Physics 827: Quantum Mechanics I

(Notes prepared for AU 2011)

Text: R. Shankar, Principles of Quantum Mechanics

Suppl. Reading: E. Merzbacher, Quantum Mechanics (advanced)
D. Griffiths, Introduction to Quantum Mechanics (upper level undergraduate)
R. Blümel, Advanced Quantum Mechanics - the classical-quantum connection (advanced)
L. Schiff, Quantum Mechanics
A. Messiah, Quantum Mechanics I, II (very comprehensive)
C. Cohen-Tannoudji et al, Quantum Mechanics I+II (very comprehensive)

Interactive Text with numerical animations:
M. Belloni / W. Christian / A. J. Cox: Physlet Quantum Physics

I. Mathematical prerequisites

I.0: Prologue:

In quantum mechanics, we describe the possible states that a system can assume as vectors in a complex (in the sense of multifaceted) vector spaces. The time evolution of these states is deterministic and controled by equations of motion such as the Schrödinger-, Dirac-, Heisenberg-, Klein-Gordon-equations, but physical observables do not evolve deterministically, but their values are distributed probabilistically, with probabilities calculated from these quantum mechanical states through a specific algorithm.

Since the states evolve deterministically, so do the probabilities for the values of physical observable, but not these observed values themselves.

Only in the classical limit (typically approached in macroscopic systems) do the probability distributions collapse approximately to δ-functions, and the observables themselves evolve deterministically.

To be able to formalize the rules for computing physical observables and the evolution of their probability distributions, we need to master the mathematics of Hilbert spaces, a special class of linear vector spaces. We will begin the course with that.

To wet your appetite, let me note already now that the quantum mechanical algorithm for computing the probability distributions for physical observables is highly non-trivial and initially quite unintuitive. The reason for this are fundamental uncertainty relations which have a simple mathematical origin but state that the measured values of certain pairs of observable can never be both predicted with certainty (i.e. with δ-function probability distribution) simultaneously. This complicates the computation of probability distribution for physical observables through wave-like interference phenomena which are entirely unexpected from classical considerations. In a sense, the quantum mechanical algorithm boils down to the calculation of "conditional probabilities", where the quantum nature of the phenomena reflects itself in the way these "conditions" are implemented.

I.1 Linear vector spaces

Defn: A linear vector space $\mathbb{V}$ is a collection of objects ("vectors") $|1\rangle, |2\rangle, \ldots, |V\rangle, \ldots, |W\rangle \ldots$ which satisfy the following conditions:

(i) if $|V\rangle, |W\rangle \in \mathbb{V}$, then $\alpha|V\rangle + \beta|W\rangle \in \mathbb{V}$, ($\alpha, \beta$ scalars)

(ii) $\alpha(|V\rangle + |W\rangle) = \alpha|V\rangle + \alpha|W\rangle$

(iii) $(\alpha+\beta)|V\rangle = \alpha|V\rangle + \beta|V\rangle$

(scalar mult. is distributive in both the vectors and scalars)

(iv) $\alpha(\beta|V\rangle) = (\alpha\beta)|V\rangle$ (scalar mult. is associative)

(v) $|V\rangle + |W\rangle = |W\rangle + |V\rangle$ (addn. is commutative)

(vi) $|V\rangle + (|W\rangle + |Z\rangle) = (|V\rangle + |W\rangle) + |Z\rangle$ (addn. is associative)

(vii) $\exists$ null vector $|0\rangle$: $|V\rangle + |0\rangle = |V\rangle$

(viii) $\forall |V\rangle$, exists inverse under addition $|-V\rangle$ such that $|V\rangle + |-V\rangle = |0\rangle$

Defn: The numbers $\alpha, \beta, \ldots$ are called the field over which $\mathbb{V}$ is defined.

Examples: $\{\alpha, \beta, \ldots\} = \mathbb{R}$ $\rightarrow$ real vector space

$\{\alpha, \beta, \ldots\} = \mathbb{C}$ $\rightarrow$ complex vector space

($\{\alpha, \beta, \ldots\} = \{\text{grassmann variables}\}$
$\quad\;\; '' \quad = \{\text{quaternions}\}$ $\rightarrow$ other linear vector spaces)

Consequences:

- $|0\rangle$ is unique: Assume $\exists\, |0'\rangle$ with $|V\rangle + |0'\rangle = |V\rangle \;\; \forall |V\rangle$

$$\Rightarrow |0\rangle + |0'\rangle = |0\rangle$$
$$= |0'\rangle + |0\rangle = |0'\rangle$$

- $0|V\rangle = |0\rangle$: $\quad |0\rangle = (0+1)|V\rangle + |-V\rangle = 0|V\rangle + |V\rangle + |-V\rangle$
$\qquad = 0|V\rangle + |0\rangle = 0|V\rangle$
since $|V\rangle + |0\rangle = |V\rangle \; \forall |V\rangle$

- $|-V\rangle = -|V\rangle$: $\quad |V\rangle + (-|V\rangle) = (1-1)|V\rangle = 0|V\rangle = 0 \quad \forall |V\rangle$
$\Rightarrow -|V\rangle$ has same properties as $|-V\rangle$.

- $|-V\rangle$ is the unique additive inverse of $|V\rangle$:

Assume $\exists |W\rangle$ with $|V\rangle + |W\rangle = |0\rangle$

Add $|-V\rangle$ on both sides: $\underbrace{|-V\rangle + |V\rangle}_{|0\rangle} + |W\rangle = |-V\rangle + |0\rangle = |-V\rangle$

$= |0\rangle + |W\rangle = |W\rangle \quad \Rightarrow |W\rangle = |-V\rangle$

Example 1: $\mathbb{V} = \mathbb{R}^3$, the space spanned by $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$, $\vec{e}_3 = (0,0,1)$ with real coefficients. $\alpha, \beta \in \mathbb{R}$

$\mathbb{V} = \mathbb{C}^3$: same basis, but complex $\alpha, \beta \in \mathbb{C}$.

We pronounce $|V\rangle$ as "ket V".

In general, there is no direction associated with $|V\rangle$, nor have we defined a magnitude for $|V\rangle$. For a set $\{|0\rangle, |1\rangle, |2\rangle, \ldots, |V\rangle, \ldots, |W\rangle \ldots\}$ to form a linear vector space, the concept of "magnitude of $|V\rangle$" is not needed.

Example 2: Consider all real 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $(a,b,c,d \in \mathbb{R})$

We know how to add them, multiply them with real scalars, form their negation (by multiplying with (-1)). The null vector $|0\rangle$ is represented by $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
$\rightarrow$ No obvious length or direction associated with each matrix!

Example 3 All functions $f(x)$ defined in $0 \le x \le L$.

$f(x) + g(x) = h(x) \in \mathbb{V}$ $\quad (|f\rangle + |g\rangle = |h\rangle)$

$\alpha f(x) + \beta g(x) \quad \in \mathbb{V}$ $\quad (\alpha|f\rangle + \beta|g\rangle \in \mathbb{V}$

$|0\rangle$ is the null function $f(x) \equiv 0 \ \forall x$

additive inverse $|-f\rangle$: $-f(x)$.

Linear independence:

Defn: A set of vectors $\{|1\rangle, \dots, |n\rangle\}$ is said to be linearly independent if from $\sum_{i=1}^{n} a_i |i\rangle = |0\rangle$ $\quad (*)$

it follows that $a_i = 0$ for all $i = 1, 2, \dots, n$.

Otherwise, they are linearly dependent.

Linearly independent vectors can not be expressed in terms of each other. If the vectors $|1\rangle, \dots, |n\rangle$ are linearly dependent, then at least 2 coefficients in $(*)$ must be nonzero. Let's assume that $a_4 \neq 0$:

$$(*) \Rightarrow |4\rangle = -\sum_{\substack{i=1 \\ i\neq 4}}^{n} \frac{a_i}{a_4} |i\rangle$$

This expresses $|4\rangle$ in terms of the others.

Exercise: Are the vectors $|1\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$|2\rangle = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$|3\rangle = \begin{pmatrix} -2 & -1 \\ 0 & -2 \end{pmatrix}$

linearly independent?

Write down $|0\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = a_1|1\rangle + a_2|2\rangle + a_3|3\rangle$

$$= \begin{pmatrix} 0 & a_1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} a_2 & a_2 \\ 0 & a_2 \end{pmatrix} + \begin{pmatrix} -2a_3 & -a_3 \\ 0 & -2a_3 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_2-2a_3 & a_1+a_2-a_3 \\ 0 & a_2-2a_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$\Rightarrow \quad a_2 = 2a_3 \quad$ and $\quad a_1+a_2-a_3 = a_1+a_3 = 0$

$\Rightarrow \quad \begin{matrix} a_2 = 2a_3 \\ a_1 = -a_3 \end{matrix}$ works for any a_3. For example, $a_3=1$:

$$-|1\rangle + 2|2\rangle + |3\rangle = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} -2 & -1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Answer: they are linearly dependent.

Dimension of a vector space:

Defn: A vector space has dimension n if it can accomodate a maximum of n linearly independent vectors. It will be denoted as $\mathbb{V}^n(\mathbb{R})$ or $\mathbb{V}^n(\mathbb{C})$ depending on whether it is a vector space over the field of real or complex numbers.

Example 2x2 matrices form a 4-dimensional vector space.

Proof: $|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $|2\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; $|3\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$; $|4\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

are linearly independent. (Why?)

$\rightarrow \dim \mathbb{V} \geq 4$

But since any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a|1\rangle + b|2\rangle + c|3\rangle + d|4\rangle$ can be written in terms of these 4 vectors, there are no other vectors in $\mathbb{V}$ (i.e. no other 2x2 matrices) that are linearly independent of $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$.

$\Rightarrow \dim \mathbb{V} = 4$ $\square$

If a,b,c,d are real (complex) $\Rightarrow$ real (complex) 4-dim vector space

Theorem Any vector $|V\rangle$ in an n-dimensional vector space can be written as a linear combination of n linearly independent vectors $|1\rangle, \ldots, |n\rangle$.

(Proof by contradiction: if this were not the case, we would have found an $n+1^{st}$ linearly indep. vector ↯).

Defn: A set of n linearly independent vectors in an n-dim. vector space is called a basis.

Hence, any $|V\rangle$ can be written as $|V\rangle = \sum_{i=1}^{n} v_i |i\rangle$. (**)

Defn: The coefficients v_i in this basis expansion are called the components of $|V\rangle$ in the basis $\{|i\rangle \mid i=1,\ldots,n\}$.

Theorem The expansion (**) is unique.

Proof by contradiction: $|V\rangle = \sum_{i=1}^{n} v_i |i\rangle = \sum_{i=1}^{n} v_i' |i\rangle$

$\Rightarrow |0\rangle = \sum_{i=1}^{n} (v_i - v_i')|i\rangle$

Since $\{|i\rangle\}$ are linearly independent, all coefficients $v_i - v_i'$ must vanish ◆

While the components of $|V\rangle$ in a given basis are unique, they will be different in a different basis. Different bases in $\mathbb{V}$ are like different coordinate systems in $\mathbb{R}^3$. The vector $|V\rangle$ exists in an abstract way, independent of the choice of basis, but its components depend on the choice of basis. Mathematicians like to think about vectors without reference to a choice of basis (in a "coordinate free" manner); physicists in general prefer to represent vectors by their components in a basis set.

The "physics constitution" guarantees every physicist the fundamental right to choose any basis of her liking for such a representation. We will learn the rules of basis-transformations that guarantee that any (correct) physical conclusion is <u>independent of the choice of representation</u> (or basis).

But beware: Depending on the problem at hand (its symmetries, mathematical structure, external constraints), results may be much easier to derive in one basis-representation than in another. As Steven Weinberg said[*)]: "You can choose any basis you want, but if you choose the wrong one, you'll be sorry." So exercise your constitutional right carefully!

In a basis representation, the procedures of adding vectors and multiplying them by a scalar become really easy: you simply perform the same operations on the coefficients:

$$\left.\begin{aligned} |V\rangle &= \sum_{i=1}^{n} v_i\,|i\rangle \\ |W\rangle &= \sum_{i=1}^{n} w_i\,|i\rangle \end{aligned}\right\} \quad |V\rangle + |W\rangle = \sum_{i=1}^{n}(v_i+w_i)\,|i\rangle\,; \quad \alpha|V\rangle = a\sum_{i=1}^{n} v_i\,|i\rangle = \sum_{i=1}^{n}(a v_i)\,|i\rangle\,.$$

---

*) "3rd Law of Progress in Theoretical Physics", paraphrased from S. Weinberg, in "Asymptotic Realms of Physics", A. Guth, K. Huang, R. L. Jaffe (Eds.), MIT Press, 1983