Instruction on how to use these notes.

These notes must be used with *care*: they are preliminary and their aim is just to provide a concise summary of the topics covered during the lectures. Only few detailed step-by-step derivations are included and the general approach is to explain the general concepts through examples.

Please try to complement these notes by reading the relevant parts of a Quantum Mechanics book such as

- Modern quantum mechanics J.J. Sakurai. QM Library: QC174.1 SAK
- Quantum mechanics Cohen-Tannoudj, Diu, Laloe. QM Library: QC174.1 COH

The parts written in small fonts lie outside the main syllabus of the course: they are not strictly necessary for the rest of the material presented nor will appear in the exam paper – these parts are included just for your own curiosity.

`Please let me know if you have any question regarding these notes, spot typos or mistakes.`

1 Some useful algebraic structures.

1.1 Groups.

A *group* is a set of elements G together with an operation $\bullet$ that combines any two elements and gives another element of the group.

- Closure: $\forall\ a, b \in G,\ a \bullet b \in G$.
- Associativity: $\forall\ a, b, c \in G,\ (a \bullet b) \bullet c = a \bullet (b \bullet c)$.
- Existence of the identity element: $\exists\ e \in G$ such that, $\forall\ a \in G,\ a \bullet e = e \bullet a = a$
- Existence of the inverse: $\forall\ a \in G\ ,\ \exists\ b \in G$ such that $a \bullet b = b \bullet a = e$.

The concept of group is particularly important in physics because the set of symmetries of a physical system is a group. In this case the product of two elements consists just in performing the two symmetry operation in sequence: the result is a possibly new operation that leaves the system invariant.

1.1.1 Examples and exercises.

- The set of integer numbers $\mathbf{Z}$ (that is $\ldots, -2, -1, 0, 1, 2, \ldots$), together with the standard addition form a group $(\mathbf{Z}, +)$

This group enjoys an additional property, that is the operation is commutative: $\forall\ a, b \in \mathbf{Z}$ we have $a + b = b + a$. This type of groups is called <u>Abelian</u>.

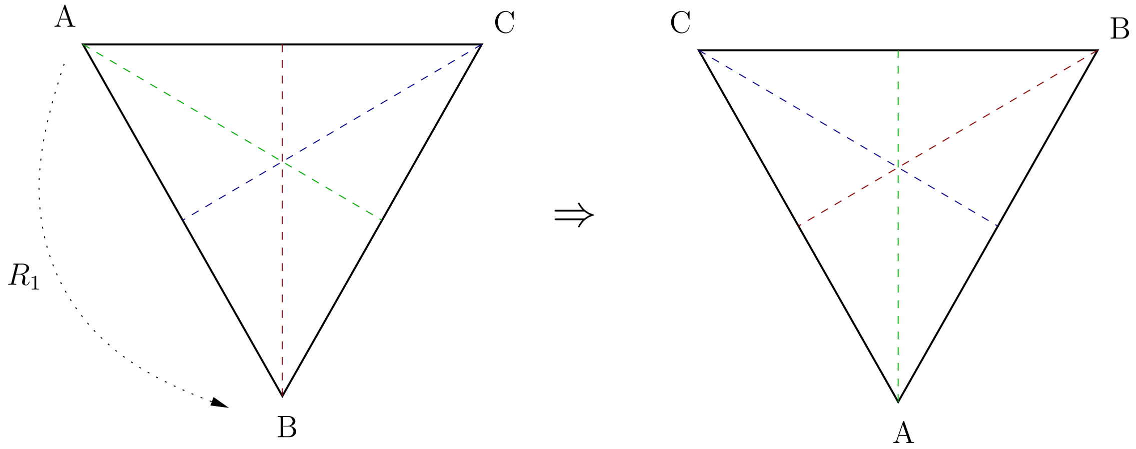
- Consider an equilateral triangle: a rotation by 120 degrees ($2\pi/3$ radians) around the center of the triangle leaves the object invariant.

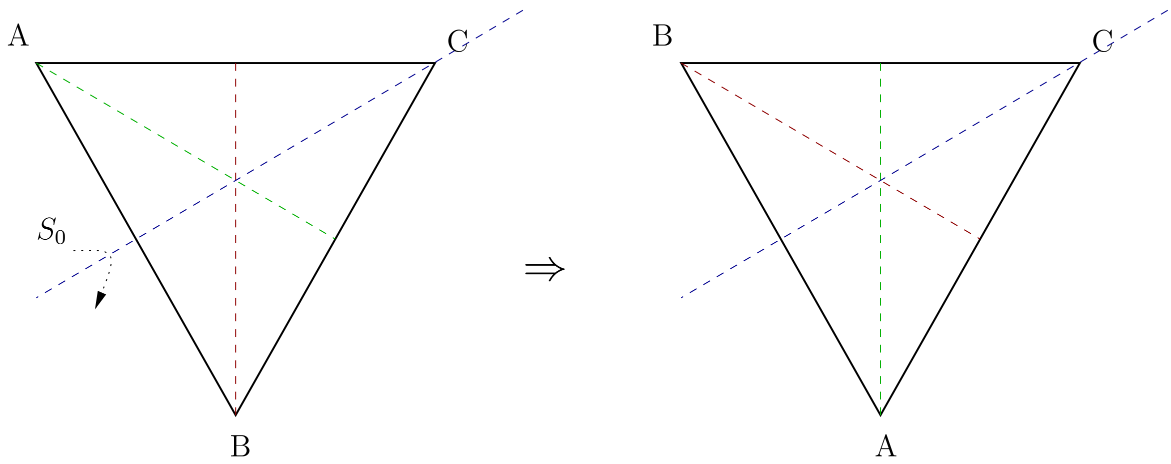
Notice that these operations are *not* the only symmetries of the triangle! We can perform also reflection along the three altitudes and leave the triangle unchanged. The group generated by all symmetries is called D_3 (it's one of the Dihedral groups). This group contains a finite number of elements and is not Abelian (non-Abelian). See exercise below.

- Consider a sphere: any rotation around an axis passing through the origin of the sphere leaves the sphere unchanged. The set of all these rotations forms a group that we will analyze in some detail in this course.

This group contains an infinite numbers of elements, since the angle of the rotation is a continuous parameter. As we will see this group is non-Abelian.

Exercise: Consider the group D_3. This group is generated by the following two operations:





1. Compose these two symmetries in all possible ways and write down all elements of D_3. How many elements are contained in D_3 ?

2. Write down all possible products between two elements in D_3 and prove that the group is non-Abelian.

1.2 Vector spaces.

A vector space is a set of elements that can be summed and rescaled: it represents an abstraction of the usual Euclidean space. To be precise, a *real vector space* V is an Abelian group with respects to the addition; moreover each element of $v \in V$ can be multiplied (rescaled) by a real number a and $av \in V$. This scalar multiplication must have the following properties:

- Distributivity with respect to the vector addition: $\forall\, a \in \mathbf{R}$ and $\forall\, v,\ w \in V$ we have $a(v+w) = av + aw$;
- Distributivity with respect to the real number addition: $\forall\, a,\ b \in \mathbf{R}$ and $\forall\, v \in V$ we have $(a+b)v = av + bv$;
- $\forall\, a,\ b \in \mathbf{R}$ and $\forall\, v \in V$ we have $a(bv) = (ab)v$;
- Multiplication by the identity and zero: $\forall\, v \in V$ we have $1v = v$ and $0v = 0$.

A *complex vector space* is defined in a similar way just by allowing rescaling of the vectors with complex, instead of real numbers. Then in the axioms above a, b will belong to $\mathbf{C}$.

Vector spaces play a central role in physics. In classical mechanics the position of a point-particle is specified by a vector in the Euclidean space. In Quantum mechanics the state of a system is specified by the wavefunction which, as we will see, is an element in a complex vector space.

Let me now recall the concept of linearly independent vectors

- A set of vectors $\{v_1, v_2, \ldots, v_m\}$ is *linearly independent* if $a_1 v_1 + \ldots + a_m v_m = 0$ (with $a_i \in \mathbf{R}$ for real vector spaces, while $a_i \in \mathbf{C}$ for complex spaces) implies that $a_1 = \ldots = a_m = 0$.
- If it exists a maximum number of linearly independent vectors n, then n is the dimension of the vector space. A set of n linearly independent vectors is called *basis*.

This implies that a vector space V of dimension n can be "represented"[1] as the Euclidean space ($\mathbf{R}^n$ for real vector spaces and $\mathbf{C}^n$ for complex ones). Consider a basis $\{v_1, v_2, \ldots, v_n\}$ for this space; then we have the following

<u>Theorem</u>: each vector $v \in V$ has a *unique* decomposition in terms of $\{v_1, v_2, \ldots, v_n\}$

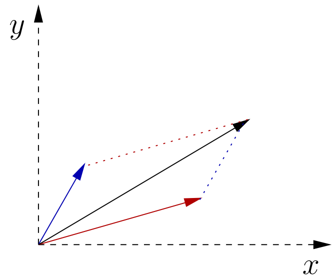
$$v = a_1 v_1 + \ldots + a_n v_n \ . \tag{1.1}$$

Proof "by contradiction". Suppose that there are two different such decompositions: $v = \sum_{i=1}^n a_i v_i$ and $v = \sum_{i=1}^n b_i v_i$ with $a_i \neq b_i$ at least for one value of i. Then we can take the difference between these two decompositions and get: $0 = \sum_{i=1}^n (a_i - b_i) v_i$. This contradicts the hypothesis that $\{v_1, v_2, \ldots, v_n\}$ forms a basis. The n numbers a_i are called *coordinates* of v in the basis $\{v_1, v_2, \ldots, v_n\}$.

[1] We will see next week what this means exactly.

1.2.1 Examples and exercises.

- The 2-dimensional Euclidean space is the standard example of a vector space.



The 2-dimensional Euclidean space: the elements of this real vector space are arrows on a plane whose length can be rescaled and that can be summed in the usual way.

<u>**Exercises**</u>:

1) Consider the set of polynomial with real coefficients of degree 2: $3x+4$ and $x^2+\sqrt{2}$ are examples of such polynomials and $\sum_{i=0}^{2} a_i x^i \equiv a_2 x^2 + a_1 x + a_0$ with $a_i \in \mathbf{R}$ is the most general element.

- Show that this set forms a vector space with the standard addition between polynomial and with the scalar multiplication with any real number b defined as: $b(\sum_{i=0}^{2} a_i x^i) = \sum_{i=0}^{2}(b a_i) x^i$.
- What is the dimension of this vector space ?

2) Consider the following pairs of vectors in $\mathbf{C}^3$:

$$a) \qquad v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad , \quad w = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \quad ; \tag{1.2}$$

$$b) \qquad v = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad , \quad w = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} \quad ; \tag{1.3}$$

$$c) \qquad v = \begin{pmatrix} 1 \\ i \\ -i \end{pmatrix} \quad , \quad w = \begin{pmatrix} i \\ -1 \\ 1 \end{pmatrix} \quad . \tag{1.4}$$

Check whether these pairs of vectors linearly are linearly independent.

1.3 Scalar products and Hilbert Spaces.

We are interested in vector spaces that have an additional structure: a **scalar product** (sometimes I will use the equivalent denomination "inner product"). The inner product is a bilinear map from $V \times V \to \mathbf{R}$ ($V \times V \to \mathbf{C}$ for complex vector spaces). Let me focus on the complex case:

- linearity: $\forall\ a_i \in \mathbf{C}$ and $\forall\ w, v_i \in V$ we have $(w, a_1 v_1 + a_2 v_2) = a_1 (w, v_1) + a_2 (w, v_2)$ and $(a_1 v_1 + a_2 v_2, w) = \bar{a}_1 (v_1, w) + \bar{a}_2 (v_2, w)$
- conjugation symmetry: $\forall\ w, v \in V$ we have $(w, v) = \overline{(v, w)}$.

Notice that this last property implies that $(v, v) \in \mathbf{R}\ \forall\ v \in V$. Then it makes sense to require that

- $\forall\ v \neq 0$ in V we have $(v, v) > 0$

When this additional property is satisfied the scalar product is said to be *positive definite*. The scalar product of a vector with itself is called *norm*: $||v||^2 \equiv (v, v)$. In our applications to Quantum Mechanics we will be focusing on positive definite scalar products. On the contrary, for instance in special relativity one deals with a vector space with a non-positive definite scalar product.

Some useful definitions and properties:

- If two vectors v_1, v_2 have vanishing scalar product $(v_1, v_2) = 0$, then they are said to be orthogonal.
- An orthogonal basis for V is a basis $\{v_1, v_2, \ldots\}$ for which $(v_i, v_j) = 0\ \forall i \neq j$. If in addition we have $(v_i, v_i) = 1\ \forall i$, then the basis is called orthonormal. In symbols we have $(v_i, v_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.
- The decomposition (1.1) of any vector in terms of an orthonormal basis is given by

$$v = \sum (v_i, v) v_i \quad \Rightarrow \quad a_i = (v_i, v)\ . \tag{1.5}$$

Theorem (Schwarz inequality). Take any two vectors $v_1,\ v_2$ of a Hilbert space. Then we have

$$|(v_1, v_2)|^2 \leq ||v_1||^2 ||v_2||^2\ . \tag{1.6}$$

Proof: Consider the vector $w = v_1 + a v_2$, where a is an arbitrary complex (real) number. Then we have $||w||^2 \geq 0$, which implies

$$0 \leq ||w||^2 = (v_1, v_1) + |a|^2 (v_2, v_2) + a (v_1, v_2) + \bar{a} (v_2, v_1)\ . \tag{1.7}$$

Now we can take

$$a = -\frac{(v_2, v_1)}{(v_2, v_2)}\ . \tag{1.8}$$

In this case the last two terms of (1.7) become equal and opposite to the second one. Then we can immediately see that (1.7) reduces to (1.6).

Roughly speaking, a **Hilbert space** is a vector space with a positive definite inner product.

If the vector space is infinite dimensional, we also require that:

1. The norm of each vector is finite: $\forall\ v$ we have $(v, v) < \infty$.
2. Any Cauchy sequence of vectors has a limit vector in V. A sequence of vectors v_k with $k = 1, 2, \ldots$ is Cauchy if $||v_m - v_n||^2$ becomes arbitrary small when m, n are big. In formal terms: $\forall\ \epsilon > 0\ \exists\, k \in \mathbf{N}$ such that $\forall\ m, n > k$ we have $||v_m - v_n||^2 < \epsilon$. Spaces satisfying this condition are said to be Cauchy complete.

   In order to appreciate the meaning of the Cauchy completeness, let us apply to the case of the set of numbers, which is simpler than an infinite dimensional vector space, but still capture the essence of this requirement. Consider the set of rational numbers, i.e. the number that can be expressed as ratio of two integers. You can prove (by contradiction) that $\sqrt{2}$ is not a rational number. However it is easy to find a sequence of rational number that are closer and closer to $\sqrt{2}$. Consider, for instance, the binomial expansion of

   $$2\sqrt{1-\frac{1}{2}} = \lim_{N\to\infty} 2\left(1 - \sum_{n=1}^{N} \frac{c_n}{n!}\left(\frac{1}{2}\right)^n\right)$$

   where $c_1 = \frac{1}{2}$ and $c_n = \frac{(1)(3)\ldots(2n-3)}{2^n}$ for $n > 1$. If we take a large, but fixed N, the r.h.s. is rational, as it is a sum of rational numbers. Choosing bigger values of N makes the sum closer to $\sqrt{2}$, so the sequence of numbers labelled by N is Cauchy. However its limit is $\sqrt{2}$ and thus it is outside the space of rational number. So the space of rational number is not Cauchy complete and to satisfy this requirement we need to consider real numbers. Hilbert spaces share the same Cauchy completeness property of real numbers.
3. The space has a countable orthonormal basis. Hilbert spaces satisfying this requirement are often called *separable*. Mathematicians consider also non-separable Hilbert spaces which satisfy the first two requirements, but not the last one; these non-separable spaces do not arise in Quantum Mechanics and so we will ignore them .

1.3.1 Examples and exercises.

- Consider the infinite dimensional generalization of the vectors in $\mathbf{C}^3$, that is the vectors v are just infinite arrays of complex numbers: $v = (v_1, v_2, v_3, v_4, \ldots)$. The scalar product between two vectors of this kind is defined to be $(w, v) = \sum_{k=1}^{\infty} \bar{w}_k v_k$. Thus in order to satisfy property 1, we focus only on the vectors for which $\sum_{k=1}^{\infty} |v|_k^2 < \infty$. One can show that this set of vectors forms a Hilbert space (that is property 2 and 3 are satisfied). This Hilbert space is usually named l^2.

  Exercise: Consider an set $\{v_1, v_2, \ldots\}$ of orthogonal vectors. Show that these vectors are linearly independent.

2 Linear maps.

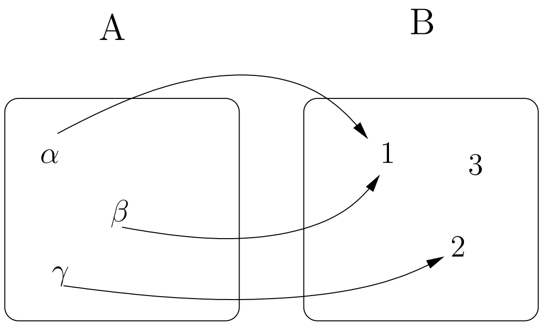
2.1 Functions between two sets.

A function between two sets ($f : A \to B$) is a map that associates each element of the first set (A) one element of the second set (B).

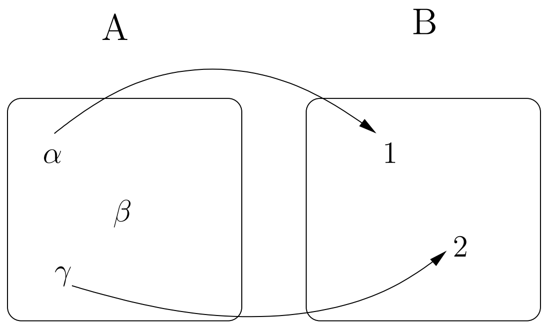
In a formal language: $\forall\ a \in A\ \exists!\ b \in B$ such that $f(a) = b$. Some definitions:

- A function $f : A \to B$ is said surjective (or "onto") if all elements of B are images of some element in A: $\forall\ b \in B\ \exists a \in A$ such that $f(a) = b$
- A function $f : A \to B$ is said injective if any two elements in A have different images in B: $\forall\ a_1, a_2 \in A\ f(a_1) = f(a_2)$ implies $a_1 = a_2$.
- a function is bijective if it is both surjective and injective.

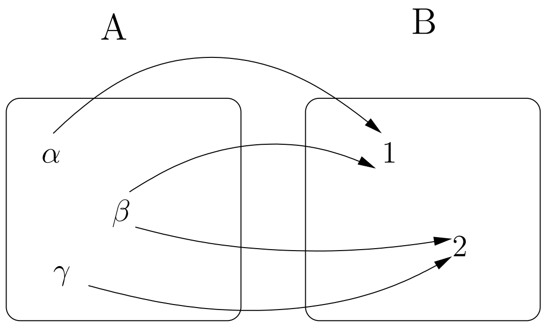
2.1.1 Examples.



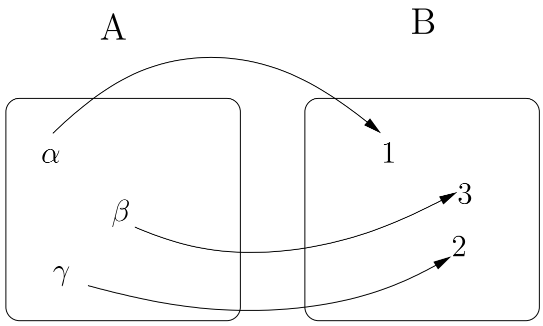
This is a function.



This is not a function



This is not a function



This function is bijective.

2.2 Linear functionals and Dirac's notation.

From now on we will very often indicate the vectors of an abstract vector space by using Dirac's notation $|\psi\rangle$ (*ket vector*). Depending on the type of the vector space V one is considering $|v\rangle$ can be a n-tuple of number, a polynomial or an element of l^2 (see the example in the previous Section). This notation is useful when one considers the dual space of linear functionals.

A linear functional is a function χ between a real (complex) vector space V and the real (complex) numbers which has the following property

- $\forall\ |\psi_1\rangle,\ |\psi_2\rangle \in V$ and $\forall\ a,\ b \in \mathbf{R}\ (\mathbf{C})$ we have $\chi(a|\psi_1\rangle + b|\psi_2\rangle) = a\chi(|\psi_1\rangle) + b\chi(|\psi_2\rangle)$.

The dual space V^* is the set of all possible linear functionals. In the Dirac's notation each linear functional is represented by a $\langle\chi|$ (*bra vector*).

To every ket corresponds a bra. In a vector space with a scalar product, it is easy to define a function that maps the vectors of V into elements of V^*. For each $|\phi\rangle \in V$ consider the scalar product between $|\phi\rangle$ and any other element of V. This is a linear functional mapping V in $\mathbf{R}$ ($\mathbf{C}$) that is completely specified by $|\phi\rangle$. Thus we can represent this linear functional with $\langle\phi|$

$$(|\phi\rangle, |\psi\rangle) \equiv \langle\phi|\psi\rangle\ . \tag{2.1}$$

From now on we will often indicate the scalar products between two vectors by using Dirac's notation, that is by using the left hand side of Eq. (2.1). Let us recall the main properties of the scalar product

$$\begin{aligned} \langle\phi|\chi\rangle &= \overline{\langle\chi|\phi\rangle}\ , \\ \langle\phi|a_1\chi_1 + a_2\chi_2\rangle &= a_1\langle\phi|\chi_1\rangle + a_2\langle\phi|\chi_2\rangle\ , \\ \langle a_1\phi_1 + a_2\phi_2|\chi\rangle &= \bar{a}_1\langle\phi_1|\chi\rangle + \bar{a}_2\langle\phi_2|\chi\rangle\ , \\ \langle\chi|\chi\rangle &> 0\ , \quad \forall\ |\chi\rangle \neq 0\ . \end{aligned} \tag{2.2}$$

The correspondence between ket and bra vectors is *antilinear*: if the bra vectors corresponding to $|\psi_1\rangle$ and $|\psi_2\rangle$ are $\langle\psi_1|$ and $\langle\psi_2|$, then the bra vector corresponding to $a_1|\psi_1\rangle + a_2|\psi_2\rangle$ is $\bar{a}_1\langle\psi_1| + \bar{a}_2\langle\psi_2|$

$$a_1|\psi_1\rangle + a_2|\psi_2\rangle \Rightarrow \bar{a}_1\langle\psi_1| + \bar{a}_2\langle\psi_2|\ . \tag{2.3}$$

Question: Is there a ket corresponding to every bra?
The answer is yes for finite dimensional vector spaces with a scalar product, while for infinite dimensional spaces the situation is subtler.

Let us first focus on the simple finite dimensional case. We can choose an orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle\}$ which means that we have $\langle\psi_j|\psi_i\rangle = \delta_{ij}$. For each linear functional $\langle\chi|$ we can build a vector as follows

$$\sum_{i=1}^{n} \overline{\langle\chi|\psi_i\rangle}\ |\psi_i\rangle \equiv |\chi\rangle\ . \tag{2.4}$$

Notice that the bra associated to ket just defined in (2.4) is the original linear functional $\langle\chi|$, as it can be seen by using (2.1). Proof: take any vector $|\phi\rangle$, this can be decomposed in a unique way on the basis $|\psi_i\rangle$ ($|\phi\rangle = \sum c_i|\psi_i\rangle$); then the bra associated to $|\chi\rangle$ acts as follow

$$(|\chi\rangle, |\phi\rangle) = \sum_{i=1}^{n} \langle\chi|\psi_i\rangle \; (|\psi_i\rangle, |\phi\rangle) = \sum_{i,j=1}^{n} \langle\chi|\psi_i\rangle\delta_{ij}c_j = \langle\chi|\phi\rangle \; . \tag{2.5}$$

Now the question is: what can go wrong in the case of an infinite dimensional vector space V? In this case the sum in (2.4) becomes an infinite series and the problem is that this series might not have a limit in V even if it is a combination of vectors in V. It is possible to construct an explicit example of such a situation in the case V is *not* a Hilbert space and in particular does not satisfy the second requirement in the previous notes. This example is important for our applications to Quantum Mechanics (see below the example about Dirac's delta for some more details).

2.2.1 Examples and exercises.

- Bra for finite dimensional Hilbert spaces.

Consider the space $\mathbf{C}^3$: the element of this space are just column vectors with a triplet of complex numbers (see for instance 1.2). We use the standard scalar product

$$(|v\rangle, |w\rangle) \equiv \bar{v}_1 w_1 + \bar{v}_2 w_2 + \bar{v}_3 w_3 = \begin{pmatrix} \bar{v}_1 & \bar{v}_2 & \bar{v}_3 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \; . \tag{2.6}$$

Then, by using Eq. (2.1), it is clear that the bra corresponding to the vector $|v\rangle$ is simply the row-vector $\begin{pmatrix} \bar{v}_1 & \bar{v}_2 & \bar{v}_3 \end{pmatrix}$.

- Dirac's delta.

Consider the space of the following "nice" function $f : \mathbf{R} \to \mathbf{C}$: f is infinitely differentiable and goes to zero very quickly[2] as $|x| \to \infty$. This set of functions form a complex vector space[3] V_f with scalar product defined as

$$\int_{-\infty}^{\infty} \bar{g}(x) f(x) dx \; . \tag{2.7}$$

[2] $\forall\; n, m = 1, 2, \ldots$ we have $|x^n d^m f/dx^m| \to 0$ as $|x| \to \infty$.

[3] See the exercise below.

Consider the mapping $f(x) \to f(x=0)$. This is a linear functional[4] that we will call $\langle\delta_0|$. We can represent this functional by using the scalar product (2.7) and the Dirac's delta

$$\langle\delta_0|f\rangle \equiv \int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \ . \tag{2.8}$$

Notice that there is *no* standard function g for which $\int_{-\infty}^{\infty} \bar{g}(x) f(x) = f(0)$ for any $f(x) \in V_f$. This shows that this linear functional cannot be represented by the scalar product of an element in V, but requires a new object (the Dirac's delta in this case).

Exercise

- Prove the statement: V_f is a vector space.
- Prove the statement: $\langle\delta_0|$ is a linear functional.

* Show explicitly that V_f is not a Hilbert space.

2.3 Linear operators.

Consider two vector spaces V and W (they are not necessarily different, we can have $V = W$). A linear operator is a function A from V to W satisfying

- $\forall |\chi_i\rangle \in V$ we have $A(|a_1\chi_1 + a_2\chi_2\rangle) = a_1 A(|\chi_1\rangle) + a_2 A(|\chi_2\rangle) \in V'$

In the case $V = W$ we can define the product of two operators in a simple way just by acting on the vectors in an ordered way:

$$\forall\ \chi \in V \quad AB(|\chi\rangle) \equiv A\Big(B(|\chi\rangle)\Big) \ .$$

Any real (complex) finite dimensional vector space V is isomorphic to $\mathbf{R}^n$ ($\mathbf{C}^n$). This means that there is a injective linear map between V and $\mathbf{R}^n$ (or $\mathbf{C}^n$).

In order to see this let us take a basis for V: $\{|v_1\rangle, \ldots, |v_n\rangle\}$. Then any vector $|v\rangle \in V$ can be decomposed along this basis

$$|v\rangle = \sum_i c^i |v_i\rangle \quad \Rightarrow \quad \begin{pmatrix} c^1 \\ c^2 \\ \vdots \\ c^n \end{pmatrix} \leftrightarrow |v\rangle \ , \tag{2.9}$$

where c^i are the coordinates[5]. Thus we can associate to any $|v\rangle \in V$ a unique n-tuple of numbers; vice-versa to any n-tuple of numbers we can associate a vector simply be

[4] See the exercise below.

[5] As a notation, from now on we will use upper indices for the vector coordinates; you will see why this is convenient.

reading (2.9) in the opposite sense. Thus the map is injective. In order to complete the proof that this is an isomorphism between V and $\mathbf{R}^n$ (or $\mathbf{C}^n$), see the first exercise below.

2.3.1 Examples and exercises.

- In the case of finite dimensional vector spaces, any linear operator $A : V \to V'$ can be represented by a matrix.

In order to see this let us take a basis for V $\{|v_1\rangle, \ldots, |v_n\rangle\}$ and one for W $\{|w_1\rangle, \ldots, |w_m\rangle\}$. Then we have

$$A|v\rangle = \sum_{i=1}^{n} c^i A|v_i\rangle \ . \tag{2.10}$$

Now let us focus on each $A|v_i\rangle$: these vectors belong to W so they can be decomposed along the $|w_j\rangle$ basis

$$A|v_i\rangle = \sum_{j=1}^{m} |w_j\rangle a^j{}_i \quad \Rightarrow \quad A|v\rangle = \sum_{i=1}^{n}\sum_{j=1}^{m} a^j{}_i c^i \ |w_j\rangle \ . \tag{2.11}$$

By using the isomorphism introduced above, we can

- represent the kets $|v\rangle \in V$ as column vectors with n numbers,
- represent the kets $|W\rangle \in W$ as column vectors with m numbers,
- represent the linear operator $\hat{A}$ as the matrix $a^j{}_i$

$$\hat{A} \leftrightarrow a^j{}_i \equiv \langle v_j|\hat{A}|v_i\rangle$$

  where j is the row index and i is the column index.

- A projector is a linear operator P from a vector space V to itself ($P : V \to V$) such that $P^2 = P$. This definition generalized to an abstract vector space the idea of projection in the standard Euclidean space. Dirac's notation provides a simple way to write projectors in a simple way. Consider a vector with norm 1: $|v\rangle$ and consider also the associated bra $\langle v|$. We can define a projector $P_v \equiv |v\rangle\langle v|$ which acts as follow

$$\forall \ |w\rangle \in V \quad P_v|w\rangle \equiv (\langle v|w\rangle) \ |v\rangle \ . \tag{2.12}$$

  This is a projector since $P_v^2 = |v\rangle\langle v|v\rangle\langle v| = P_v$. If you take V to be, for instance, the standard 2-dimensional vector space and $|v\rangle$ to be the versor along the x-axis, then you can see that (2.12) is indeed the projection on this axis.

It is straightforward to generalize (2.12) when you deal with more vectors: if you have a set of orthonormal vectors $\{|v_1\rangle, \ldots, |v_m\rangle\}$, then you can define a projector on the plane generated by these vectors as follow

$$P_m = \sum_{i=1}^{m} |v_i\rangle\langle v_i| \ . \tag{2.13}$$

Exercise

1) Prove that the map in (2.9) is linear.

2) Consider a set of orthonormal vectors $|v_1\rangle, \ldots$ that forms a basis for V. Prove that the associated projector, as in (2.13), is the identity operator: $P = 1$.

3) Consider a finite dimensional complex Hilbert space and the isomorphism (2.9) with $\mathbf{C}^n$. Derive how the scalar product between two vectors $|v\rangle$ and $|w\rangle$ is written in terms of their coordinates.

2.3.2 Action of a linear operator on a bra.

So far we have discussed the action of linear operators on the kets in a vector space V. Let us focus on operators from V to V. By using the scalar product, it is simple to define an action also on the linear functionals (that is the bras). Consider a linear operator A, then for any bra $\langle\phi|$ we can associate a new bra $\langle\phi'|$ defined as follow

$$\forall\ |\psi\rangle \in V \qquad \langle\phi'|\psi\rangle \equiv \langle\phi|\,(A|\psi\rangle) = \langle\phi|A|\psi\rangle \ . \tag{2.14}$$

The correspondence $\langle\phi| \to \langle\phi'| \equiv \langle\phi|A$ is linear:

$$(a_1\langle\phi_1| + a_2\langle\phi_2|)A = a_1\langle\phi_1|A + a_2\langle\phi_2|A \ . \tag{2.15}$$

2.3.3 Examples.

- We have seen that in the case of a finite dimensional Hilbert space, linear operators can be represented by standard matrices. The action of an operator on a vector is then represented by the standard left multiplication of the corresponding matrix on the vector coordinates $A|v\rangle \to \sum_i a^j{}_i c^i$. The action on the bras defined above corresponds to right matrix multiplication: $\langle v|A \to \sum_j \bar{c}_j a^j{}_i$.

2.4 Hermitian and self-Adjoint operators.

Let us consider an operator A defined in a subvector space $W \subseteq V$ (of course we might have $W = V$). If for any pair of vectors $|v_1\rangle$, $|v_2\rangle \in W$ we have that $\langle Av_1|v_2\rangle = \langle v_1|Av_2\rangle$ then A is a Hermitian operator.

By using the previous paragraph A defines also a linear operator on the bras. Now, by using the bra/ket relation, we can define a new operator $A^\dagger$ on the kets ($A^\dagger$ is called the *adjoint* of A): $\forall\, |\psi\rangle \in V$ we define $A^\dagger|\psi\rangle$ as the ket corresponding[6] to the linear functional $\langle\psi|A$ acting on W. Even if this definition might seem abstract we will see that it is just the generalization of the standard Hermitian conjugation for (possibly infinite dimensional) Hilbert spaces. Let me summarize the main properties of the Adjoint operation:

$$(A+B)^\dagger = A^\dagger + B^\dagger \;, \tag{2.16}$$

$$\forall a \in \mathbf{C} \quad \Rightarrow \quad (aA)^\dagger = \bar{a}A^\dagger \;, \tag{2.17}$$

$$(A\,B)^\dagger = B^\dagger A^\dagger \;. \tag{2.18}$$

The first two properties follow from the linearity of A and B on the bra and from the antilinearity of the bra/ket relation. Eq. (2.18)

An operator is Hermitian if $A^\dagger = A$.

Subtlety: notice that in this definition I have been vague on the domain of definition of A and $A^\dagger$. In the case of infinite dimensional spaces the domains of the two operators might be different, even if the $A^\dagger$ and A are equal on the common part of the two domains. A self-adjoint operator is an Hermitian operator for which the domains of A and $A^\dagger$ are also equal. Except for the example below, we will use the words "Hermitean" and "self-adjoint" as equivalent and asssume that there are no subtelties with the definition of the domain of the operators.

2.4.1 Examples and exercises.

- Consider a finite dimensional Hilbert space V. We know that is isomorphic to $\mathbf{C}^n$ and that each linear operator A in V is mapped in a matrix $a^j{}_i$ in $\mathbf{C}^n$. Then Hermitian operators just correspond to Hermitian matrices. Moreover there are no subtleties with the definitions of the domains (as they always coincide with the whole vector space); then the operators corresponding to Hermitian matrices are also self-adjoint.

[6]There is a subtlety: there might be no corresponding ket; in this case we just eliminate $|\psi\rangle$ from the domain of $A^\dagger$.

- In the case of infinite dimensional vector spaces, one has to pay some attention to the domain where the linear operator are defined. For instance, consider the space V of smooth functions $(\psi(x))$ in $\mathbf{R}$ which are square integrable and the operator position (x): it is not guaranteed that $x\psi(x)$ is an element of V and so x is not defined over the whole V. Example: $\psi(x) = A\frac{x}{1+x^2}$.
- A Hermitian, but not self-adjoint operator. Consider the wavefunctions you have seen in the problem of the infinite potential well

$$|n\rangle \equiv \sin\left(\frac{n\pi x}{L}\right) \ , \quad \text{with} \ \ x \in [0, L] \ .$$

  Consider the vector space W generated by any finite linear combination of these functions with the standard scalar product $\int_0^L \bar{g}(x)\, f(x)\, dx$. Then the operator $P = -i\frac{d}{dx}$ is Hermitian, but cannot be extended to be a self-adjoint operator.

Exercise

- Prove that $\langle\phi|A^\dagger|\psi\rangle = \overline{\langle\psi|A|\phi\rangle}$ (of course suppose that $|\phi\rangle$ belongs to the domains of A and $|\psi\rangle$ to the domain of $A^\dagger$).

3 Eigenvalues and Eigenvectors.

3.1 Definition.

Let complex V be a vector space. Consider a linear operator A from V to itself $A : V \to V$. A vector $|v\rangle \in V$ that satisfy

$$A|v\rangle = \lambda|v\rangle, \tag{3.1}$$

for some complex number λ is called eigenvector and λ is called eigenvalue[7]. Of course, since A is linear, one can rescale $|v\rangle$ by an arbitrary number (as in $|v'\rangle = c|v\rangle$) and build a new eigenvector ($|v'\rangle$) with the same eigenvalue. It is also possible that a vector $|w\rangle$, that is linearly independent from $|v\rangle$, is an eigenvector with the same eigenvalue (that is we might have $A|w\rangle = \lambda|w\rangle$). It is straightforward to prove that the set of all eigenvectors with the same eigenvalue form a vector space (called eigenspace) that is a subspace of V.

[**Revision from MT2/MT3**]. If V is a finite dimensional vector space, we have a clear algorithm to find the eigenvalues and the eigenvectors.

- Any finite dimensional vector space V is isomorphic to $\mathbf{C}^n$ and any linear operator from V to itself can be represented as a matrix $a^j{}_i$ acting on $\mathbf{C}^n$.
- The eigenvalues are the solutions of the following equation: $\det\left(a^j{}_i - \lambda\delta^j{}_i\right) = 0$ (this is a polynomial equation whose degree is equal to the dimension of the vector space).

[7] The trivial solution $|v\rangle = 0$ is neglected and does not count as an eigenvector.

- For each eigenvalue we can find the corresponding eigenvector by solving the following set of n linear equations

$$\sum_{i=1}^{n} a^j{}_i c^i - \lambda c^j = 0 \ , \qquad j = 1, 2, \ldots, n \ . \tag{3.2}$$

Notice that eigenvectors with *different* eigenvalues form a set of linearly independent vectors.

3.1.1 Examples and exercises.

Exercise. Consider the following matrices as operators form $\mathbf{C}^2$ to itself

$$M = \begin{pmatrix} 1 & -i \\ 0 & 1 \end{pmatrix} \ , \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{3.3}$$

Find all possible eigenvectors and the corresponding eigenvalues for M and the σ_i's.

3.2 Eigenvectors of self-adjoint operators.

Let us start from the case of finite dimensional vector spaces that you studied in MT2/MT3. In this case, self-adjoint operators can be represented simply as Hermitian matrices. We have the following theorem.

> [T1] The eigenvectors of an Hermitian matrix A form a complete basis for $\mathbf{C}^n$ and the corresponding eigenvalues are always real.

(Sketch of a) Proof: the eigenvalue equation is a polynomial equation (of degree n) then it has at least one complex root (the "Fundamental theorem of algebra"). This means that there is at least one eigenvector $|v_1\rangle$. Since A is Hermitian, then A maps the space orthogonal to $|v_1\rangle$ ($V_1^\perp$) into itself: if $(|v_1\rangle, |w\rangle) = 0$ then $(|v_1\rangle, A|w\rangle) = (A|v_1\rangle, |w\rangle) = \lambda_1(|v_1\rangle, |w\rangle) = 0$. Then A restricted to $V_1^\perp$ is just a $(n-1) \times (n-1)$ matrix and we can repeat the same steps recursively to find n eigenvalues and eigenvectors. Since the eigenvectors are linearly independent, they form a basis. Notice that the basis just constructed is an *orthogonal basis*. Of course we are free to rescale the eigenvectors as we want and (3.1) is always satisfied, thus we can make the eigenvector basis *orthonormal*. So by using an exercise given in *week 2*, we can state the completeness of the eigenvectors $|v_i\rangle$ of a Hermitian matrix in the following way:

$$\sum_i |v_i\rangle\langle v_i| = 1 \ , \tag{3.4}$$

Finally notice that the eigenvalues of A are real

$$\lambda_1||v_1||^2 = \langle v_1, |Av_1\rangle = \langle Av_1|, v_1\rangle = \bar{\lambda}_1||v_1||^2 \ . \tag{3.5}$$

Now the question is what happens if we deal with self-adjoint operators defined on an infinite dimensional Hilbert space. As you can see from the example below this nice theorem cannot hold exactly in the same form. We can consider a weaker version of (3.1): look for a bra $\langle v_\lambda|$ such that

$$\langle v_\lambda|A|w\rangle = \lambda\langle v_\lambda|w\rangle \ , \tag{3.6}$$

for all $|w\rangle$ in the domain of A (λ is real, as in (3.5)). As we have seen in the example 2.2.1 on the Dirac's delta not all bras satisfying (3.6) do have a corresponding ket. Thus for a self-adjoint operator we have two cases:

- the possible eigenvalues satisfying (3.1) form a discrete set;
- the "eigenvalues" satisfying (3.6), but not (3.1), form a a continuous set.

It turns out that, even if the bras with continuous eigenvalues do not have corresponding ket, their integral over a finite region of values of λ does. In particular if $c(\lambda)$ is a smooth function that is non-zero only in a finite region of the possible λ's, then the linear functional $\int d\lambda c(\lambda)\langle v_\lambda|$ has a corresponding vector that we will indicate with $\int d\lambda c(\lambda)|v_\lambda\rangle$. With an abuse of notation, physicists commonly use also the symbol $|v_\lambda\rangle$, even if there is no $|v_\lambda\rangle$ corresponding to the bra in (3.6)! The idea is that this object yields standard vectors when integrated.

At this point we can state the infinite dimensional analogue of the theorem [T1]. Consider a self-adjoint operator: the eigenvectors $|v_i\rangle$ corresponding to discrete eigenvalues together with those related to continues eigenvalues ($|v_\lambda\rangle$) form a complete set

$$\sum_i |v_i\rangle\langle v_i| + \int d\lambda|v_\lambda\rangle\langle v_\lambda| = 1 \ . \tag{3.7}$$

The orthonormality condition reads as follow

$$\langle v_i|v_j\rangle = \delta_{ij} \quad , \qquad \langle v_{\lambda_1}|v_{\lambda_2}\rangle = \delta(\lambda_1 - \lambda_2) \ . \tag{3.8}$$

3.2.1 Examples and exercises.

- Consider the momentum operator in quantum mechanics $-i\hbar\frac{d}{dx}$ that act on differentiable, square integrable (wave)functions. There is no solution to the eigenvector equation

  $$-i\hbar\frac{d\psi(x)}{dx} = \lambda\psi(x) \ . \tag{3.9}$$

  It is clear that the only possibility is to choose $\psi(x) = \mathrm{e}^{\frac{i\lambda x}{\hbar}}$, but this function is not square integrable regardless whether λ is real or imaginary.

4 The postulates of quantum mechanics

1 At a fixed time t_0, the state of a physical system is defined by a vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$.

2 Every physical (measurable) quantity $\mathcal{A}$ is described by a self-adjoint operator A (also called "observable").

3 The result of a measurement of the physical quantity $\mathcal{A}$ is always one the eigenvalues of the corresponding operator A.

4 The probability of finding the eigenvalue a in a measurement is $||P_a|\psi\rangle||^2$, where $|\psi\rangle$ has unit norm and P_a is the projector on the space of eigenvectors of eigenvalue a.

5 After a measurement of $\mathcal{A}$ yielding the value a (an eigenvalue of A), then the state of the system change from $|\psi\rangle$ to $P_a|\psi\rangle/||P_a|\psi\rangle||$.

6 The time evolution of the system is described

$$H|\psi(t)\rangle = i\hbar\frac{d}{dt}|\psi(t)\rangle \ , \tag{4.1}$$

where H is the observable associated to the energy of the system (Hamiltonian).

4.0.2 Examples and exercises.

Consider the operator defined on $\mathbf{C}^3$ and the ket $|\phi\rangle$:

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \ , \quad |\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \ . \tag{4.2}$$

The operator A has the following three eigenvectors:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \ , \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \ , \quad |\psi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \ . \tag{4.3}$$

with eigenvalues 2, -2 and 2 respectively. Thus the projectors on the two eigenspaces are

$$P_{(2)} = |\psi_2\rangle\langle\psi_2| + |\psi_3\rangle\langle\psi_3| \ , \quad P_{(-2)} = |\psi_1\rangle\langle\psi_1| \ . \tag{4.4}$$

If the state of our physical system is described by $|\phi\rangle$, then I can compute the probabilities of measuring ± 2 by decomposing $|\phi\rangle$ on the basis (4.3)

$$|\phi\rangle = P_{(2)}|\phi\rangle + P_{(-2)}|\phi\rangle = \sum_{i=1}^{3}\langle\psi_i|\phi\rangle|\psi_i\rangle \ . \tag{4.5}$$

The probability of finding -2 is $||P_{(-2)}|\phi\rangle||^2 = 1/4$, while that for 2 is $||P_{(2)}|\phi\rangle||^2 = 3/4$.

Exercise. Consider the Hilbert space $\mathbf{C}^2$ and the observable σ_1 in (3.3). If a quantum mechanical system is described by the state

$$|\psi\rangle = \begin{pmatrix} 1 \\ 2 \end{pmatrix} , \tag{4.6}$$

- What is the probability, in a physical measure, of finding as a result the first and the second eigenvalue?
- If the result of this measure is the positive eigenvalues, what are the possible results in a subsequent measure of the observable $\sin\theta\sigma_2 + \cos\theta\sigma_3$? What are the probabilities of finding each result?

5 Some simple quantum mechanical system.

We will focus on some simple quantum mechanical systems that have a time independent Hamiltonian. In this case it is simple to describe in general how a vector $|\psi(t_0)\rangle$, representing the system at the time t_0, evolves with time. If $\hat{H}$ is time independent, one can check that the following state

$$|\psi(t)\rangle = \mathrm{e}^{-\frac{i\hat{H}}{\hbar}(t-t_0)}|\psi(t_0)\rangle \equiv \sum_{n=0}^{\infty} \frac{1}{n!}\left[\frac{-i\hat{H}}{\hbar}(t-t_0)\right]^n |\psi(t_0)\rangle \tag{5.1}$$

solves the time evolution equation of the postulate 6. In order to compute explicitly $|\psi(t)\rangle$ it is clearly convenient to decompose $|\psi(t_0)\rangle$ along the complete basis of the Hamiltonian eigenvector. So one of the tools we need is the set of the solutions of the "time independent Schroedinger equation" $\hat{H}|\psi_E\rangle = E|\psi_E\rangle$. Often this problem can be simplified by exploiting the following observation.

Two observables $\hat{A}$ and $\hat{B}$ that commute ($[\hat{A},\hat{B}] = 0$) have a common set eigenspaces. This means that we can find projectors $P_{(a,b)}$ that project at the same time on the subspace of eigenvalue a for the first operator $\hat{A}$ and the the subspace of eigenvalue b for the second operator $\hat{B}$. You can convince yourself that this is reasonable, by looking at the simple case of finite dimensional Hilbert spaces: in this case if two Hermitian matrices commute, they have a common set of eigenvectors. Sketch of a proof: Suppose that the $|v_a\rangle$ is the only eigenvector of eigenvalue a of the Hermitian matrix A. If $[A,B] = 0$, it is easy to see that also $B|v_a\rangle$ is an eigenvector of A with eigenvalue a: $A(B|v_a\rangle) = BA|v_a\rangle = a(B|v_a\rangle)$. This means that $B|v_a\rangle$ must be proportional to $|v_a\rangle$, in formulae: $B|v_a\rangle = b|v_a\rangle$, which implies that $|v_a\rangle$ is also an eigenvector of B.