follows

$$Ax = A \lim_n x_n = \lim_n Ax_n = \alpha \lim_n x_n = \alpha x, \text{ i.e. } x \in V_\alpha$$

The set of the eigenvectors of an operator, $\sigma_A = \{\alpha\} \subset \mathbb{C}$ is called the *spectrum* of A. Of interest are also the expectation values of A, previously introduced. For any $x \in \mathcal{H}$, $E_A(x)$ is the (non-linear) functional defined by

$$x \in \mathcal{H}, x \neq 0 \ x \rightsquigarrow E_A(x) := \frac{(x|Ax)}{(x|x)}$$

[15]Traditionally, one writes $\alpha - A$ to mean the operator $\alpha 1 - A$.

We call the set of the possible expectation values, $\mathsf{Im}E_A \subset \mathbb{C}$, the *range* of A; clearly, $\sigma_A \subset \mathsf{Im}E_A$, since if x is an eigenvector with eigenvalue α, $E_A(x) \equiv \alpha$.

Notice that for $A \in L(\mathcal{H})$ and $x \neq 0$

$$|E_A(x)| = \left|\frac{(x|Ax)}{(x|x)}\right| \leq ||A||_\infty$$

that is, the range (and the spectrum) of a bound operator is bound[16].

Simple properties arise for *normal* operators $N \in L(\mathcal{H})$. An operator is said to be normal if $[N, N^\dagger] = 0$; important examples are of course the self-adjoint operators and the unitary operators. It is not hard to show that N is normal if and only if

$$||Nx|| = ||N^\dagger x|| \ \forall x \in \mathcal{H}$$

This is a consequence of the following simple result: if A and B are self-adjoint operators, then $A = B$ if and only if $(x|Ax) = (x|Bx)$ for any $x \in \mathcal{H}$ [17]. Thus, sufficiency can be proved by noticing $||N^\dagger x||^2 = (N^\dagger x|N^\dagger x) = (x|NN^\dagger x)$ and analogously $||Nx||^2 = (x|N^\dagger Nx)$, where both $N^\dagger N$ and $NN^\dagger$ are self-adjoint; necessity is trivial.

It thus follows that if x is eigenvector of N with eigenvalue ν, x is also eigenvector of $N^\dagger$ with eigenvalue ν^*: indeed, the operator $A = \nu - N$ is normal and $||(\nu - N)x|| = 0 = ||(v^* - N^\dagger)x||$. It further follows that if $A = A^\dagger$ the spectrum is real, $\sigma_A \subset \mathbb{R}$, a result which could be also anticipated since we have already seen that for a self-adjoint operator $\mathsf{Im}E_A \subset \mathbb{R}$.

Same cases are of particular interest. For a projector $P = P^2 = P^\dagger$ which projects onto $\mathsf{Im}P$, if x is eigenvector with eigenvalue π we have

$$P(1 - P)x = \pi(1 - \pi)x = 0$$

i.e. $\pi = 0, 1$ are the only possible eigenvalues. Of course, $\pi = 1$ if and only if $x \in \mathsf{Im}P$ and $\pi = 0$ if and only if $x \in (\mathsf{Im}P)^\dagger = \mathsf{Im}(1 - P)$. Similarly, for a unitary operator U, if x is eigenvector with eigenvalue η

$$(U^\dagger U - 1)x = (\eta^*\eta - 1)x = 0$$

i.e. the only possible eigenvalues are those of unit modulus, $|\eta| = 1$.

Eigenvectors of normal operators have interesting properties, too. If x_ν and $x_{\nu'}$ are eigenvectors corresponding to different eigenvalues, $\nu \neq \nu'$, then $(x_\nu|x_{\nu'}) = 0$. The proof is simple:

$$\nu'(x_\nu|x_{\nu'}) = (x_\nu|Nx_{\nu'}) = (N^\dagger x_\nu|x_{\nu'}) = (\nu^* x_\nu|x_{\nu'}) = \nu(x_\nu|x_{\nu'})$$

[16]For any normed space $\mathcal{M}$ ($\mathbb{C}$ in this case), a bound set $B \subset \mathcal{M}$ is a set that can be included in a sphere, *i.e.* $B \subset S_R$ where $S_R = \{x \in \mathcal{M} | \, ||x|| < R\}$ for some $R > 0$.

[17]The proof is simple, just use the condition $(x + \lambda y|(A - B)(x + \lambda y) = 0$ for $\lambda = 1, i$. Notice though that self-adjointness is essential for this property to hold: indeed, we are just asking equality between "diagonal" elements.

i.e. if $\nu \neq \nu'$ necessarily $(x_\nu|x_{\nu'}) = 0$. In other words, different eigenspaces of a normal operator are orthogonal to each other, $V_\nu \perp V_{\nu'}$ for $\nu \neq \nu'$.

There remains to characterize the spectrum σ_A, which is a quite subtle issue. In finite-dimensional Hilbert space ($\mathsf{dim}\mathcal{H} < \infty$) we know from algebra that the distinct eigenvalues are at most $\mathsf{dim}\mathcal{H}$ in number since they are the zeros of a polynomial of degree $\mathsf{dim}\mathcal{H}$, the so-called *characteristic polynomial*, $p_A(\alpha)$. The latter is defined as

$$p_A(\alpha) = \det(\alpha - \mathbb{A})$$

where $\mathbb{A}$ is a matrix representation of the operator, *i.e.*

$$\mathbb{A}_{nm} = (u_n|Au_m)$$

where $\{u_n\}$ is an orthonormal *basis* of $\mathcal{H}$. Note that the polynomial is basis-dependent but its zeros are not.

This finite-dimensional case is a special case of a more general result which holds for any *compact* operator[18], a special class of bound operators, $L_c(\mathcal{H}) \subset L(\mathcal{H})$. The important result is

Theorem If $A \in L_c(\mathcal{H})$ is a compact operator σ_A is at most numerable ($\text{card}\sigma_A \leq \aleph_0$). If $\text{card}\sigma_A = \aleph_0$ there exists a sequence of eigenvalues $\{\alpha_n\}$ such that $\alpha_n \to 0$, though $\alpha = 0$ is not necessarily an eigenvalue.

The finite-dimensional case discussed above follows from the property that *any* operator in a finite-dimensional vector space is compact.

In general, however, this result cannot be extended to arbitrary vector spaces and operators in those spaces, and we have to admit the possibility that in infinite-dimensional Hilbert spaces operators can have a "continuous" (non-numerable) part of the spectrum. These eigenvalues are *not* true eigenvalues – *i.e.,* there are no "proper" eigenvectors attached to them– and $G_A(\lambda)$ is actually *defined* for λ in the continuos spectrum, though not "nicely behaving". It is however much more useful to consider them as *improper* eigenvalues because in that way one can write a sort of *spectral representation* which exactly parallels the one rigorously holding for compact operators, provided self-adjointness is satisfied. In the following we use, when needed, σ_A^d and σ_A^c for the "discrete" and the "continuos" parts of the spectrum, respectively.

[18] An operator is said to be *compact* if it transforms *bound* subsets of $\mathcal{H}$ in *relatively compact* subsets. A bound set is a set which can be included in a sphere; a relatively compact set is a set K whose closure $\bar{K}$ is *compact*. A set is said to be compact if for *any* open covering $\{\Omega_\alpha\}_{\alpha\in I}$ (Ω_α open set, $\{\Omega_\alpha\}_{\alpha\in I}$ is a covering if $\bigcup_\alpha \Omega_\alpha \supset K$) there exists a *sub-covering* with a *finite* number of elements, $\{\Omega_\alpha\}_{\alpha=1}^n \supset K$. Note that a relatively compact set is always bound and thus a compact operator is necessarily a bound operator.

6.2 Spectral representation of observables

Let A be a compact, self-adjoint operator, σ_A its spectrum and $\{V_\alpha\}_{\alpha\in\sigma_A}$ its eigenspaces. We have seen above that if $\alpha \neq \alpha'$ $V_\alpha \perp V_{\alpha'}$ holds, *i.e.* if P_α is the projector in the α-th eigenspace ("eigenprojector") it holds

$$P_\alpha P_{\alpha'} = \delta_{\alpha\alpha'} P_\alpha$$

Let then $V = \sum^{\oplus}_{\alpha\in\sigma_A} V_\alpha$ be the space spanned by the *projector* $P = \sum_{\alpha\in\sigma_A} P_\alpha$. For any $x \in V$

$$x = \sum_{\alpha\in\sigma_A} P_\alpha x$$

and

$$Ax = \sum_{\alpha\in\sigma_A} AP_\alpha x = \sum_{\alpha\in\sigma_A} \alpha P_\alpha x$$

i.e.

$$A = \sum_{\alpha\in\sigma_A} \alpha P_\alpha$$

holds *in* V. This is the spectral representation of A in terms of its eigenprojectors. A fundamental theorem is

Theorem If $A \in L_c(\mathcal{H})$ is a compact, normal operator, $V = \sum^{\oplus}_{\alpha\in\sigma_A} V_\alpha \equiv \mathcal{H}$.

Thus, for any such operator

$$1 = \sum_{\alpha\in\sigma_A} P_\alpha$$

$$A = \sum_{\alpha\in\sigma_A} \alpha P_\alpha$$

hold in $\mathcal{H}$. The first equation is a *completeness* relation: it says that for any operator of this kind there exists a *complete* orthonormal set of eigenvectors. The second equation expresses the action of the operator in terms of its elementary components, the eigenvectors, which define the "main" directions of the operator. This is the *spectral representation* of the operator A.

For the expectation values we have, for any $x \in \mathcal{H}$ ($||x|| = 1$ in the following),

$$1 = \sum_{\alpha\in\sigma_A} ||x_\alpha||^2 = \sum_{\alpha\in\sigma_A} (x|P_\alpha x) = \sum_{\alpha\in\sigma_A} p_\alpha(x)$$

where we have introduced

$$1 \geq p_\alpha(x) := (x|P_\alpha x) \geq 0$$

and

$$E_A(x) = \sum_{\alpha\in\sigma_A} \alpha(x|P_\alpha x) \equiv \sum_{\alpha\in\sigma_A} \alpha p_\alpha(x)$$

Thus, $p_\alpha(x)$ can be considered the "fraction" of x in the $\alpha-$eigenspace and $E_A(x)$ can be regarded as the *average* value of $\{\alpha\}_{\alpha\in\sigma_A}$ as given by an $x-$dependent probability distribution $\{p_\alpha\}_{\alpha\in\sigma_A}$ ($p_\alpha \equiv p_\alpha(x) = E_{P_\alpha}(x)$).

Notice that

$$0 \le P_\alpha \le 1, \quad \sum_{\alpha\in\sigma_A} P_\alpha = 1, \quad P_\alpha P_{\alpha'} = \delta_{\alpha\alpha'} P_\alpha$$

or equivalently,

$$V_\alpha \subset \mathcal{H}, \quad \sum_{\alpha\in\sigma_A}^{\oplus} V_\alpha = \mathcal{H}, \quad V_\alpha \perp V_{\alpha'} \text{ for } \alpha \neq \alpha'$$

are suggestive of a "probability measure". Indeed, in general, a *probability measure* can be defined as follows: let Ω be a space (the *sample* space) and $\mathcal{A}$ a $\sigma-$*algebra* on that space. $\mathcal{A}$ is a set of subsets of Ω, $\mathcal{A} = \{A_n,\ A_n \subset \Omega\}$ with the following properties

- $\Omega, \phi \in \mathcal{A}$
- $A_1, A_2 \in \mathcal{A} \ \Rightarrow\ A_1 \bigcup A_2,\ A_1 \bigcap A_2,\ A_1 \backslash A_2 \in \mathcal{A}$
- $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A} \ \Rightarrow\ \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

The sets $\Omega \supset A_i \in \mathcal{A}$ are called *events* and a probability measure is a function

$$\mu : \mathcal{A} \to \mathbb{R}^+$$

$$\mathcal{A} \ni A_i \rightsquigarrow \mu(A_i) \in \mathbb{R}^+$$

such that

- $\mu(\Omega) = 1$, $\mu(\phi) = 0$; $0 \le \mu(A) \le 1$
- for any sequence $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $A_n \bigcap A_m = \phi$ for $n \neq m$, $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$

Notice that the fact the μ is real-valued is not essential, provided is "positive" valued in some sense. The results obtained above for compact operators indeed identify P_α as a projector-valued probability for a vector lying in V_α. In this case $\Omega \equiv \sigma_A$, the events are sets of eigenvalues $\omega = \{..\alpha_i, \alpha_j..\}$ with their associated spaces $V_\omega = ..V_i \oplus V_j..$ and projectors P_ω. The map defined by

$$\mu : \mathcal{A} \to L(\mathcal{H})$$

$$\mathcal{A} \ni \omega \rightsquigarrow \mu(\omega) = P_\omega$$

satisfies

- $\mu(\sigma_A) = 1$, $\mu(\phi) = 0$; $0 \le P_\omega \le 1$

- for any sequence $\{\omega_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $\omega_n \bigcap \omega_m = \phi$ for $n \neq m$, $\mu(\bigcup_{n\in\mathbb{N}} \omega_n) = \sum_{n\in\mathbb{N}} \mu(\omega_n) = \sum_{n\in\mathbb{N}} P_{\omega_n}$.

and is thus a measure.

The results obtained so far can be generalized to generic self-adjoint operators, which we now call *observables*, provided we abandon the condition that the spectrum is solely discrete. As already mentioned above, this creates serious problems if we look for true eigenvectors, as can be seen by the following heuristic argument. Let $x_\alpha \in V_\alpha$ be an eigenvector of the self-adjoint operator A; it satisfies

$$E_A(x_\alpha) = \alpha \;\; p_{\alpha'}(x_\alpha) = \delta_{\alpha\alpha'}$$

If we have to generalize the completeness relation and the spectral representation to

$$1 = \sum_{\alpha\in\sigma_A^d} P_\alpha + \int_{\sigma_A^c} P_\alpha d\alpha$$

$$A = \sum_{\alpha\in\sigma_A^d} \alpha P_\alpha + \int_{\sigma_A^c} \alpha P_\alpha d\alpha$$

we necessarily have, for $\alpha \in \sigma_A^c$,

$$E_A(x_\alpha) = \alpha = \int_{\sigma_A^c} \alpha' p_{\alpha'}(x_\alpha) d\alpha'$$

i.e.

$$p_{\alpha'}(x_\alpha) = (x_\alpha|P_{\alpha'} x_\alpha) = \delta(\alpha' - \alpha)$$

This implies that x_α *cannot be normalizable* ($||x_\alpha||^2 = \lim_{\alpha'\to\alpha} p_{\alpha'}(x_\alpha)$), or in other words, $x_\alpha \notin \mathcal{H}$ (!).

Fortunately, these problems can be overcome and it turns out that one can introduce *improper* vectors[19] x_α which do *not* belong to $\mathcal{H}$ but are such that

$$x_{\alpha,\Delta\alpha} = \frac{1}{\sqrt{\Delta\alpha}} \int_\alpha^{\alpha+\Delta\alpha} x_{\alpha'} d\alpha'$$

belongs to $\mathcal{H}$ and

$$\lim_{\Delta\alpha\to 0} (x_{\alpha,\Delta\alpha}|x_{\alpha,\Delta\alpha}) < \infty$$

[19]The possibility of "extending" the Hilbert space is related to the following observations. First, for any vector space $\mathcal{E}$ we can define a "non-orthodox" dual space $\tilde{\mathcal{E}}^*$ on the basis of a non-orthodox topology. For instance, we can set by definition $\alpha \in \tilde{\mathcal{E}}^*$ if α is the weak-limit of some sequence $\{\alpha_n\} \subset \mathcal{E}^*$, *i.e.* if there exists a sequence of ordinary functionals such that $\lim_{n\to\infty} < \alpha_n, x >= \alpha(x)$, $\forall x \in \mathcal{E}$ (clearly, if the sequence converges properly, it also weakly convergent or, in other words, $\mathcal{E}^* \subset \tilde{\mathcal{E}}^*$). Secondly, if $\mathcal{E} \subset \mathcal{H}$ properly, then it holds $\tilde{\mathcal{E}}^* \supset \mathcal{E}^* \supset \mathcal{H}^*$. Thus if the $\mathcal{E}$ space is dense in the Hilbert space, $\bar{\mathcal{E}} = \mathcal{H}$, the extended dual $\tilde{\mathcal{E}}^*$is larger yet close to $\mathcal{H}^* \cong \mathcal{H}$; vectors in this extended space correspond to the improper vectors.

The orthogonality properties can be re-stated in the form

$$(x_\alpha|x_{\alpha'}) = \delta_{\alpha\alpha'} \quad \alpha, \alpha' \in \sigma_A^d$$

$$(x_\alpha|x_{\alpha'}) = 0 \quad \alpha \in \sigma_A^d, \alpha' \in \sigma_A^c$$

$$(x_\alpha|x_{\alpha'}) = \delta(\alpha - \alpha') \quad \alpha, \alpha' \in \sigma_A^c$$

Correspondingly, the "projectors" P_α (for $\alpha \in \sigma_A^c$) become *improper* projectors

$$P_\alpha P_{\alpha'} = \delta(\alpha - \alpha')P_\alpha \quad \alpha, \alpha' \in \sigma_A^c$$

such that

$$P_{\alpha,\Delta\alpha} = \frac{1}{\sqrt{\Delta\alpha}} \int_\alpha^{\alpha+\Delta\alpha} P_{\alpha'} d\alpha'$$

is a true projector. The quantity $P_\alpha d\alpha := dP_\alpha$ is then called a *differential projector*.

Thus, one can introduce a generalized, orthonormal complete set of special vectors for each observables. We write a "ket" $|\psi\rangle$ for any proper or improper vector and $\langle\phi|\psi\rangle$for the corresponding proper or improper scalar product between two kets, $|\phi\rangle$ and $|\phi\rangle$. For the observable A with eigenvalues $\{\alpha_1, \alpha_2..\alpha_n, ..\} = \sigma_A^d$ and $\{\alpha\} = \sigma_A^c$ we write $|\alpha_n\rangle$ for a proper eigenvector and simply $|\alpha\rangle$for an improper one. On account of the possible degeneracies we write

$$P_{\alpha_n} = \sum_{k=1}^{g_n} |\alpha_n, k\rangle \langle\alpha_n, k|$$

$$P_\alpha = \sum_{k=1}^{g(\alpha)} |\alpha, k\rangle \langle\alpha, k|$$

and, eventually, re-write the completeness relation and the spectral decomposition above in terms of these vectors. Clearly, the formulation in terms of projectors is much simpler and can be further shortened by writing

$$1 = \int_{\sigma_A} P_\alpha d\mu(\alpha) \quad A = \int_{\sigma_A} \alpha P_\alpha d\mu(\alpha)$$

$$P_\alpha P_{\alpha'} = \delta_\mu(\alpha - \alpha')P_\alpha$$

where it is meant that μ turns the integral into a sum when $\alpha \in \sigma_A^d$.

Example: Position and momentum operators

Let q, p be the ordinary position and momentum operators of quantum mechanics and $|q_0\rangle$ an eigenvector of q with eigenvalue q_0. $|q_0\rangle$ cannot have a finite norm, since from the fundamental commutator

$$[q, p] = i\hbar$$

it follows

$$i\hbar \langle q_0|q_0\rangle = \langle q_0|[q,p]|q_0\rangle = q_0 \langle q_0|p|q_0\rangle - \langle q_0|p|q_0\rangle q_0$$

i.e. if $\langle q_0|q_0\rangle < \infty$ necessarily $\langle q_0|q_0\rangle = 0$ follows. Thus the only possibility left to us is that *both* $\langle q_0|q_0\rangle$ and $\langle q_0|p|q_0\rangle$ are infinite. Let us determine the spectrum of q. From the above commutator the following equations

$$[q, p^n] = i\hbar n p^{n-1}$$

and

$$[q, f(p)] = i\hbar \frac{\partial f}{\partial p}(p)$$

are readily proved. The first follows by induction; the second holds for any analytic function f

$$f(p) = \sum_{n=0}^{\infty} c_n p^n$$

upon applying the first equation term by term. Let us then consider $\xi \in \mathbb{R}$ and the following *unitary* operator

$$T_\xi(p) = e^{-\frac{i}{\hbar}\xi p}$$

Using the above equation for $f \equiv T$ we obtain

$$[q, T_\xi(p)] = i\hbar \frac{\partial T_\xi}{\partial p}(p) = \xi T_\xi(p)$$

or, equivalently,

$$qT_\xi(p) = T_\xi(p)(q+\xi)$$

$$q\left(T_\xi(p)\,|q_0\rangle\right) = (q_0+\xi)\left(T_\xi(p)\,|q_0\rangle\right)$$

This shows that $T_\xi(p)\,|q_0\rangle$ is eigenvector of q (with the same norm of $|q_0\rangle$ since T_ξ is unitary) and $q_0+\xi$ is eigenvalue, *for arbitrary* $\xi \in \mathbb{R}$. Therefore, the spectrum of q is continuous, $\sigma_q \equiv \mathbb{R}$, and the eigenvectors can be generated by $T_\xi(p)$ once an eigenvector $|q_0\rangle$ is given,

$$|q_0+\xi\rangle := T_\xi(p)\,|q_0\rangle$$

Clearly, the same holds for p with the help of the operator $T_\zeta(x) = e^{+\frac{i}{\hbar}\zeta x}$. We write

$$1 = \int_{\mathbb{R}} dq\,|q\rangle\langle q| = \int_{\mathbb{R}} dp\,|p\rangle\langle p|$$

$$q = \int_{\mathbb{R}} dqq\,|q\rangle\langle q| \qquad p = \int_{\mathbb{R}} dpp\,|p\rangle\langle p|$$

Correspondingly,

$$\psi(q) := \langle q|\psi\rangle$$

$$\psi(p) := \langle p|\psi\rangle$$

are the *representations* of $|\psi\rangle$ in the q, p eigenvector basis, respectively.

Example: (improper) momentum vectors

In the above coordinate (Schrdinger) representation (set now $x \equiv q$ as usual) the momentum operator reads as

$$\langle x|p\psi\rangle = -i\hbar\frac{\partial\psi}{\partial x}(x) = -i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle$$

Thus, the coordinate representation of the operator p follows upon setting $|\psi\rangle = |x'\rangle$, namely

$$\langle x|p|x'\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|x'\rangle = -i\hbar\frac{\partial}{\partial x}\delta(x - x')$$

The eigenvectors $|p\rangle$ in this representation follows from

$$\langle x|p|p\rangle = p\,\langle x|p\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle$$

i.e.

$$\langle x|p\rangle \propto e^{\frac{i}{\hbar}px}$$

Clearly,

$$\langle p|p\rangle = \int_{\mathbb{R}} dx\,\langle p|x\rangle\,\langle x|p\rangle = \int_{\mathbb{R}} dx|\,\langle x|p\rangle\,|^2 = \infty$$

but the vectors

$$\psi_p(x) = \frac{1}{\sqrt{\Delta p}}\int_p^{p+\Delta p} e^{\frac{i}{\hbar}px}dp$$

are proper vectors with finite norm even in the limit $\Delta p \to 0$. To see this, consider for simplicity the case $p = -\Delta p/2$,

$$\psi_{-\Delta p/2}(x) = \frac{1}{\sqrt{\Delta p}}\int_{-\Delta p/2}^{+\Delta p/2} e^{\frac{i}{\hbar}px}dp \equiv \frac{1}{\sqrt{\Delta p}}\frac{2\hbar}{x}sin\left(\frac{\Delta p}{2\hbar}x\right)$$

We have

$$\int_{\mathbb{R}} |\psi_{-\Delta p/2}(x)|^2 dx = \frac{4\hbar^2}{\Delta p}\int_{\mathbb{R}} sin^2\left(\frac{\Delta p}{2\hbar}x\right)\frac{dx}{x^2} < \infty$$

since the integrand goes as $\propto \left(\frac{\Delta p}{2\hbar}\right)^2$ when $x \to 0$ and as $\propto x^{-2}$ when $x \to \infty$. Specifically, the integral remains finite even when $\Delta p \to 0$,

$$\int_{\mathbb{R}} |\psi_{-\Delta p/2}(x)|^2 dx = 2\hbar\int_{\mathbb{R}}\frac{sin^2(\xi)}{\xi^2}d\xi$$

Thus, the above function represents a legitimate "improper" eigenfunction of p. The normalization $N(p)$ in

$$\langle x|p\rangle = N(p)e^{\frac{i}{\hbar}px}$$

can be fixed by the requirement

$$\begin{aligned}\langle p|p'\rangle &= \delta(p-p') = N^*(p)N(p)\int_{\mathbb{R}} dx e^{-\frac{i}{\hbar}px}e^{+\frac{i}{\hbar}p'x} = \\ & N^*(p)N(p)2\pi\delta\left(\frac{p-p'}{\hbar}\right) \equiv |N(p)|^2 2\pi\hbar\delta(p-p')\end{aligned}$$

i.e. we can set

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}px}$$

Example: free particle

For a particle in 3D the Hamiltonian operator reads simply as

$$H = \frac{p^2}{2m}$$

Thus

$$H|\boldsymbol{p}\rangle = \frac{p^2}{2m}|\boldsymbol{p}\rangle$$

shows that $|\boldsymbol{p}\rangle$'s are eigenvectors of H and

$$H = \int d^3p \frac{p^2}{2m}|\boldsymbol{p}\rangle\langle\boldsymbol{p}|$$

Such vectors are normalized as

$$\langle\boldsymbol{p}|\boldsymbol{p}'\rangle = \delta_3(\boldsymbol{p}-\boldsymbol{p}')$$

We can also write ($E_p = p^2/2m$ implies $dE_p = pdp/m$; furthermore, $\hat{\boldsymbol{p}}$ below is the unit vector running over the sphere which specifies the direction of $\boldsymbol{p}$)

$$\begin{aligned}H &= \int d^3p\frac{p^2}{2m}|\boldsymbol{p}\rangle\langle\boldsymbol{p}| = \int dp p^2 \int d\hat{\boldsymbol{p}}\frac{p^2}{2m}|p\hat{\boldsymbol{p}}\rangle\langle p\hat{\boldsymbol{p}}| = \\ &= \int_0^\infty dE_p mpE_p\int d\hat{\boldsymbol{p}}|p\hat{\boldsymbol{p}}\rangle\langle p\hat{\boldsymbol{p}}| \equiv \int_0^\infty dE_pE_p\int d\hat{\boldsymbol{p}}|E_p\hat{\boldsymbol{p}}\rangle\langle E_p\hat{\boldsymbol{p}}|\end{aligned}$$

where

$$|E_p\hat{\boldsymbol{p}}\rangle := \sqrt{mp}\,|p\hat{\boldsymbol{p}}\rangle$$

In this form

$$P(E_p) = \int d\hat{\boldsymbol{p}}|E_p\hat{\boldsymbol{p}}\rangle\langle E_p\hat{\boldsymbol{p}}|$$

must be the (improper) projector on the energy E_p. To see this, notice that $|E_p\hat{\boldsymbol{p}}\rangle$ is of course eigenvector of H with eigenvalue E_p and, from

$$\delta(E_p - E_{p'}) = \delta\left(\frac{p^2 - p'^2}{2m}\right) = 2m\delta((p-p')(p+p')) \equiv \frac{m}{p}\delta(p-p'),$$

it follows

$$\begin{aligned}\langle E_p\hat{\boldsymbol{p}}|E_{p'}\hat{\boldsymbol{p}}'\rangle &= m\sqrt{pp'}\,\langle p\hat{\boldsymbol{p}}|p'\hat{\boldsymbol{p}}'\rangle \equiv mp\delta_3(\boldsymbol{p}-\boldsymbol{p}') = \\ &= mp\frac{1}{p^2}\delta(p-p')\delta_2(\hat{\boldsymbol{p}}-\hat{\boldsymbol{p}}') \equiv \delta(E_p - E_{p'})\delta_2(\hat{\boldsymbol{p}}-\hat{\boldsymbol{p}}')\end{aligned}$$

Here we have used

$$d^3\boldsymbol{p}\delta_3(\boldsymbol{p}-\boldsymbol{p}') = d\hat{\boldsymbol{p}}dpp^2\delta_3(\boldsymbol{p}-\boldsymbol{p}') = d\hat{\boldsymbol{p}}dp\delta(p-p')\delta_2(\hat{\boldsymbol{p}}-\hat{\boldsymbol{p}}')$$

6.3 Functions of observables and commuting observables

Let now A be a seld-adjoint operator, σ_A its spectrum and

$$A = \int_{\sigma_A} \alpha P_\alpha d\mu(\alpha)$$

its spectral representation, with

$$1 = \int_{\sigma_A} P_\alpha d\mu(\alpha)$$

$$P_\alpha P_{\alpha'} = \delta_\mu(\alpha - \alpha')P_\alpha$$

Let then $f : \sigma \to \mathbb{C}$ be a "measurable" function on the spectrum of A; we define a *function of the operator* as

$$f(A) := \int_{\sigma_A} f(\alpha)P_\alpha d\mu(\alpha)$$

This definition is sound since for $f(\alpha) = \alpha^n$ it gives back the results expected from algebraic considerations: for $n = 0$ it gives back the completeness relation, for $n = 1$ the spectral representation and for $n = 2$

$$A^2 = \left(\int_\sigma \alpha P_\alpha d\mu(\alpha)\right)^2 = \int_\sigma\int_\sigma \alpha\alpha' P_\alpha P_{\alpha'} d\mu d\mu' = \int_\sigma \alpha\alpha' P_{\alpha'}\delta(\alpha-\alpha')d\mu d\mu' \equiv \int_\sigma \alpha^2 P_\alpha d\mu$$

The function

$$f(\alpha) = \delta_\mu(\alpha_0 - \alpha)$$

allows us to re-write the projector on the eigenvalue α_0 (either proper or improper, depending on whether α_0 is a proper or an improper eigenvalue),

$$\delta_\mu(\alpha_0 - A) := \int_\sigma \delta_\mu(\alpha_0 - \alpha) P_\alpha d\mu \equiv P(\alpha_0)$$

so, it is usual to write the completeness relation and the spectral representation in the form

$$1 = \int_{\sigma_A} \delta_\mu(\alpha - A) d\mu$$

$$A = \int_{\sigma_A} \alpha \delta_\mu(\alpha - A) d\mu$$

Example For a free particle in 3D one usually writes

$$H = \int_0^{+\infty} dE \delta(E - H)$$

with

$$\begin{aligned} \delta(E - H) &= \int d\hat{\boldsymbol{p}} \, |E\hat{\boldsymbol{p}}\rangle \langle E\hat{\boldsymbol{p}}| \equiv \int d\hat{\boldsymbol{p}} mp \, |\boldsymbol{p}_E\rangle \langle \boldsymbol{p}_E| \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} |Elm\rangle \langle Elm| \end{aligned}$$

where in the last line we have used the angular momentum eigenstates $|Elm\rangle$.

We can also check that if $0 \notin \sigma_A$ (*i.e.* if $\mathsf{Ker} A = \{0\}$) the function $f(\alpha) = \alpha^{-1}$ correctly defines the inverse

$$\left(\int_{\sigma_A} \alpha^{-1} P_\alpha d\mu \right) A = \int_{\sigma_A} \alpha^{-1} P_\alpha A d\mu = \int_{\sigma_A} \alpha^{-1} A P_\alpha d\mu \equiv \int_{\sigma_A} \alpha^{-1} \alpha d\mu = 1$$

where we have used $[A, P_\alpha] = 0$ and $AP_\alpha = \alpha P_\alpha$ which are easily proved for any operator which admits a spectral representation.

With the same token, one can define e^A, $(1-A)^{-1}$, etc. and verify that if A is bound and f analytic this definition is equivalent to the power-series representation. However, it is worth stressing here that the above definition of a function of operator based on its spectral representation is *more general*, since it does *not* require the analyticity of f to work; this is evindent, for instance, for the case above where we used $f(\alpha) = \delta_\mu(\alpha_0 - \alpha)$. Obviously, $f(\alpha)^*$ is related to $f(A)^\dagger$ since

$$f(A)^\dagger = \left(\int_\sigma f(\alpha) P_\alpha d\mu \right)^\dagger \equiv \left(\int_{\sigma_A} f(\alpha)^* P_\alpha d\mu \right) = f^*(A)$$

Furthermore, if f and g are functions of the same observable, it holds

$$[f(A), g(A)] = 0$$

as can be easily checked with a direct calculation (just notice that $[P_\alpha, P_{\alpha'}] = \delta_\mu(\alpha - \alpha')(P_\alpha - P_\alpha) \equiv 0$).

Conversely, it is easy to see that if two *observable* commute with each other they have a common set of eigenvectors. Indeed, suppose $[A, B] = 0$ and let $|\alpha\rangle$ be an eigenvector of A with eigenvalue α; we have

$$0 = [A, B] |\alpha\rangle = A (B |\alpha\rangle) - \alpha (B |\alpha\rangle)$$

that is, $B |\alpha\rangle$ is eigenvector of A with the same eigenvalue α. Now, if V^A_α is non-degenerate we simply have $B |\alpha\rangle = \beta |\alpha\rangle$ for some constant β ($\beta \in \mathbb{R}$ since $B = B^\dagger$) and thus $|\alpha\rangle$ is a common eigenvector. Otherwise, and more generally, consider B_α, the *restriction* of B to the Hilbert space V^A_α: such operator is still self-adjoint and its eigenvectors $|\beta; \alpha\rangle$ satisfy

$$B_\alpha |\beta; \alpha\rangle \equiv B |\beta; \alpha\rangle = \beta |\beta; \alpha\rangle$$

(with $\beta \in \mathbb{R}$) *i.e.* they are simultaneous eigenvectors of A *and* B.

In such cases, therefore, one can introduce a common set of eigenprojectors P_γ such that

$$P_\gamma P_{\gamma'} = \delta_\mu(\gamma - \gamma') P_\gamma$$
$$\int_\sigma P_\gamma d\mu_\gamma = 1$$
$$AP_\gamma = \alpha(\gamma) P_\gamma \qquad BP_\gamma = \beta(\gamma) P_\gamma$$
$$A = \int_\sigma \alpha(\gamma) P_\gamma d\mu_\gamma \quad B = \int_\sigma \beta(\gamma) P_\gamma d\mu_\gamma$$

(for some appropriate set σ and functions $\alpha : \sigma \to \sigma_A$ and $\beta : \sigma \to \sigma_B$) and consider the functions of *both* observables

$$f(A, B) := \int_\sigma f(\alpha(\gamma), \beta(\gamma) P_\gamma d\mu_\gamma$$

(A and B are special cases with $f(\alpha, \beta) = \alpha$ and $f(\alpha, \beta) = \beta$, respectively).

With the same token, if $[A, B] = 0$ *and every* eigenvector of A is eigenvector of B then $B \equiv f(A)$ for some function f. Indeed, we have seen above that if $|\alpha\rangle$ is an eigenvector of A, the same is true for $B |\alpha\rangle$. If V^A_α's are non-degenerate, we simply have the desired result, since $B |\alpha\rangle \equiv \beta(\alpha) |\alpha\rangle$ holds for any α; otherwise, we have

$$B |\beta_n; \alpha\rangle = \beta_n(\alpha) |\beta_n; \alpha\rangle$$

for *any* $\{|\beta_n;\alpha\rangle\} \subset V_\alpha^A$, *i.e.* *however* we choose a complete, orthonormal set $\{|\beta_n;\alpha\rangle\}$ of vectors in V_α^A. Thus, for any such complete set, and *arbitrary* coefficients $\{..c_n..\} = \boldsymbol{c} \in l^2(\mathbb{C})$

$$\sum_n c_n\beta_n |\beta_n;\alpha\rangle = B\left(\sum_n c_n |\beta_n;\alpha\rangle\right) = \beta(\boldsymbol{c};\alpha)\left(\sum_n c_n |\beta_n;\alpha\rangle\right)$$

that is

$$c_n(\beta_n(\alpha) - \beta(\boldsymbol{c};\alpha)) = 0 \quad n \in \mathbb{N}$$

from which it follows that $\beta_1 = \beta_2 = .. = \beta(\boldsymbol{c};\alpha)$ must be independent on the vectors and can only depend on α, *i.e.* again $\beta = \beta(\alpha)$.

In general, if $\{A_i\}_{i=1}^N$ is a set of commuting observables, $[A_i, A_j] = 0$, we say that the set is *complete* if any other commuting observable is a function of these observables,

$$[A_i, B] = 0 \quad i = 1, 2, ..N \;\Rightarrow\; B = f(A_1, A_2, ..A_N)$$

Equivalently, the set is complete if and only if each common eigenspace is *non-degenerate*. Indeed, if $V_{\boldsymbol{\alpha}}$ ($\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ..\alpha_N\}$) were such degenerate eigenspace common to all the observables in the set, the projectors $P_{\boldsymbol{\alpha};1}$,$P_{\boldsymbol{\alpha};2}$..$P_{\boldsymbol{\alpha};g}$ defined by any basis in $V_{\boldsymbol{\alpha}}$ would define operators that commute with all A_i's *without* being a function of A_i's. Conversely, if $V_{\boldsymbol{\alpha}}$ is non-degenerate and B commutes with any A_i we can write

$$B\,|\alpha_1, \alpha_2, ..\alpha_N\rangle = \beta(\alpha_1, \alpha_2, ..\alpha_N)\,|\alpha_1, \alpha_2, ..\alpha_N\rangle$$

thereby showing that B is necessarily a function of A_i's.