Notes on Quantum Kinematics

J. M. B. Lopes dos Santos*

September 15, 2021

Departamento de Física e Astronomia,
Faculdade de Ciências da Universidade do Porto
4169-007 Porto

A brief review of core concepts on the description of a quantum system

1 Introduction

These notes are intended as a review of fundamental concepts in Quantum Mechanics (QM), as a basis for some extensions of previously acquired knowledge. I will not dwell on History, or give many details to motivate concepts. I will assume some familiarity with Quantum Mechanics and strive to bring out a core of important concepts. Our focus will be on quantum kinematics, *i.e.,* on the mode of description of a quantum system. Many important discoveries of quantum theories succeeded the laying out of the foundations in the second half of the 1920 decade, but the concepts we are about to discuss stand as they were, without change, from these early beginnings. In this review, I will first address the mathematical structure separated from the physical content of the theory, a route I would no advocate in introductory courses.

2 Vector Spaces

2.1 Fundamentals

The fundamental idea of quantum mechanics is that the space of states of a physical system is a vector space, with complex scalars (in the field of complex numbers, using mathematicians nomenclature). The abstract notion of vector space is therefore crucial to understanding the mathematical structure of QM. A state will be denoted, according do Dirac notation,

$$|\psi\rangle \tag{1}$$

*jlsantos@fc.up.pt

and refereed to as a *ket.* What is it exactly? A vector in 3D space? A function? A matrix?

In the context of abstract vector spaces, we do not care about the nature of these objects, only on what we can do with them. Hence, if $|\psi\rangle$ is a vector, it it is an element of a set $\mathfrak{H} = \{|\xi\rangle, |\eta\rangle, \dots\}$ such that

$$\lambda |\psi\rangle \in \mathfrak{H}, \qquad \text{if } |\psi\rangle \in \mathfrak{H} \text{ and } \lambda \in \mathbb{C} \tag{2}$$

($\mathbb{C}$ is the field of complex numbers) and

$$|\psi\rangle + |\phi\rangle \in \mathfrak{H} \qquad \text{if } |\psi\rangle, |\phi\rangle \in \mathfrak{H} \tag{3}$$

In addition, there are some axioms that we specify for these operations—the existence of a null vector, commutativity of addition, associativity of multiplication by a scalar, etc.—, that have by now become so ingrained that I will dispense from stating them explicitly.

The notion that a state is an element of a vector space, immediately leads to a distinctive property of the description quantum systems

> ***Principle of superposition***: *If $|\psi\rangle$ and $|\phi\rangle$ are possible states of a physical system (whatever that means) so is*
>
> $$\lambda |\psi\rangle + \eta |\phi\rangle, \qquad \lambda, \eta \in \mathbb{C} \tag{4}$$

In classical physics no such concept is possible. A system of N particles, to be specific, has a state given by a point in phase space ($6N$ dimensional space); there is no meaning that can be attributed to a linear combination of two such points. The only situation involving more than one point of phase space has to do with uncertainty of which state the system is in. One can, for instance, specify that uncertainty by defining a probability density in phase space. But in QM, $a|\psi\rangle + b|\phi\rangle$ is a state with the same standing as $|\psi\rangle$ or $|\phi\rangle$, *i.e.,* another element of $\mathfrak{H}$.

We now proceed to state some results for vector spaces, which should be familiar to all. A good reference to review this matter is Tom Apostol's Calculus Vol II chapter 1 [1].

Linear Independence A set of vectors $\{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_r\rangle\}$is said to be linearly *independent* if

$$\sum_{i=1}^{r} \lambda_i |\phi_i\rangle = 0 \tag{5}$$

implies $\lambda_i = 0$ for $i = 1, \dots, r$.

Space spanned by set of vectors Let $\mathcal{B} = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_r\rangle\}$ be a set of linearly independent vectors. Any vector of the form

$$\sum_{i=1}^{r} \lambda_i |\phi_i\rangle, \qquad \lambda_i \in \mathbb{C} \tag{6}$$

is said to be spanned by $\mathcal{B}$. The set of vectors spanned by $\mathcal{B}$ is itself a vector space. It is easily seen that if $|\psi\rangle$ and $|\phi\rangle$ are spanned by $\mathcal{B}$ so is any linear combination of $|\psi\rangle$ and $|\phi\rangle$.

Basis A vector space $\mathfrak{H}$ spanned by d linear independent vectors, $\mathcal{B} = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_d\rangle\}$ is a finite dimensional space, of dimension d and $\mathcal{B}$ is a *basis* of $\mathfrak{H}$. An important results, is that all basis of $\mathfrak{H}$ have the same number of elements, otherwise we could not define their number to be a property of the space. This can be shown using[1]

> **Theorem:** If $\mathcal{B} = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_r\rangle\}$ is a set of linearly independent vectors, any $r+1$ vectors in the space spanned by $\mathcal{B}$, $L(\mathcal{B})$ are linearly dependent.

If follows from the concept of linear independence that two different linear combinations are two different vectors. If the $\{|\phi_i\rangle : i = 1, 2, \ldots\}$ are a linearly independent set and

$$|\psi\rangle = \sum_{i=1}^{r} \lambda_i |\phi_i\rangle \tag{7}$$

$$|\psi\rangle = \sum_{i=1}^{r} \eta_i |\phi_i\rangle \tag{8}$$

we have

$$0 = \sum_{i=1}^{r} (\lambda_i - \eta_i) |\phi_i\rangle \tag{9}$$

and linear independence implies

$$\lambda_i = \eta_i \qquad \forall_i \tag{10}$$

This implies that there a unique way to express any vector in a given basis

$$|\psi\rangle = \sum_{i=1}^{d} \lambda_i |\phi_i\rangle \tag{11}$$

So given a basis, there is a a one-to-one correspondence between kets—elements of $\mathfrak{H}$—, and column vectors with d complex entries

$$|\psi\rangle \longleftrightarrow [\lambda_1, \ldots, \lambda_d]^T \tag{12}$$

We say that a basis defines a *representation.* Before proceeding, let me call your attention to the possibility of a vector space having no *finite* basis. Mathematically, infinite dimensional spaces are much trickier. One has to admit linear combinations of infinite sets of vectors and, of course, that can only be understood in terms of suitable limiting procedures. We will plod on describing results for finite dimensional spaces, aware, of course, that that is not the whole story. Luckily in many applications in QM suitable spaces $\mathfrak{H}$ are finite dimensional, and also most of the concepts we are about to consider carry over to infinite d.

2.2 Dual space

A complex functional in $\mathfrak{H}$ is a linear map from $\mathfrak{H}$ to $\mathbb{C}$.

$$\begin{aligned} \mathcal{L} : \mathfrak{H} &\to \mathbb{C} \\ |\psi\rangle &\to \mathcal{L}\left[|\psi\rangle\right] \end{aligned} \tag{13}$$

Linearity means, of course,

$$\mathcal{L}\left[\lambda\left|\psi\right\rangle + \eta\left|\phi\right\rangle\right] = \lambda\mathcal{L}\left[|\psi\rangle\right] + \eta\mathcal{L}\left[|\psi\rangle\right] \tag{14}$$

Note that on the LHS addition is in $\mathfrak{H}$ and on the RHS in $\mathbb{C}$. The space of linear functionals on $\mathfrak{H}$ is also a vector space of the same dimension as $\mathfrak{H}$ called the dual $\mathfrak{H}^*$. We merely define $\mathcal{L} = \lambda\mathcal{L}_1 + \eta\mathcal{L}_2$ by

$$\mathcal{L}\left[|\psi\rangle\right] = \lambda\mathcal{L}_1\left[|\psi\rangle\right] + \eta\mathcal{L}_2\left[|\psi\rangle\right] \tag{15}$$

The null vector in $\mathfrak{H}^*$ is the functional

$$\mathcal{L}\left[|\psi\rangle\right] = 0, \qquad \forall : |\psi\rangle \in \mathfrak{H} \tag{16}$$

and given a basis $\mathcal{B} = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_d\rangle\}$ of $\mathfrak{H}$ we define, $\mathcal{B}^* = \{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_d\}$ by

$$\mathcal{L}_i\left[|\phi_j\rangle\right] := \delta_{ij} \tag{17}$$

easily shown to be basis of $\mathfrak{H}^*$, called the *dual basis* of $\mathcal{B}$

Exercise 1.
Prove this last statement as follows:

1. Show that any functional is uniquely defined by its action on a basis of $\mathfrak{H}$

$$\mathcal{L}\left[|\phi_j\rangle\right] = \lambda_j \tag{18}$$

2. express $\mathcal{L}$ as linear combination of the $\mathcal{L}_i$

2.3 Scalar product

The vector space structure contains no notion of distance or angles. We have yet no way of calculation the "size" of a vector or determine if two vectors are close, or orthogonal because we do not know how to calculate the "size" of their difference, or the "angle" between them. In other words we have no metric properties. QM mechanics requires these concepts.

To endow a vector space with metrical properties we require an operation that maps two vectors to a complex number—a scalar product—,

$$\begin{aligned} \mathfrak{H} \otimes \mathfrak{H} &\to \mathcal{C} \\ |\psi\rangle, |\phi\rangle &\to (|\psi\rangle, |\phi\rangle) \in \mathcal{C} \end{aligned} \tag{19}$$

with the following properties:

1. $(|\psi\rangle, |\psi\rangle) \geq 0$
2. $(|\psi\rangle, |\psi\rangle) = 0$ implies that $|\psi\rangle$ is the null vector
3. $(|\psi\rangle, |\phi\rangle) = (|\phi\rangle, |\psi\rangle)^*$
4. Linearity in the rightmost ket

   $$(|\psi\rangle, \lambda |\phi_1\rangle + \eta |\phi_2\rangle) = \lambda (|\psi\rangle, |\phi_1\rangle) + \eta (|\psi\rangle, |\phi_2\rangle) \tag{20}$$

   which, by 3, implies anti-linearity in the leftmost ket

   $$(\lambda |\phi_1\rangle + \eta |\phi_2\rangle, |\psi\rangle) = \lambda^* (|\phi_1\rangle, |\psi\rangle) + \eta^* (|\phi_2\rangle, |\psi\rangle) \tag{21}$$

The Dirac notation for this scalar product, which we will use henceforth, is

$$(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle \tag{22}$$

The "length" or norm[1] of a ket is

$$\||\psi\rangle\|^2 = \langle\psi|\psi\rangle \tag{23}$$

and by definition is non-negative. The distance between two kets is the norm of their difference. Angles can also be defined using an important result, the Schwarz inequality

$$|\langle\psi|\phi\rangle|^2 \leq \langle\phi|\phi\rangle\langle\psi|\psi\rangle$$

We *define* the angles between two vectors by

$$|\langle\psi|\phi\rangle|^2 = \langle\phi|\phi\rangle\langle\psi|\psi\rangle \cos^2\theta \tag{24}$$

since Schwarz inequality implies that $\cos^2\theta \leq 1$. Kets with zero scalar product are called orthogonal.

[1] In mathematics textbooks the norm is the square root of the $\langle\psi|\psi\rangle$

If our basis is orthonormal (orthogonal vectors of unit norm)

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \tag{25}$$

we can express the components of a ket by scalar products

$$|\psi\rangle = \sum_i \lambda_i |\phi_i\rangle \tag{26}$$

$$\langle\phi_j|\psi\rangle = \sum_i \lambda_i \langle\phi_j|\phi_i\rangle = \sum_i \lambda_i \delta_{ji} = \lambda_j \tag{27}$$

Exercise 2.
You can prove Schwarz inequality by computing the norm of

$$|\psi\rangle - \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle} |\phi\rangle \tag{28}$$

and using the fact that it must be non-negative. You can also show that the equality only occurs if $|\phi\rangle$ and $|\psi\rangle$ are linearly dependent. Note that this vector is the component of $|\psi\rangle$ orthogonal to $|\phi\rangle$.

2.4 Bra Space

The existence of a scalar product induces a natural bijective correspondence between vectors in $\mathfrak{H}$ (kets) and in $\mathfrak{H}^*$ (bras). Natural, in this context, means basis independent. For any ket $|\phi\rangle \in \mathfrak{H}$ we define a bra, a functional on $\mathfrak{H}$, by

$$\begin{gathered} |\phi\rangle \to \mathcal{L}_\phi \\ \mathcal{L}_\phi\left[|\psi\rangle\right] = \langle\phi|\psi\rangle \end{gathered} \tag{29}$$

This is clearly a linear functional, since the scalar product is linear in the rightmost ket. We now make one new step in Dirac notation by denoting $\mathcal{L}_\phi$ by $\langle\phi|$ and calling it a *bra,* an element of $\mathfrak{H}^*$. So the ket–bra correspondence is indicated by using the same letter inside each half–bracket. We can now read the symbol $\langle\phi|\psi\rangle$ in two ways:

$$\langle\phi|\psi\rangle = \mathcal{L}_\phi\left[|\psi\rangle\right]; \qquad \text{the bra that corresponds to } |\phi\rangle \text{ with argument } |\psi\rangle \tag{30}$$

$$\langle\phi|\psi\rangle = \left(|\phi\rangle\,, |\psi\rangle\right); \qquad \text{the scalar product of kets } |\phi\rangle \text{ by } |\psi\rangle \tag{31}$$

The relation between $|\phi\rangle$ and $\langle\phi|$ is *hermitian conjugation*

$$\langle\phi| = \left(|\phi\rangle\right)^\dagger \tag{32}$$

The reason is simple. Because the scalar product is anti-linear in the leftmost ket,

$$\left(\lambda\,|\phi_1\rangle + \eta\,|\phi_2\rangle\,,\right)|\psi\rangle = \lambda^*\langle\phi_1|\psi\rangle + \eta^*\langle\phi_2|\psi\rangle, \tag{33}$$

which means

$$\left(\lambda \left|\phi_1\right\rangle + \eta \left|\phi_2\right\rangle\right)^\dagger = \lambda^* \left\langle\phi_1\right| + \eta^* \left\langle\phi_2\right| . \tag{34}$$

If we use an basis in $\mathfrak{H}$, $\mathcal{B} = \{\left|\phi_1\right\rangle, \left|\phi_2\right\rangle, \ldots, \left|\phi_d\right\rangle\}$ and

$$\left|\psi\right\rangle = \sum_i \lambda_i \left|\phi_i\right\rangle \tag{35}$$

$$\left\langle\psi\right| = \sum_i \lambda_i^* \left\langle\phi_i\right| . \tag{36}$$

We naturally represent bras by row vectors. When we use dual basis in each space,

$$\left|\psi\right\rangle \rightarrow [\lambda_1, \ldots, \lambda_d]^T \tag{37}$$

$$\left\langle\psi\right| \rightarrow [\lambda_1^*, \ldots, \lambda_d^*] \tag{38}$$

and the scalar product in a matrix product of components if $\mathcal{B} = \{\left|\phi_1\right\rangle, \left|\phi_2\right\rangle, \ldots, \left|\phi_d\right\rangle\}$ is orthonormal.[2]

$$\langle\phi|\psi\rangle = [\eta_1^*, \ldots, \eta_d^*] \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_d \end{bmatrix} = \sum_i \eta_i^* \lambda_i \tag{39}$$

The economy of Dirac's notation is starting to show. To compute a scalar product, a braket, just write a bra (left half of a braket) followed by a ket (right half).

2.5 Operators

Linear Transformations between $\mathfrak{h}$ and $\mathfrak{U}$ preserve linear combinations

$$\begin{aligned} M : \mathfrak{h} &\rightarrow \mathfrak{U} \\ M\left(\lambda u + \eta v\right) &= \lambda M(u) + \eta M(v) \end{aligned} \tag{40}$$

In QM almost always

$$\begin{aligned} M : \mathfrak{h} &\rightarrow \mathfrak{h} \\ M\left(\lambda \left|\phi\right\rangle + \eta \left|\psi\right\rangle\right) &= \lambda M\left(\left|\phi\right\rangle\right) + + \eta M\left(\left|\psi\right\rangle\right) \end{aligned} \tag{41}$$

M is an *linear operator,*

$$\left|\psi\right\rangle \rightarrow \hat{M} \left|\psi\right\rangle \tag{42}$$

This notation is very economical for *composition of transformations*

$$\begin{aligned} M : \mathfrak{h} &\rightarrow \mathfrak{h} \\ N : \mathfrak{h} &\rightarrow \mathfrak{h} \\ N \circ M : \mathfrak{h} &\rightarrow \mathfrak{h} \\ N\left(M\left(\left|\xi\right\rangle\right)\right) &= \hat{N}\left(\hat{M}\left|\xi\right\rangle\right) = \hat{N}\hat{M}\left|\xi\right\rangle \end{aligned} \tag{43}$$

[2] Note that if $\mathcal{B} = \{\left|\phi_1\right\rangle, \left|\phi_2\right\rangle, \ldots, \left|\phi_d\right\rangle\}$ is orthonormal, its dual is $\mathcal{B}^* = \{\left\langle\phi_1\right|, \left\langle\phi_2\right|, \ldots, \left\langle\phi_d\right|\}$

which defines the *product* of operators as the composition of its transformations. Composition of maps is obviously non-commutative and so is the product of operators.

Here is another example of Dirac notation: the symbol $|\psi\rangle\langle\phi|$ is linear operator defined by

$$\hat{M} = |\psi\rangle\langle\phi| \tag{44}$$

$$\hat{M}|\xi\rangle = |\psi\rangle\langle\phi|\xi\rangle. \tag{45}$$

Linearity of $\hat{M}$ follows from linearity of scalar product.

Let a a linear functional $\langle\phi|$ act on $\hat{M}|\xi\rangle$

$$\langle\phi|\left(\hat{M}|\xi\rangle\right) \tag{46}$$

This is a linear functional on $\mathfrak{h}$, mapping

$$|\xi\rangle \to \langle\phi|\left(\hat{M}|\xi\rangle\right) \tag{47}$$

So this a a bra (element of $\mathfrak{h}^*$) defined, using the power of Dirac notation, simply as

$$\langle\psi| = \langle\phi|\hat{M} \tag{48}$$

Hence two equivalent views on the same symbol

$$\langle\phi|\hat{M}|\xi\rangle, \qquad \text{the bra } \langle\phi| \text{ acting on } \hat{M}|\xi\rangle \tag{49}$$

$$\langle\phi|\hat{M}|\xi\rangle, \qquad \text{the bra } \langle\phi|\hat{M} \text{ acting on } |\xi\rangle \tag{50}$$

2.5.1 Projectors

$$\hat{P}_\phi = |\phi\rangle\langle\phi| \tag{51}$$

If $|\phi\rangle$ is normalized

$$\hat{P}_\phi|\xi\rangle = |\phi\rangle\langle\phi|\xi\rangle \tag{52}$$

is the $|\phi\rangle$ component of $|\xi\rangle$ in an basis of which $|\phi\rangle$ is one of the states. $\hat{P}_\phi$ is the *projector* on $|\phi\rangle$.

$$\begin{aligned} |\xi\rangle &= \sum_i \lambda_i |\phi_i\rangle \\ &= \sum_i |\phi_i\rangle\langle\phi_i|\xi\rangle \end{aligned} \tag{53}$$

leads to

$$\sum_i |\phi_i\rangle\langle\phi_i| = \hat{1} \qquad \text{(identity operator)} \tag{54}$$

This is the *identity resolution* in base $\{|\phi_i\rangle : i = 1, 2, \dots\}$.

2.5.2 Operators and square matrices

Typical Dirac notation manipulation

$$\begin{aligned}\hat{M}|\xi\rangle &= \hat{1}\hat{M}\hat{1}|\xi\rangle \\ &\left(\sum_i |\phi_i\rangle\langle\phi_i|\right)\hat{M}\left(\sum_j |\phi_j\rangle\langle\phi_j|\right)|\xi\rangle \\ &\sum_i\sum_j |\phi_i\rangle\langle\phi_i|\hat{M}|\phi_j\rangle\langle\phi_j|\xi\rangle \end{aligned} \tag{55}$$

Representation

$$|\xi\rangle \to \lambda_j = \langle\phi_j|\xi\rangle \tag{56}$$

$$|\psi\rangle = \hat{M}|\xi\rangle \to \eta_j := \langle\phi_j|\psi\rangle \tag{57}$$

$$\hat{M} \to M_{ij} = \langle\phi_i|\hat{M}|\phi_j\rangle \tag{58}$$

so

$$|\psi\rangle = \hat{M}|\xi\rangle \to \eta_i = \sum_j M_{ij}\lambda_j \tag{59}$$

$$\begin{bmatrix}\eta_1\\ \vdots \\ \eta_d\end{bmatrix} = \begin{bmatrix} M_{11} & \dots & M_{1d} \\ \vdots & \ddots & \vdots \\ M_{d1} & \dots & M_{dd}\end{bmatrix}\begin{bmatrix}\lambda_1\\ \vdots \\ \lambda_d\end{bmatrix} \tag{60}$$

Exercise 3.
Use Dirac notation and the resolution of the identity to show that, in given representation, the matrix of the product $\hat{N}\hat{M}$ is the usual matrix product of the matrices of N and M

$$(NM)_{ij} = \sum_k N_{ik}M_{kj} \tag{61}$$

2.5.3 Operators from a representation

Expressing operator from its representation

$$\begin{aligned}\hat{M} = \hat{1}\hat{M}\hat{1} &= \left(\sum_i |\phi_i\rangle\langle\phi_i|\right)\hat{M}\left(\sum_j |\phi_j\rangle\langle\phi_j|\right) \\ &= \sum_{i,j}|\phi_i\rangle\langle\phi_i|\hat{M}|\phi_j\rangle\langle\phi_j| = \sum_{i,j}|\phi_i\rangle M_{ij}\langle\phi_j| \end{aligned} \tag{62}$$

This expression captures the two "lives" of $\hat{M}$ as a operator in ket or in bra space; put a ket at its right and get a ket; put it after a bra and get a new bra.

Exercise 4.
Use Dirac notation to show that $\langle\psi|\,\hat{M}$, in a basis where $\langle\psi| = [\lambda_1, \ldots, \lambda_d]$, is the row matrix $[\eta_1, \ldots, \eta_d]$,

$$[\eta_1, \ldots, \eta_d] = [\lambda_1, \ldots, \lambda_d] \begin{bmatrix} M_{11} & \ldots & M_{1d} \\ \vdots & \ddots & \vdots \\ M_{d1} & \ldots & M_{dd} \end{bmatrix} \tag{63}$$

with the usual convention of matrix products.

2.5.4 Hermitian conjugation

Of kets and bras

$$(|\xi\rangle)^\dagger = \langle\xi| \tag{64}$$

$$(\langle\xi|)^\dagger = |\xi\rangle \tag{65}$$

We saw above, properties of scalar product imply,

$$(\lambda\,|\xi\rangle + \eta\,|\psi\rangle)^\dagger = \lambda^*\,\langle\xi| + \eta^*\,\langle\psi| \tag{66}$$

$$(\lambda\,\langle\xi| + \eta\,\langle\psi|)^\dagger = \lambda^*\,|\xi\rangle + \eta^*\,|\psi\rangle \tag{67}$$

Of operators

$$\left(\hat{M}\,|\xi\rangle\right)^\dagger = \langle\xi|\,\hat{M}^\dagger \tag{68}$$

$$\left(\langle\xi|\,\hat{M}\right)^\dagger = \hat{M}^\dagger\,|\xi\rangle \tag{69}$$

These definitions imply

$$\langle\psi|\,\hat{M}\,|\phi\rangle^* = \left(\langle\phi|\,\hat{M}^\dagger\right)|\psi\rangle = \langle\phi|\,\hat{M}^\dagger\,|\psi\rangle \tag{70}$$

Exercise 5.
Prove that

$$\left(\hat{N}\hat{M}\right)^\dagger = \hat{M}^\dagger\hat{N}^\dagger \tag{71}$$

Exercise 6.
Show that the above definition of $\hat{M}^\dagger$, which is independent of any representation, implies that its matrix, in any given representation (basis) is the matrix of $\hat{M}$ with its elements transposed and conjugated

$$M^\dagger_{ij} = M^*_{ji} \tag{72}$$

2.6 Eigenvalues and Eigenvectors

Definition

$$\hat{M} |\psi\rangle = m |\psi\rangle \tag{73}$$

- $|\psi\rangle$ is an *eigenvector* of $\hat{M}$
- m is the corresponding *eigenvalue*

An eigenvector defines an invariant "direction" or ray in Hilbert space[3], under the map defined by $\hat{M}$. The eigenvalue $m = |m| e^{i\theta}$ defines a scale factor and a phase change. Properties of eigenvalues of *Hermitian* operators are especially important. An Hermitian operator is one that is equal to its hermitian conjugate

$$\hat{M}^\dagger = \hat{M} \tag{74}$$

a) Eigenvalues of hermitian operators are real;

b) Eigenvectors of different eigenvalues are orthogonal (not true in general for non-hermitian operators):

c) In the spaces of interest in QM, the Hermitian operators are complete, in the sense that one can define a basis for the space with its eigenvectors. Given an hermitian operator $\hat{M}$ one can choose a basis for the space $\mathfrak{h}$ $\mathcal{B} = \{|\phi_1\rangle, \ldots, |\phi_d\rangle\}$ where

$$\hat{M} |\phi_i\rangle = m_i |\phi_i\rangle \tag{75}$$

The $m_i's$ are not necessarily all distinct. One can always choose the basis to be orthonormal and in that case the matrix representing $\hat{M}$ is diagonal

$$\begin{aligned}
\hat{M} &= \sum_{i,j} |\phi_i\rangle \langle\phi_i| \hat{M} |\phi_j\rangle \langle\phi_j| \\
&= \sum_{i,j} |\phi_i\rangle m_j \langle\phi_i|\phi_j\rangle \langle\phi_j| \\
&= \sum_i m_i |\phi_i\rangle \langle\phi_i|
\end{aligned} \tag{76}$$

Exercise 7.

Prove statements **a)** and **b)** concerning the eigenvalues of Hermitian operators

---

[3] A *ray* is the one dimensional subspace $\{m |\psi\rangle : m \in \mathcal{C}\}$

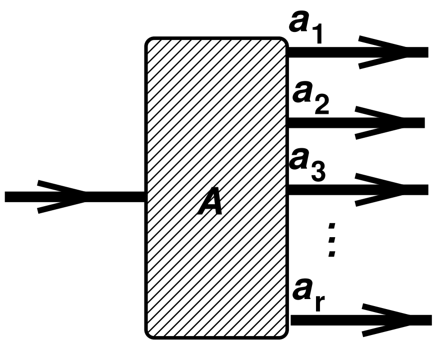

Figure 1: A filter to select systems with a definite value of quantity A.

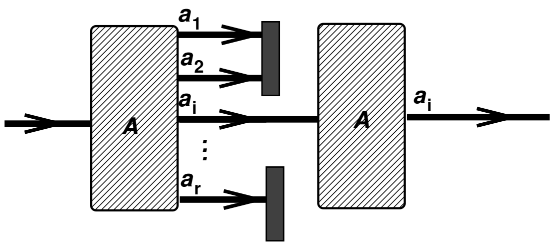

Figure 2: Systems emerging form a given output channel fo the A-filter, have a predictable value of A.

3 The Physical Content of the theory

3.1 Filtering

Repeated experiments on identical physical systems; the metaphor of a beam. Assume you can select—filter— systems according a physical property called A. For simplicity let A take r finite possible values. Represent measurement of A as a device with one input channel and r output ones [Fig. (1)].

We will ignore dynamics and assume the value of A can be checked to be a definite property of each output, in the sense that a repetition of filter A on a single beam out of the first filter, leaves only one output channel occupied [Fig. (2)].

A single physical quantity does not completely specify a state. We assume we can juxtapose filter for other observables. So we select one value of A and investigate the value of B [Fig. (3)]. Did we loose the information on B? We can check by repeating measurement of A. Two possibilities arise:

a) It is possible to arrange a device such that the information on the value of A is retained, *no matter what value of A one has.* This is illustrated in Fig. 4 , and A and B are called compatible.

b) No matter how we proceed, we find that after measuring, filtering, B we are unable to predict with certainty the value of A. In that case, *for at least one* output channel of the first A filter and *one output* channel of the B filter, at least two output channels of the second A filter are populated. Fixing the value of B leads

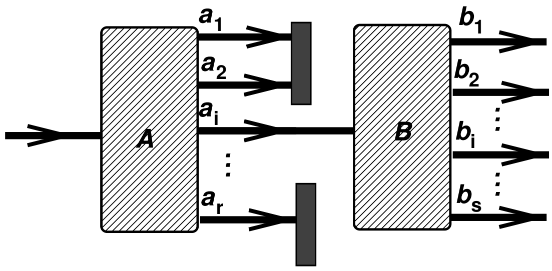

Figure 3: Measuring B after A. Is the information on A lost?

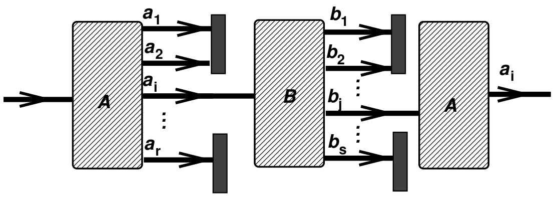

Figure 4: Measuring B after A. If A and B are compatible it is possible to build a device where we keep information on A upon measuring B

to loss of information on A; A and B are *incompatible.*

3.2 Complete sets of compatible observables in classical and quantum theory

For the moment ignore the possibility of incompatibility, and choose a second observable which is compatible with A. We can join the A and B filters and to an AB filter which selects according to values of A *and* of B. We now have a device with output channels (ij) characterized by a_i, b_j, the values of A and B [Fig. (5)].

We can keep doing this, finding variables that are compatible with previous ones, improving the specification of system variables until ... we reach one of two possibilities, which, in this language, distinguish a classical and quantum theories:

Classical: Adding another quantity Z to the already one used one yields no new infor-

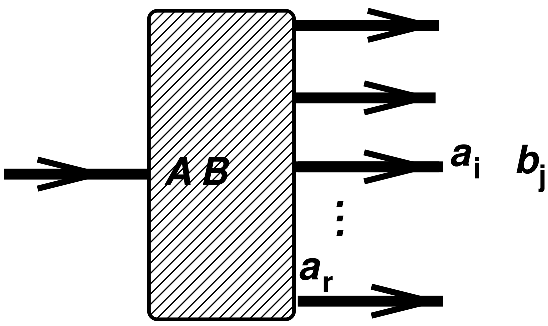

Figure 5: A filter to select systems with definite values of quantity A *and* B.

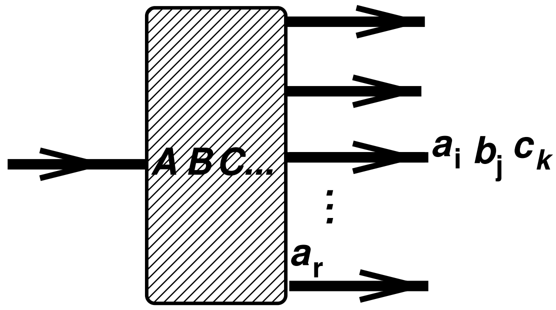

Figure 6: A filter to select systems with definite values of quantity *A and B*.

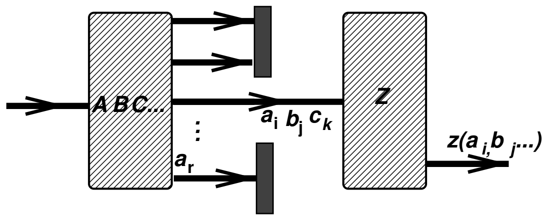

Figure 7: A filter to select systems with definite values of quantity *A and B*.

mation in the sense its filter with one input from the "complete" one will yield only one output channel. This means that the values of the set of complete variables, $a_i, b_j, c_k, \dots$ determine the value of all others, $z = F(a_i, b_j, c_k, \dots)$ [Fig.(7)]. This is the situation in classical physics. All observables are functions defined in phase space

$$A = f_A\,(q_1, p_1; q_2, p_2, \dots, q_N, p_N) \tag{77}$$

The set $\{q_1, p_1; q_2, p_2, \dots, q_N, p_N\}$ is a complete specification of the system state. Once we know (filter) a state with definite values of these variables we can predict with certainty the value of any physical observable of the system.

But this is only one possibility. There is another.

Quantum: Even though there no more observables compatible with the previous one—other than those that are functions of them—, there are still physical quantities whose values remain uncertain; observables incompatible with the complete set. We cannot add to the specification of the state, without destroying information on the values of $A, B, C, \dots$; in that sense the set is complete. But fixing its values still leaVes physical properties undefined. This how reality behaves. It is quantum, not classical, and quantum theories have to deal with this feature [Fig. (8)].

3.3 Quantum states and Statistical Postulate

Now we can understand what we mean by a quantum state:

A **Quantum State** of a physical system is a specification of a Complete Set of Compatible Observables (CSCO).

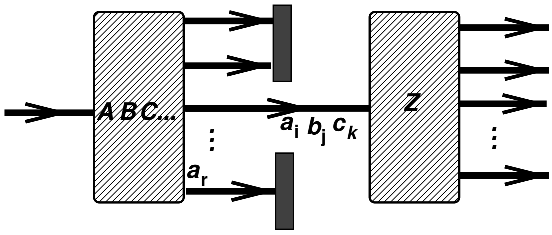

Figure 8: A filter to select systems with definite values of quantity A *and* B.

Let us use vector notation to specify all the values of a CSCO

$$a_i, b_j, c_k, \cdots \to \mathbf{a}_l \tag{78}$$

With a combined filter of a CSCO, physical systems with different values, $\mathbf{a} \neq \mathbf{b}$ of the same CSCO can be distinguished, in the sense that, with probability one, they are in different output channels. Such states are called orthogonal.

But, as stated before, incompatibility of observables means that we can build more than one CSCO, say $CSCO_1$ and $CSCO_2$, where the variables in the first are incompatible with those of the second. Let us say we denote $\mathbf{a}_i$ the values of $CSCO_1$ and $\mathbf{z}_r$ the values of $CSCO_2$. How is a state $|\mathbf{z}_r\rangle$ related to the $|\mathbf{a}_i\rangle$?

It is time to state state the basic structure of any quantum theory:

1. **Space of States** The states $|\mathbf{a}_i\rangle$ with definite values of *any* CSCO form a basis for a single linear vector space of states $\mathfrak{h}$ with complex scalars.

2. **Scalar Product** This space has a scalar product and we may take each of these basis to be orthonormal
$$\langle \mathbf{a}_i | \mathbf{a}_j \rangle = \delta_{ij} \tag{79}$$

3. **Principle of superposition**: Any quantum state, *i.e.,* any state which corresponds to a complete specification of a CSCO, $|\mathbf{z}\rangle$, is a linear combination of any of the basis, of other CSCO
$$|\mathbf{z}\rangle = \sum_i \lambda_i \, |\mathbf{a}_i\rangle \tag{80}$$

4. **Statistical Postulate:** a physical system prepared in state $|\psi\rangle$, given as a linear superposition of the basis of some CSCO,
$$|\psi\rangle = \sum_i \lambda_i \, |\mathbf{a}_i\rangle \tag{81}$$
has a probability $p(\mathbf{a}_i)$
$$p(\mathbf{a}_i) = |\lambda_i|^2 = |\langle \mathbf{a}_i | \psi \rangle|^2 \tag{82}$$
of yielding the value $\mathbf{a}_i$ upon being measured for the values of this CSCO. Naturally we are assuming
$$\langle \psi | \psi \rangle = \sum_i \lambda_i^* \lambda_j \langle \mathbf{a}_i | \mathbf{a}_j \rangle = \sum_i |\lambda_i|^2 = 1$$

since the $\mathbf{a}_i$ exhaust all possible values of the measurement. Note that we can relax this condition and consider physical systems defined by rays as long as we rewrite the statistical postulate as

$$p(\mathbf{a}_i) = \frac{|\lambda_i|^2}{\langle\psi|\psi\rangle} = \frac{|\langle\mathbf{a}_i|\psi\rangle|^2}{\langle\psi|\psi\rangle} \tag{83}$$

3.4 Observables as operators. Compatibility and commutation

Let A be a physical property belonging to some CSCO $A, B, C, \dots$. Denote by $\{\boldsymbol{\xi}_i : i = 1, 2, \dots\}$ the vectors of values of this CSCO. Assume the value of A is $\alpha(\boldsymbol{\xi}_i)$. We can associate with A a linear operator $\hat{A}$ defined by

$$\hat{A} = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| \tag{84}$$

which is another way of saying that the basis of this CSCO are eigenvectors of $\hat{A}$, and the eigenvalues are the values of the observable

$$\hat{A}\, |\boldsymbol{\xi}_i\rangle = \alpha(\boldsymbol{\xi}_i)\, |\boldsymbol{\xi}_i\rangle \tag{85}$$

If a is an eigenvalue of $\hat{A}$, its degeneracy is just the number of states that have

$$a = \alpha(\boldsymbol{\xi}_i). \tag{86}$$

Observables which are compatible, and therefore can be members of the same CSCO are represented by commuting operators because, by definition, they have a common eigenvector basis

$$\hat{A} = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| \tag{87}$$

$$\hat{B} = \sum_i |\boldsymbol{\xi}_i\rangle\, \beta(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| \tag{88}$$

$$\hat{A}\hat{B} = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\beta(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| = \hat{B}\hat{A} \tag{89}$$

these operators are hermitian because we choose $\alpha(\boldsymbol{\xi}_i)$ to be real.

$$\hat{A}^\dagger = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha^*(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| = \hat{A} \tag{90}$$

Observables which are incompatible— also called, after Bohr, *complementary*—, cannot have a common basis. They may have common eigenstates, but not a common *basis*; their operators cannot commute, because, otherwise they would have a common basis.

$$\hat{A} = \sum_i |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\, \langle\boldsymbol{\xi}_i| \tag{91}$$

$$\hat{Z} = \sum_i |\boldsymbol{\eta}_i\rangle\, \zeta(\boldsymbol{\eta}_i)\, \langle\boldsymbol{\eta}_i| \tag{92}$$

$$\hat{A}\hat{Z} = \sum_{i,j} |\boldsymbol{\xi}_i\rangle\, \alpha(\boldsymbol{\xi}_i)\zeta(\boldsymbol{\eta}_j)\langle\boldsymbol{\xi}_i|\boldsymbol{\eta}_j\rangle\, \langle\boldsymbol{\eta}_j| \tag{93}$$

$$\hat{Z}\hat{A} = \sum_{i,j} |\boldsymbol{\eta}_j\rangle\, \alpha(\boldsymbol{\xi}_i)\zeta(\boldsymbol{\eta}_j)\langle\boldsymbol{\eta}_j|\boldsymbol{\xi}_i\rangle\, \langle\xi_i| \tag{94}$$

The commutator in the first basis is

$$\langle\boldsymbol{\xi}_k|\, \hat{A}\hat{Z} - \hat{Z}\hat{A}\, |\boldsymbol{\xi}_l\rangle = [\alpha(\boldsymbol{\xi}_k) - \alpha(\boldsymbol{\xi}_l)]\, \langle\boldsymbol{\xi}_k|\, \hat{Z}\, |\boldsymbol{\xi}_l\rangle \tag{95}$$

In accordance with the statistical postulate for a state $|\psi\rangle = \sum_i \lambda_i\, |\boldsymbol{\xi_i}\rangle$, the probability that A has value a is

$$p(a) = \sum_i \lambda_i^2 \delta_{a,\alpha\boldsymbol{\xi}_i)}. \tag{96}$$

In plain English $p(a)$ is the sum of probabilities of being in a state $|\boldsymbol{\xi}_i\rangle$ of a basis of a CSCO which includes A, over all states for which $\alpha(\boldsymbol{\xi}_i) = a$. The number of such states in called the degeneracy of a. The expectation value of A is

$$\begin{aligned}
\langle A\rangle &= \sum_a p(a)a = \sum_a \sum_i a\lambda_i^2 \delta_{a,\alpha(\boldsymbol{\xi}_i)} \\
&= \sum_i \alpha(\boldsymbol{\xi}_i)\lambda_i^2 = \sum_i \alpha(\boldsymbol{\xi}_i)\langle\psi|\boldsymbol{\xi}_i\rangle\langle\boldsymbol{\xi}_i|\psi\rangle \\
&= \sum_i \alpha(\boldsymbol{\xi}_i)\langle\psi|\boldsymbol{\xi}_i\rangle\alpha(\boldsymbol{\xi}_i)\langle\boldsymbol{\xi}_i|\psi\rangle = \langle\psi|\, \hat{A}\, |\psi\rangle
\end{aligned} \tag{97}$$

4 Uncertainty relations

If $\hat{A}$and $\hat{B}$ are Hermitian and incompatible,

$$\left[\hat{A}, \hat{B}\right] = i\hat{C} \tag{98}$$

(thus defined, $\hat{C}$ is also hermitian, check it) it follows that

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4}\langle\hat{C}\rangle^2 \tag{99}$$

where the variance of $\hat{A}$,ΔA^2 is

$$\Delta A^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2 = \langle\psi|\left(\hat{A} - \langle\hat{A}\rangle\right)^2|\psi\rangle \tag{100}$$

Proof

$$|\psi_A\rangle = \left[\hat{A} - \langle\psi|\, \hat{A}\, |\psi\rangle\right]|\psi\rangle = \left[\hat{A} - \langle\hat{A}\rangle\right]|\psi\rangle \tag{101}$$

$$|\psi_B\rangle = \left[\hat{B} - \langle\psi|\, \hat{B}\, |\psi\rangle\right]|\psi\rangle = \left[\hat{B} - \langle\hat{B}\rangle\right]|\psi\rangle \tag{102}$$

the norms are

$$\langle\psi_A|\psi_A\rangle = \langle\psi|\left[\hat{A}-\langle\hat{A}\rangle\right]\left[\hat{A}-\langle\hat{A}\rangle\right]|\psi\rangle = \Delta A^2 \tag{103}$$

$$\langle\psi_B|\psi_B\rangle = \langle\psi|\left[\hat{B}-\langle\hat{B}\rangle\right]\left[\hat{B}-\langle\hat{B}\rangle\right]|\psi\rangle = \Delta A^2 \tag{104}$$

Schwarz inequality

$$\begin{aligned}
\Delta A^2\Delta B^2 &\geq |\langle\psi_A|\psi_B\rangle|^2 = \langle\psi_A|\psi_B\rangle\langle\psi_B|\psi_A\rangle \\
&\geq \langle\psi|\left[\hat{A}-\langle\hat{A}\rangle\right]\left[\hat{B}-\langle\hat{B}\rangle\right]|\psi\rangle\,\langle\psi|\left[\hat{B}-\langle\hat{B}\rangle\right]\left[\hat{A}-\langle\hat{A}\rangle\right]|\psi\rangle \\
&\geq \left(\langle\hat{A}\hat{B}\rangle-\langle\hat{A}\rangle\langle\hat{B}\rangle\right)\left(\langle\hat{B}\hat{A}\rangle-\langle\hat{A}\rangle\langle\hat{B}\rangle\right) \\
&\geq \left\langle\left[\frac{1}{2}\left(\left\{\hat{A},\hat{B}\right\}+\left[\hat{A},\hat{B}\right]\right)-\langle\hat{A}\rangle\langle\hat{B}\rangle\right]\right\rangle\left\langle\left[\frac{1}{2}\left(\left\{\hat{A},\hat{B}\right\}-\left[\hat{A},\hat{B}\right]\right)-\langle\hat{A}\rangle\langle\hat{B}\rangle\right]\right\rangle \\
&\left\langle\left[\frac{1}{2}\left\{\hat{A},\hat{B}\right\}-\langle\hat{A}\rangle\langle\hat{B}\rangle\right]\right\rangle^2-\frac{1}{4}\left\langle\left[\hat{A},\hat{B}\right]\right\rangle^2 \\
&\geq \left\langle\left[\frac{1}{2}\left\{\hat{A},\hat{B}\right\}-\langle\hat{A}\rangle\langle\hat{B}\rangle\right]\right\rangle^2+\frac{1}{4}\left\langle\hat{C}\right\rangle^2
\end{aligned} \tag{105}$$

One easily checks that

$$\left[\frac{1}{2}\left\langle\left\{\hat{A},\hat{B}\right\}\right\rangle-\langle\hat{A}\rangle\langle\hat{B}\rangle\right]\in\mathbb{R} \tag{106}$$

Hence the firsts term of the LHS is positive and

$$\Delta A^2\Delta B^2 \geq \frac{1}{4}\langle\hat{C}\rangle^2. \tag{107}$$

.

References

[1] T.M. Apostol. *Calculus, Volume 2*. Blaisdell book in pure and applied mathematics. Wiley, 1967.