Lecture Notes

Quantum Mechanics

S. Murugesh

IIST-Trivandrum

Last update: August 17, 2010

Contents

Chapter 1

Mathematical Preliminaries

We shall start with a brief review of some basic math ideas we will frequently invoke, without discussing them in detail (See *Topics in Algebra* by I. N. Herstein for a complete rehash). Particularly,

1.1 Vector Spaces

For the time being it suffices to know that a vector space is a Set of elements, say S, with a property *called* addition, denoted $+$, *defined on* the set S. We say it is *called* addition, because we are free to define the *addition*, but as long as it is *commutative, associative and linear*. We shall denote such a vector space by $(S, +)$, or simply S. The term *defined on* implies that

$$+ : S \times S \to S. \tag{1.1}$$

$$\forall s_1, s_2 \in S \Rightarrow s_1 + s_2 \in S. \tag{1.2}$$

Elements of such a vector space are then called *vectors*.
Examples:

1. Let S be $\mathbb{R}^1$, the real line. Evidently, S is a vector space with $+$ being the usual arithmetic addition.

2. The set $\mathbb{N}$ of all natural numbers is a vector space, under usual addition.

3. The set of all n'th order real polynomials forms a vector space under addition.

4. The set of all possible complex periodic piece-wise smooth functions (with period, say, L) forms vector space.

5. The set S of all $N \times M$ real matrices forms a vector space, where $+$ is the matrix addition. The most common being the Euclidean n-dimensional space, which is the set of all column (or row) matrices.

Completeness:
Understood that the operation *addition* just *maps* two elements (vectors) in S to a third element in S, a subset of $C \subset S$, is said to be *Complete* if any vector in S can be written as a linear combination of vectors in C. I.e., a subset $C = \{c_i\}, i = 1, 2, ...n$, is said to be complete if

$$\forall s \in S \;\; \exists \;\; \{a_i\}, i = 1, 2, ...n : a_i c_i = s. \tag{1.3}$$

a_i are referred to as **scalars** and belong in some field.[1,2] Such a subset C is said to **span** the vector space S.
Examples:

1. The set of functions defined as $C = \{x^0, x^1, x^2, ...x^n\}$ forms a complete set for the vector space S of all possible polynomials upto order n.

2. The infinite set of functions defined by $C = \{\exp(in\pi x/L)\}, n = 0, 1, 2, ...\infty$, is a complete set in the vector space of all complex periodic piece-wise smooth functions of period L.

3. The vector space S of all 2×2 real matrices is spanned by the subset

$$C = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}.$$

   C is thus complete in S.

For a given vector space S, the minimal value of n - the number of elements in the complete set C- is called the **dimension** of the vector space. In fact, if C in eg.1 spans the vector space of all polynomials of order n, then so does $C' = \{x^0, x^1, x^2, ...x^n, x^1 + x^2\}$. We say the set C' is **over-complete**. Further, notice that the complete set need not be unique. For, again in eg.1 above, the vector space of polynomials can also be spanned by the complete set $C'' = \{x^0+x^1, x^0-x^1, x^2, ...x^n\}$. The complete subsets C and C' give two different **basis** for the vector space S. Choosing one over the other amounts to merely a change of basis vectors.
Linear Independence:
A set of vectors $\{s_i\}, i = 1, 2, ..m$ is said to be linearly independent if

$$a_i s_i = 0 \Rightarrow a_i = 0 \; for \; all \; i, \tag{1.4}$$

where a_i are all scalars.
The following statements must be apparent:

1. The maximum number of linearly independent vectors possible in a n-dimensional vector space is n.

2. Any set of n linearly independent vectors in a n-dimensional vector space forms a complete set in that vector space, and thus forms a basis.

[1]The nature of $\{a_i\}$ tells the nature of the vector space. Thus, if $\{a_i\}$ can only be real, then the vector space is a *real* vector space. This takes us into the subject of *Fields*, which we shall deliberately avoid.

[2]We have adopted the Einstein summation convention, wherein summation is assumed when repeated indices occur unless stated otherwise. Thus $c_i a_i \equiv \sum_i c_i a_i$.

1.2 Inner product

Just as the 'addition' used in the previous section was a map of any two given elements of the set to a third element of the set S, the *inner product* (or dot product), indicated by a '$\cdot$', is a map of any two given elements of the set to a element of the underlying *field* $\mathbb{F}$ (remember the elements of the field are *scalars* and can be real or complex depending on the nature of the vector space).

$$\cdot : S \times S \to \mathbb{F}. \tag{1.5}$$

$$\forall s_1, s_2 \in S \Rightarrow s_1 \cdot s_2 \in \mathbb{F}. \tag{1.6}$$

A vector space endowed with a valid inner product is called a *inner product space*. Euclidean inner product is the most common, defined in the $\mathbb{R}^n$, this is:

$$\forall \vec{u} = \{u_1, u_2, ...u_n\}, \vec{v} = \{v_1, v_2, ...v_n\} \in \mathbb{R}^n \Rightarrow \vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + ...u_n v_n \in \mathbb{F}. \tag{1.7}$$

No matter what the specific definition of the scalar product, it should obey certain tenets:

1. The $\cdot$ product treats both vectors alike, i.e.,

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}. \tag{1.8}$$

2. For any two scalars a and b , and vectors $\vec{u}, \vec{v}$ and $\vec{w}$,

$$(a\vec{u} + b\vec{v}) \cdot \vec{w} = a\vec{u} \cdot \vec{w} + b\vec{v} \cdot \vec{w}. \tag{1.9}$$

3. The $\cdot$ product is positive definite

$$\vec{u} \cdot \vec{u} \geq 0 \tag{1.10}$$

$$\vec{u} \cdot \vec{u} = 0 \quad \Longleftrightarrow \quad \vec{u} = 0, \tag{1.11}$$

   and also defines the norm, or *length*, of a vector

$$||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}}. \tag{1.12}$$

Two vectors are said to be **orthogonal** iff their $\cdot$ product vanishes. A vector $\vec{u}$ is **normal** if $\vec{u} \cdot \vec{u} = 1$. A set of vectors $\vec{u}_i, i = 1, 2, ..n$ are **orthonormal** if

$$\vec{u}_i \cdot \vec{u}_j = \delta_{ij}. \tag{1.13}$$

The following statements should be obvious:

1. Orthogonal vectors are also linearly independent. The converse, however, need not be true.
2. In a n-dimensional vector space the maximal number of orthonormal vectors possible is n.
3. Given any n linearly independent vectors in n-dimensional space, a suitable $\cdot$ product can always be defined such that the n given vectors are orthonormal.

Notice that if the vectors $\vec{u}, \vec{v}$ in Eq. (1.7) are seen as row matrices u and v of order $1 \times n$, their dot product can also be written as

$$\vec{u} \cdot \vec{v} = u^T v, \tag{1.14}$$

where u^T is the transpose matrix of u of order $n \times 1$, and the $\cdot$ product becomes just the usual matrix multiplication. Some crucial differences have to remembered:

1. u and u^T do not belong to the same vector space.

2. One has to be careful that the transposed matrix goes in the left, for

$$u^T v \neq v u^T. \tag{1.15}$$

A **co-vector space** of a vector space S, denoted S^*, consists of elements that map vectors to scalars. The entire co-vector space S^* maps S *on to* the field $\mathbb{F}$.

$$\tilde{u} \in S^* \rightarrow \tilde{u}(\vec{v}) \in \mathbb{F} \forall \vec{v} \in S, \tag{1.16}$$

and

$$S^*(S) = \mathbb{F}. \tag{1.17}$$

We note from Eq. (2.12), a sample of a co-vector, and a co-vector space. If S is the vector space of all $1 \times n$ row vectors, the co-vector space S^* is the set of all *transpose* vectors consisting of $n \times 1$ column vectors. The action of a co-vector on a vector, in this case, is merely the usual matrix multiplication. Evidently the vector space and its co-space have the same dimension.

1.3 Dirac Bra-Kets

In Section-I we realized that vectors can be represented in different forms - numbers, real or complex functions, polynomials, matrices. In physics vectors are a mathematical representation of physical quantities. Which of these representations will be appropriate will be determined by the physical quantity in question. In classical physics *displacement, velocity, momentum, force* are some common vectors that can be represented using *ordered set of functions*, or, a column or row matrix of functions. However, the ideas discussed so far apply to vectors in any representative form. A common notation for vectors , no matter what representation we imply, is provided by Dirac's Bra and Ket notation - one that will be used almost always in Quantum Mechanics.

A vector $\vec{u}$ will denoted by notation referred to as a *Ket*, and denoted by $|u\rangle$, while its conjugate $\tilde{u}$, will be a *Bra* denoted by $\langle u|$. While $|u\rangle$ is a element of some vector space, $\langle u|$ is a element of its conjugate vector space. The action of a conjugate vector on any vector is a scalar, the *Bra-Ket* (braket henceforth), denoted by $\langle u|v\rangle$.

Although we have discussed brakets for vector spaces with a dot product here, truly we have Hilbert spaces in mind. Hilbert spaces are little more than vector spaces, a discussion we shale postpone for the time being.

Chapter 2

Quantum Mechanics: Fundamentals

Quantum Mechanics is the physics at sub-atomic scale. As we shall discuss soon, the behavior at this scale is fundamentally different from the mechanics of macroscopic objects governed by Newton's three laws - what we shall hitherto refer to as classical physics.

Through several experiments in the first two decades of the 20th century it was increasingly becoming clear that the behavior at sub-atomic level cannot be completely described by classical physics. We discuss here one such experiment, and point out the inadequacies of classical physics, and the striking features that form the crux of quantum mechanics.

2.1 Stern-Gerlach (SG) Experiment

A schematic diagram of the experiment is shown below. A hot gas of Silver atoms is allowed to escape through a pin hole. This beam is collimated and made to pass through a magnetic field. The atom has 47 electrons with one unpaired electron, giving the atom the effective magnetic moment of a single electron. We take the direction of magnetic field to be the z-direction.

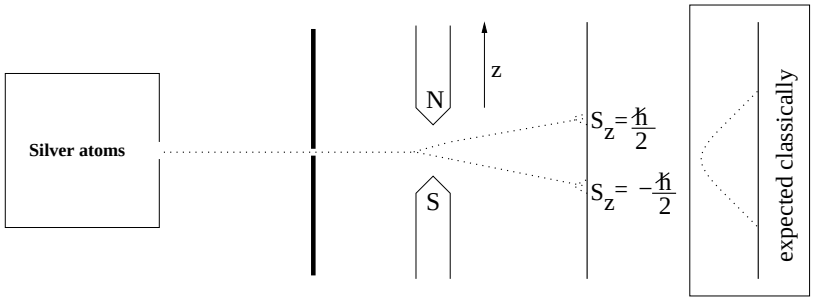

Figure 2.1: The Stern-Gerlach experiment: A beam from a hot gas of Silver atoms is collimated and passed through a magnetic field field along the z-direction. The atoms are finally collected on a screen. A rough classically expected scenario is shown in the box.

The interaction potential for a atom is given by $E = -\vec{\mu}\cdot\mathbf{B}$, where $\vec{\mu}$ is the magnetic moment of the atom. The magnetic moment is directly related to the spin angular momentum vector, $\vec{\mu} = \gamma\mathbf{S}$,

where $\gamma = (eg/2m)$ is the gyromagnetic ratio. Owing to inhomogeneity in the field, the force experienced by the atom is

$$\mathbf{F} = -\nabla E. \tag{2.1}$$

The field being in the z-direction, we have

$$F_z = -\mu_z \frac{\partial}{\partial z} B_z. \tag{2.2}$$

The other two components of $\mathbf{F}$ can be ignored for the discussion that follows. The magnetic moments of each atom in the beam entering the magnetic field is random, hence the beams depositing on the screen are expected to be distributed symmetrically, and continuously about the center, decreasing with the distance. However, *the deposits are noticed only at two spots equally intense, and equidistant, about the center*. The S_z magnitude of atoms corresponding to the two spots can be back calculated (using simple kinematics) and found to be $\pm\hbar/2$, where $\hbar$ is the Planck's constant h divided by 2π.

The experiment reveals a fundamental feature of all quantum systems: Though the magnetic moments of the atoms in the source are random, *the measured value of a physical quantity* (in this case z-component of spin-angular momentum $\mathbf{S}$) *can assume only one of a set of possible values* (in this case $\pm\hbar/2$). We shall call such special values as **eigenvalues** of the physical quantity. The total number of such possible values for a given physical quantity will depend on the system in question.

Let us label the setup consisting of the collimator and magnetic field in the z-direction as a black box S_z, and indicate the two beams as $|S_z^{\pm}\rangle$, pronounced *ket* $S_z^{\pm}$. Figure 2.2 shows the results from a few additional scenarios.

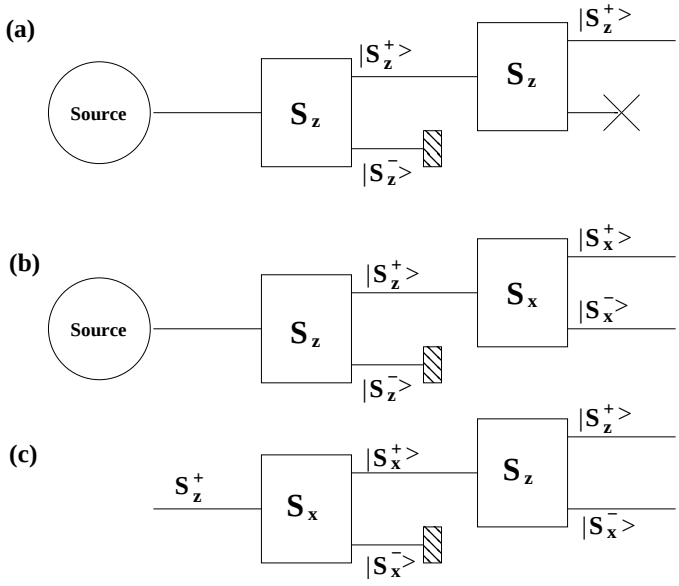

Figure 2.2: Three additional scenarios: The black box with S_x refers to the SG set up with applied magnetic field along the x-direction.

The observations have a analog in polarized light. Linearly polarized light can exist in two states *vertical* and *horizontal*, referring to the direction of polarization of the electric field. Thus the field of a vertically($\hat{\mathbf{i}}$) polarized light propagating in the z direction (angular frequency ω, wave vector k in vacuum) can be written as $\mathbf{E_i} = E\cos(kx - \omega t)\hat{\mathbf{i}}$. A vertical (horizontal) polarizer allows only vertically (horizontally) polarized light, quite like the situation in Figure 2.2a. On the contrary, if the light were polarized in a direction 45 degrees about the vertical,

$$\mathbf{E_e} = E\cos(kx - \omega t)(\hat{\mathbf{i}} + \hat{\mathbf{j}})/\sqrt{2} = \mathbf{E_i} + \mathbf{E_j}, \tag{2.3}$$

both vertical or horizontal polarizers can let the respective components pass through. The scenario involving linearly polarized light in Figure 2.3.

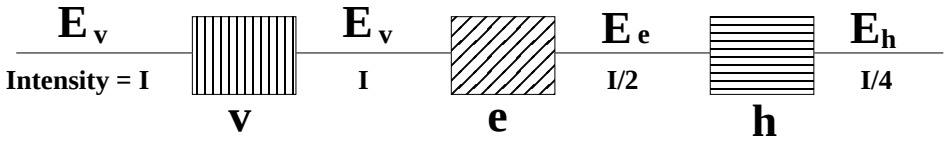

Figure 2.3: Vertically polarized light passing through an array of three polarizers, vertical (v), angular at 45 degrees (e) and horizontal (h). The intensity I is un-diminished at the first polarizer, reduces to $I/2$, and $I/4$ subsequently. Evidently, the scenario is analogous to Figure 2.2c. No light will emerge from the polarizer h on removing the polarizer e from the array, similar to Figure 2.2a.

2.2 Fundamental postulates:

The analogy with polarized light suggests that (from second stage of Figure 2.2c),

$$|S_x^{\pm}\rangle = \frac{1}{\sqrt{2}}(|S_z^{+}\rangle \pm |S_z^{-}\rangle). \tag{2.4}$$

In analogy with light we have also the used factor $1/sqrt2$, the reason for which will be clear soon. As the $|S_z^{\pm}\rangle$ states will be used frequently, we shall use the simpler notation $|\pm\rangle$ to refer to these two states. I. e., each atom in the beam entering the box S_z are in a state $|S_x^{+}\rangle$ which is a *linear superposition* of two states, however, upon measurement they are found to be only in either of the two states $|+\rangle$ and $|-\rangle$ - a feature often referred to as *quantum non-determinism* or *quantum fuzziness*. This brings us to the first quantum postulate:

- A quantum system can exist in a *superposition* of states.

The primary striking feature noted in the SG experiment, that there were only two beams, as opposed to a continuum of beams, is contained in the second postulate -

- The measurement of any physical quantity, or *observable*, on a quantum system yields only one of a possible set of special values called *eigenvalues*.

The observation shown in Figure 2.2a reveals a third fundamental feature of quantum mechanics -

- Upon measurement of a observable, a quantum system *collapses* to one of the special states of the observable - *eigenstates* - corresponding to the measured *eigenvalue* (and continues to remain so).

2.3 The state space

Just as $|S_x^{\pm}\rangle$ can be written as a linear combination of $|\pm\rangle$, Eq. (2.21) can be reversed to write

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|S_x^+\rangle \pm |S_x^-\rangle). \tag{2.5}$$

But there is nothing special about the x and z directions. The y-component of spin angular momentum must be equally expressible in a similar fashion. For example, the situation shown in Figure 2.4 should be equally true.

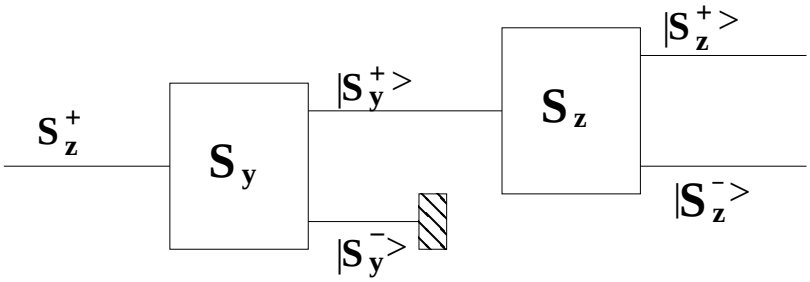

Figure 2.4: A scenario analogous to Figure 2.2c, but with magnetic field along the y-direction. The $|+\rangle$ beam again splits into two with equal numbers, with states $|S_y^{\pm}\rangle$. The beam with atoms in state $|S_y^+\rangle$ further split into two beams, again with equal number of atoms, in states $|\pm\rangle$.

Having exhausted both possibilities of linear superposition in Eqs. 2.21 and 2.5, we again look to polarized light for inspiration. Polarized light can exist in another form apart from linearly polarized light, namely *circularly polarized* light. Circularly polarized light is a linear superposition of two linearly polarized light waves out of phase by a factor $\pi/2$:

$$\mathbf{E}_{R,L} = E\cos(kx - \omega t)\hat{\mathbf{i}} + E\cos(kx - \omega t \pm \pi/2)\hat{\mathbf{j}}. \tag{2.6}$$

The subscript R(or L) stands for right (or left) circularly polarized light for a phase shift of $+$ (or $-$) $\pi/2$. This is better expressed using a complex notation as

$$\mathbf{E}_{R,L} = Re(Ee^{i(kx-\omega t)}\hat{\mathbf{i}} \mp iEe^{-i(kx-\omega t)}\hat{\mathbf{j}}). \tag{2.7}$$

Though we have prompted use of complex notation we should remember that the measured quantity in physics is always real, and hence the *real* part of the *complex* field. This provides a hint of another combination of our $|\pm\rangle$ states, that is linearly independent of $|S_x^{\pm}\rangle$ states:

$$|S_y^{\pm}\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm i|-\rangle). \tag{2.8}$$

The complex usage has been necessitated because our observations tell us that the z-component S_z can take only two values, i.e., can exist only in two states, while S_x, S_y and S_z are linearly independent components of a three dimensional vector, eigenstates of which have to be linearly independent combinations of the two S_z states. That the states involve complex numbers is not any cause of worry. After all it is only the eigenvalues, not eigenstates, that are physically measured quantities. So all we should demand is that for eigenstates to be acceptable physically for a given observable, the corresponding eigenvalues should be real.

The magnetic field in the SG experiment can be oriented in any arbitrary direction. Thus if we were to replace the box S_x in Figure 2.2b with a box, say $\hat{\mathbf{e}}$, indicating the direction of the field, the only difference in the outcome is that the two beams (there are only two beams still) are not of the same intensity, i.e., they do not carry equal number of atoms.

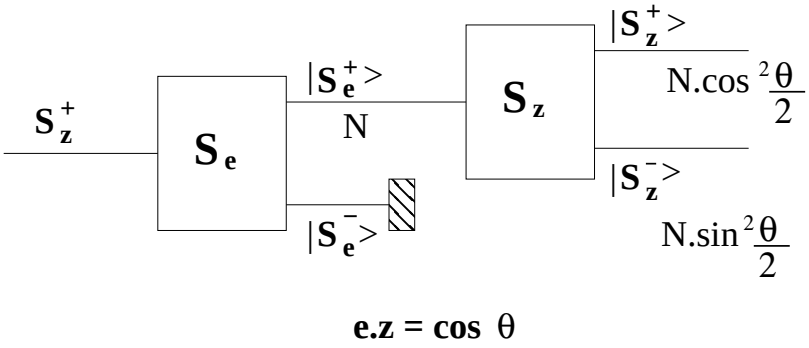

Figure 2.5: A scenario analogous to Figure 2.2c, with magnetic field along some arbitrary direction denoted by e. The $|+\rangle$ beam splits only into two, denoted by states $|\pm\rangle$, but the numbers in each beam are different. At the box S_z the beam in state $|S_e^+\rangle$ splits into two, states $|\pm\rangle$, but again with different numbers.

The observation indicates that the state $|S_e^\pm\rangle$ should again be a linear combination of states $|\pm\rangle$, but with different strengths relating to the direction e of the field with respect to z-direction. Or, any arbitrary state $|S_e\rangle$ of the incoming beam must be representable as a linear superposition of the states $|\pm\rangle$ of the form

$$|S_e\rangle = \alpha|+\rangle + \beta|-\rangle, \tag{2.9}$$

where α and β are suitable complex numbers, relating to the strengths of the two output beams.

The set of all such superpositions of the two states $|\pm\rangle$, forms a complex vector space (2-dimensional in this particular case). Any spin state of the atom is represented by a vector in this vector space called the *state space* or *Hilbert space* $\mathcal{H}$.

NOTE: It is important to note that what we refer to as the *state of the atom*, as long as we are talking about the SG set up, *is the spin-state of the atom*. This is so because our experimental set up is designed to only measure one component of spin angular momentum of the atom, not other quantities such as position, momentum, energy, or even any other component of the spin angular momentum. In fact, when ever we use the term *physical system* in quantum mechanics what we essentially mean is only the observable our experiment in question is designed to measure. Thus in

this case of the SG experiment, the 'system' refers only to the *particular component of spin angular momentum* the set up studies..

2.4 Dual space, Operators in state space and some algebra

Having familiarized with the idea of states we shall see how to work with them by discussing their algebra in deail.

A *dual conjugte* vector (or co-vector) to a *ket* $|\alpha\rangle$ is a *bra* - $\langle\alpha|$.

$$\mathbb{DC} : |\alpha\rangle = \langle\alpha|. \tag{2.10}$$

A dual vector maps a ket vector to a scalar, denoted by a braket:

$$\langle\alpha|(|\beta\rangle) = \langle\alpha|\beta\rangle \in \mathbb{C}. \tag{2.11}$$

The set of all dual vectors (*bra* vectors) form a vector space - the dual vector space($\mathcal{H}^*$). One *postulate* about these brakets is that

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*, \tag{2.12}$$

where $*$ in the superscript stands for complex conjugation.

The following statements hold true about the ket and bra vectors:

1. $c(|\alpha\rangle + |\beta\rangle) = c|\alpha\rangle + c|\beta\rangle$.
2. $\mathbb{DC} : c|\alpha\rangle = c^*\langle\alpha| = \langle\alpha|c^*$.
3. It follows from Eq. (2.12) above that $\langle\alpha|\alpha\rangle$ is *real*.
4. Further to 3, (it will be shown later) $\langle\alpha|\alpha\rangle \geq 0$, the equality valid *iff* $|\alpha\rangle = 0$.
5. Consequent to 4, we define the *norm* of a state $|\alpha\rangle$ as $\sqrt{\langle\alpha|\alpha\rangle}$.
6. A vector is said to be normalized if $\langle\alpha|\alpha\rangle = 1$.
7. Two vectors $|\alpha\rangle$ and $|\beta\rangle$ are said to be orthogonal if $\langle\alpha|\beta\rangle = 0 = \langle\alpha|\beta\rangle$.

We have already seen the state space of a *system* in the previous section. The state space is a vector space ($\mathcal{H}$), and the state of a system is represented by a vector in this state space. In quantum mechanics physical quantities - *observables*- are *operators defined on the state space*. They map one state space vector to another. I.e., given any operator $\mathbf{A}$, representing some observable,

$$\forall\ |\alpha\rangle \in \mathcal{H} \Rightarrow \mathbf{A}|\alpha\rangle \in \mathcal{H}. \tag{2.13}$$

Operators act on ket vectors from the left, and on bra vectors from the right.

$$\langle\alpha|\mathbf{A} \in \mathcal{H}^*. \tag{2.14}$$

Of particular importance are eigenstates of an observable, defined such that

$$\mathbf{A}|a_i\rangle = a_i|a_i\rangle, \quad i = 1, 2, ...n. \tag{2.15}$$

The scalar quantities a_i are the eigenvalues of the observable $\mathbf{A}$ associated with each eigenkets $|a_i\rangle$ respectively. The operators have the following properties:

1. $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$.
2. $(\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + (\mathbf{Y} + \mathbf{Z})$.
3. Two operators $\mathbf{X}$ and $\mathbf{Y}$ are equal only if their action on any arbitrary $|alpha\rangle$ is equal:

$$\mathbf{X} = \mathbf{Y} \Rightarrow \mathbf{X}|\alpha\rangle = \mathbf{Y}|\alpha\rangle, \ \forall \ |\alpha\rangle \in \mathcal{H}. \tag{2.16}$$

4. $\mathbf{X}c|\alpha\rangle = c\mathbf{X}|\alpha\rangle$.
5. $\mathbf{X}(c_1|\alpha_1\rangle + c_2|\alpha_2\rangle) = \mathbf{X}c_1|\alpha_1\rangle + \mathbf{X}c_2|\alpha_2\rangle$.
6. $\mathbf{XY} \neq \mathbf{YX}$.
7. The Hermitian adjoint of $\mathbf{X}$, denoted $\mathbf{X}^\dagger$, is defined such that

$$|\alpha\rangle \rightarrow \langle\alpha| \Rightarrow \mathbf{X}|\alpha\rangle \rightarrow \langle\alpha|\mathbf{X}^\dagger. \tag{2.17}$$

8. $\mathbf{X}(\mathbf{Y}|\alpha\rangle) = (\mathbf{XY})|\alpha\rangle$.
9. $(\mathbf{XY})^\dagger = \mathbf{Y}^\dagger\mathbf{X}^\dagger$ (show this).
10. An operator $\mathbf{X}$ is called *Hermitian* if $\mathbf{X} = \mathbf{X}^\dagger$.
11. $\langle\alpha|\mathbf{X}|\beta\rangle = \langle\beta|\mathbf{X}^\dagger|\alpha\rangle^*$.
12. For a Hermitian operator $\mathbf{X}$, $\langle\alpha|\mathbf{X}|\beta\rangle = \langle\beta|\mathbf{X}|\alpha\rangle^*$.

An **outer product** of two vectors $|\alpha\rangle$ and $|\beta\rangle$, denoted $|\alpha\rangle\langle\beta|$, is an operator defined such that

$$|\alpha\rangle\langle\beta| \cdot |\gamma\rangle = |\alpha\rangle\langle\beta|\gamma\rangle = \langle\beta|\gamma\rangle|\alpha\rangle \ \in \ \mathcal{H}. \tag{2.18}$$

The **associative axiom** lets us make sense of any acceptable combination of kets, bras and operators.

$$\mathbf{A} \cdot |\alpha\rangle \cdot \langle\beta| \cdot \mathbf{B} \cdot |\gamma\rangle = \mathbf{A}|\alpha\rangle\langle\beta|\mathbf{B}|\gamma\rangle = \langle\beta|\mathbf{B}|\gamma\rangle\mathbf{A}|\alpha\rangle. \tag{2.19}$$

Products such as $\mathbf{A}\langle\alpha|$ are not defined.

2.5 Completeness and projection operators

Note that in measuring a physical quantity associated with a system, say $\mathbf{A}$ (such as $\mathbf{S_z}$), the outcome of the measurement is only one of its eigenvalues, say $a_i, i = 1, 2, ..n$ ($\pm\hbar/2$ for $\mathbf{S_z}$). If $|a_i\rangle,\ i = 1, 2, ...n$ are the corresponding eigenstates, defined through the relation

$$\mathbf{A}|a_i\rangle = a_i|a_i\rangle, i = 1, 2, ..n, \tag{2.20}$$

we state some important results as problems.

Problem 2.1: *Show that the eigen values of a Hermitian operator are real:* $\mathbf{A} = \mathbf{A}^\dagger$ *and* $\mathbf{A}|a\rangle = a|a\rangle \Rightarrow a = a^*$.

After all any physical measurement is real. The measured values are real. Given that eigenvalues are the measured values for any physical quantity, we conclude that *all physical quantities are represented by Hermitian operators.*

Problem 2.2: *Show that the eigen states of a Hermitian operator are orthogonal if the eigenvalues are distinct: given* $\mathbf{A} = \mathbf{A}^\dagger$, $\mathbf{A}|a_i\rangle = a_i|a_i\rangle, i = 1, 2, ..n$ *with all kets of norm=1, and* $a_i = a_j\ iff\ i = j \Rightarrow \langle a_i|a_j\rangle = \delta_{ij}$.

Although stated for the case of distinct eigenvalues (we say the operator is non-degenerate), the result can also be made true even if some of the eigenvalues are same.

In the end of Section 2.3 we noted how the state space is constructed using $|S_z^{\pm}\rangle$ states - as a collection of all possible complex superpositions of the eigenkets. Indeed, this is generally true for the state space of any physical system -*the state space can be spanned by eigenkets of any Hermitian operator defined on the state space.* Or, if $\mathbf{A}$ is some Hermitian operator defined on the state space, its eigenkets $|a_i\rangle,\ i = 1, 2, ..n$ form basis for the state space. This essentially implies that any arbitrary state $|\alpha\rangle$ can be written as linear superposition of these eigenkets:

$$|\alpha\rangle = \sum_i c_i|a_i\rangle, \tag{2.21}$$

for some set of c_is. Taking inner product with $\langle a_j|$, and using the ortho normality of $|a_i\rangle$s in Eq. (2.21), we have

$$\langle a_j|\alpha\rangle = \sum_i c_i\langle a_j|a_i\rangle = \sum_i c_i\delta_{ij} = c_j. \tag{2.22}$$

Substituting back in Eq. (2.21) we have

$$|\alpha\rangle = \sum_i \langle a_i|\alpha\rangle|a_i\rangle = (\sum_i |a_i\rangle\langle a_i|)|\alpha\rangle. \tag{2.23}$$

Since this is true for any arbitrary ket $|\alpha\rangle$, we conclude from Eq. (2.23) the condition of completeness:

$$\sum_i |a_i\rangle\langle a_i| = I, \tag{2.24}$$

the identity operator. The action of each of the elements in the sum in Eq. (2.24) on a ket $|\alpha\rangle$ gives the respective ket:

$$|a_i\rangle\langle a_i| \cdot |\alpha\rangle = \langle a_i|\alpha\rangle|a_i\rangle. \tag{2.25}$$

We recognize each element in the sum $|a_i\rangle\langle a_i|$ as a **projection operator** along the direction $|a_i\rangle$.

2.6 The probability hypothesis

Clearly, a state $|\tilde{a}_i\rangle = c|a_i\rangle$, where c is some finite non-zero complex number, is also an eigen state with the same eigen value, since

$$\mathbf{A}|\tilde{a}_i\rangle = c\mathbf{A}|a_i\rangle = ca_i|a_i\rangle = a_i|\tilde{a}_i\rangle. \tag{2.26}$$

Thus *the states $|a_i\rangle$ and $|\tilde{a}_i\rangle$ correspond to the same physical state.* Naturally, it is convenient to ignore its norm, or consider only states which have a unit norm. Henceforth the physical state of the system will be represented only by kets of unit norm.

If the system is given to be in a state $|\alpha\rangle$, the probability of the system to be found in a state $|a_i\rangle$ is given by $|\langle a_i|\alpha\rangle|^2$. Thus, if

$$|\alpha\rangle = \sum_i c_i|a_i\rangle, \quad \langle\alpha|\alpha\rangle = 1 = \langle a_i|a_i\rangle, \tag{2.27}$$

then $|c_i|^2$ is the probability of the system to be in state $|a_i\rangle$.

Example: The state $|S_x^+\rangle$ is a linear combination of $|\pm\rangle$ states

$$|S_x^+\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle). \tag{2.28}$$

Thus if the beam of atoms is in state $|S_x^+\rangle$, the probability of the atoms to be in state $|+\rangle$ is $|\langle +|S_x^+\rangle|^2 = 1/2$ - essentially meaning that half the atoms in the beam will be in state $|+\rangle$.

2.7 Matrix representations

So far we have been discussing operators and states by notations, without giving any representation. The state space is a vector space, and we have seen in chapter 1 that vectors can be represented as numbers, matrices, functions, etc.,. Given that the inner product - the Dirac bracket - is a complex number, allows us to give a matrix representation for operators and states.

Consider an operator $\mathbf{X}$ *defined on* n-dimensional state space $\mathcal{H}$ spanned by eigenstates $|a_i\rangle$, $i = 1, 2, ..n$. Using the completeness condition, we may write

$$\mathbf{X} = \sum_{i,j} |a_i\rangle\langle a_i|\mathbf{X}|a_j\rangle\langle a_j|. \tag{2.29}$$

The bracket sandwiched in the middle, $\langle a_i|\mathbf{X}|a_j\rangle$, is a complex number, say x_{ij}. Since outer products are operators, Eq. (2.29) is a outer product representation of the operator $\mathbf{X}$:

$$\mathbf{X} = \sum_{i,j} x_{ij}|a_i\rangle\langle a_j|. \tag{2.30}$$

In particular, if the states $|a_i\rangle$s are eigenkets of the operator $\mathbf{A}$ with eigen values a_is, respectively, then Eq. (2.29) implies

$$\mathbf{A} = \sum_i a_i|a_i\rangle\langle a_i|. \tag{2.31}$$

Given the basis, the entire information about $\mathbf{X}$ is contained in the $n \times n$ complex numbers of the form $\langle a_i|\mathbf{X}|a_j\rangle$. This allows for a complex matrix representation of $\mathbf{X}$, with elements $\mathbf{X}_{ij} = \langle a_i|\mathbf{X}|a_j\rangle$.

$$\Gamma(\mathbf{X}) = \begin{pmatrix} \langle a_1|\mathbf{X}|a_1\rangle & \langle a_1|\mathbf{X}|a_2\rangle & . & . \\ \langle a_2|\mathbf{X}|a_1\rangle & \langle a_2|\mathbf{X}|a_2\rangle & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix}.$$

The Γ is used to indicate that the matrix is *a* representation for the operator (there are other representations possible). Evidently, the elements of this matrix depend on the basis we are working in.

Along same lines one can also give a matrix representation for an arbitrary state $|\alpha\rangle$, since

$$|\alpha\rangle = \sum_i |a_i\rangle\langle a_i|\alpha\rangle. \tag{2.32}$$

giving a column representation for $|\alpha\rangle$:

$$\Gamma(|\alpha\rangle) = \begin{pmatrix} \langle a_1|\alpha\rangle \\ \langle a_2|\alpha\rangle \\ . \\ . \end{pmatrix}.$$

The corresponding bra-vector $\langle\alpha|$ is given by the row vector

$$\Gamma(\langle\alpha|) = \begin{pmatrix} \langle\alpha|a_1\rangle & \langle\alpha|a_2\rangle & . & . \end{pmatrix}.$$

Using these results it can be directly verified that

$$\Gamma(\mathbf{X}^\dagger) = \begin{pmatrix} \langle a_1|\mathbf{X}^\dagger|a_1\rangle & \langle a_1|\mathbf{X}^\dagger|a_2\rangle & . & . \\ \langle a_2|\mathbf{X}^\dagger|a_1\rangle & \langle a_2|\mathbf{X}^\dagger|a_2\rangle & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix} = \begin{pmatrix} \langle a_1|\mathbf{X}|a_1\rangle^* & \langle a_2|\mathbf{X}|a_1\rangle^* & . & . \\ \langle a_1|\mathbf{X}|a_2\rangle^* & \langle a_2|\mathbf{X}|a_2\rangle^* & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix}.$$

Therefore,

$$\Gamma(\mathbf{X}^\dagger) = (\Gamma(\mathbf{X}))^{T*}, \tag{2.33}$$

where T on the superscript stands for *transpose* of the matrix. Here on we shall avoid using Γ when we use matrices for operators for convenience.

For a product of two operators, say $\mathbf{C} = \mathbf{AB}$, we have,

$$\mathbf{C} = \sum_{i,j,k} |a_i\rangle\langle a_i|\mathbf{A}|a_j\rangle\langle a_j|\mathbf{B}|a_k\rangle\langle a_k|. \tag{2.34}$$

Or,

$$\mathbf{C}_{ik} = \sum_j \langle a_i|\mathbf{A}|a_j\rangle\langle a_j|\mathbf{B}|a_k\rangle = \sum_j \mathbf{A}_{ik}\mathbf{B}_{kj}, \tag{2.35}$$

which is just matrix multiplication.

Problem: 2.3 *Obtain the exterior product representation for* $\mathbf{S_x}, \mathbf{S_y}$ *and* $\mathbf{S_z}$ *using* $|\pm\rangle$ *states.*

Problem: 2.4 *Show that the matrix representation for* $\mathbf{S_x}, \mathbf{S_y}$ *and* $\mathbf{S_z}$ *operators in the* $|\pm\rangle$ *basis is*

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

respectively.

Two additional operators are very useful in quantum mechanics of spin systems. In exterior product form they are defined

$$\mathbf{S}_+ = \hbar|+\rangle\langle -|, \quad \mathbf{S}_- = \hbar|-\rangle\langle +|, \tag{2.36}$$

and are called *raising* and *lowering* operators, respectively, or *ladder* operators.

Problem: 2.5 *Justify the name 'raising' and 'lowering' operators by their action on* $|\pm\rangle$ *states.*

Problem: 2.6 *Obtain the matrix representation for* $\mathbf{S}_\pm$ *operators. Using results of Problems 2.3 and 2.4 show that*

$$\mathbf{S}_\pm = \mathbf{S}_x \pm i\mathbf{S}_y. \tag{2.37}$$

Notice that $\mathbf{S}_+^\dagger = \mathbf{S}_-$.