Mathematical Foundations of Quantum Mechanics 2016-17

Dr Judith A. McGovern

Maths of Vector Spaces

This section is designed to be read in conjunction with chapter 1 of Shankar's Principles of Quantum Mechanics, which will be the principal course text book. Other on-line resources are linked from the course home page.

Another source that covers most of the material at the right level is Griffiths' Introduction to Quantum Mechanics, which has an appendix on linear algebra.

Riley's Mathematical Methods for the Physical Sciences is available as an ebook, and chapter 8 covers much of the material too. This is particularly recommended if Shankar seems initially intimidating. Unfortunately Riley does not use Dirac notation except for inner products, using boldface $\mathbf{a}$ where we would use $|a\rangle$, but if you understand the concepts from that book, the notation used here should not be a barrier. Some further comments on Riley's notation can be found in section 1.4. Riley (or a similar text such as Boas) should be consulted for revision on finding the eigenvalues and eigenvectors of matrices.

This outline omits proofs, but inserts the symbol $^{\mathbb{P}}$ to indicate where they are missing. In the early stages the proofs are extremely simple and largely consist of assuming the opposite and demonstrating a contradiction with the rules of vector spaces or with previous results. Many are in Shankar but some are left by him as exercises, though usually with hints. By the time we get on the properties of operators (existence of inverses, orthogonality of eigenstates) some of the proofs are more involved. Some of the principal proofs are on the examples sheet. Proofs from this section are not examinable, but you are advised to tackle some of them to make sure you understand the ideas.

1.1 Vector Spaces

Definition

Shankar pp 1-3, Riley 8.1, Griffiths A.1

A linear **vector space** is a set $\mathbb{V}$ of elements called **vectors**, $\{|v\rangle, |w\rangle ...\}$, for which
I) An operation, "+", is defined, which for any $|v\rangle$ and $|w\rangle$ specifies how to form $|v\rangle + |w\rangle$
II) Multiplication by a scalar is also defined, specifying $\alpha|v\rangle$
and these operations obey the following rules:

1. The result of these operations is another member of $\mathbb{V}$ (**closure**).
2. $|v\rangle + |w\rangle = |w\rangle + |v\rangle$ (vector addition is commutative)
3. $(|u\rangle + |v\rangle) + |w\rangle = |u\rangle + (|v\rangle + |w\rangle)$ (vector addition is associative)
4. $\alpha(\beta|v\rangle) = (\alpha\beta)|v\rangle$ (scalar multiplication is associative)
5. $1\,|v\rangle = |v\rangle$
6. $\alpha(|v\rangle + |w\rangle) = \alpha|v\rangle + \alpha|w\rangle$ (distributive rule 1)
7. $(\alpha + \beta)|v\rangle = \alpha|v\rangle + \beta|v\rangle$ (distributive rule 2)
8. The null or zero vector is written as $|0\rangle$ (or often, just 0), with $|0\rangle + |v\rangle = |v\rangle$
9. For every vector $|v\rangle$ there is another, denoted $|-v\rangle$, such that $|v\rangle + |-v\rangle = |0\rangle$

Note in the definition of $|-v\rangle$ the minus sign is just part of the name of the inverse vector.

The zero vector is unique. $0|v\rangle = |0\rangle$ for any $|v\rangle^{\,\mathbb{P}}$.
The inverse is unique and given by $|-v\rangle = (-1)|v\rangle^{\,\mathbb{P}}$.
We use "minus" in the following sense: $|v\rangle - |w\rangle = |v\rangle + (-1)|w\rangle = |v\rangle + |-w\rangle$.

If the scalars $\alpha, \beta ...$ are complex (written $\alpha, \beta \in \mathbb{C}$), we have a complex vector space, otherwise ($\alpha, \beta \in \mathbb{R}$) we have a real one. If we want to distinguish we write $\mathbb{V}(\mathbb{C})$ and $\mathbb{V}(\mathbb{R})$, but if we don't specify we assume it is complex. ($\mathbb{C}$ or $\mathbb{R}$ is called the **field** of the space).

These rules just confirm what you do naturally, but:

- You should not assume anything about abstract vectors that is *not* given in the definition.
- The rules apply to many things apart from traditional "arrow" vectors.
- So far there is no concept of "angle" between vectors, nor any way to measure "length".

Examples

- Ordinary 3D "arrow" vectors belong to a real vector space.[1]

[1] "arrow" vectors have length and direction in 3D, but they do not have a fixed starting point, so two vectors are added by placing the tail of second at the tip of the first; multiplication by a scalar changes the length but not the direction. In physics, displacement vectors are a better picture to keep in mind than positions.

- Real numbers form a (very simple) real vector space.
- The set $\mathbb{R}^N$ ($\mathbb{C}^N$) of sequences of N real (complex) numbers, such as $|c\rangle = (c_1, c_2, \ldots c_N)$ form a real (complex) vector space, where '+' is ordinary matrix addition, $|0\rangle = (0, 0, \ldots 0)$ and the inverse is $|-c\rangle = (-c_1, -c_2, \ldots - c_N)$.
- The set of all polynomials such as $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$, with $a_i \in \mathbb{C}$ and $x \in \mathbb{R}$, forms a complex vector space; $|0\rangle$ is the polynomial with all coefficients a_i equal to zero.
- The set of 2×2 complex matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with $a, b, c, d \in \mathbb{C}$, form a complex vector space under matrix addition (in fact any such set of $n \times m$ matrices gives a vector space).

Ket Notation

Shankar p 3, Griffiths A.1

Here we are using the **Dirac notation** for vectors, with the object $|v\rangle$ also being called a **ket**. The text between the "|" and the "⟩" is just a name or label for the ket, which can take many forms—we will see letters, numbers, symbols ($|+\rangle$, $|\heartsuit\rangle$), reminders of how the vector was formed ($|\alpha v\rangle$ for $\alpha|v\rangle$).... Sensible choices of names can help make the algebra easy to follow. The notation prevents abstract vectors being confused with simple numbers.

1.2 Linear Independence, bases and dimensions

Linear Independence

Shankar p 4, Riley 8.1.1, Griffiths A.1

Since there are infinitely many scalars, all vector spaces have infinitely many members.

If from $\mathbb{V}$ we pick n vectors $\{|x_1\rangle, |x_2\rangle, \ldots, |x_n\rangle\}$, the set is said to be **linearly dependent** if it is possible to write $\sum_{i=1}^{n} a_i |x_i\rangle = |0\rangle$ where the coefficients a_i are not all zero. It follows that at least one of the vectors can be written as a sum over the others $^{\mathbb{P}}$.

If this is not possible, the set is **linearly independent**. Any two non-parallel "arrow" vectors are linearly independent; any three arrow vectors in a plane are linearly dependent.

Dimensions and Bases

Shankar pp 5-7, Riley 8.1.1, Griffiths A.1

A vector space has **dimension** N if it can accommodate a maximum of N linearly-independent vectors. It is infinite-dimensional if there is no maximum. We use $\mathbb{V}^N$ if we want to specify the dimension.

A **basis** in a vector space $\mathbb{V}$ is a set $\{|x_1\rangle, |x_2\rangle, \ldots, |x_N\rangle\} \equiv \{|x_i\rangle\}$ of linearly-independent vectors such that every vector in $\mathbb{V}$ is a linear combination of the basis vectors $|x_i\rangle$; that is, for an arbitrary vector $|v\rangle$,

$$|v\rangle = \sum_{i=1}^{N} v_i |x_i\rangle$$

where v_i are suitable coefficients (or **components** or **coordinates**). For a given basis and vector $|v\rangle$, these components are unique$^{\mathbb{P}}$. However in *different* bases, a given vector will have *different* components.

In general components are complex, but for a real vector space (with a suitable choice of basis) they are real.

Example: In real 3-D space, using the usual notation, the vectors $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ form a basis . (These may also be written $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ or $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$.) So does any other set of three non-coplanar vectors.

Every basis in $\mathbb{V}^N$ has N elements; conversely any set of N linearly-independent vectors in $\mathbb{V}^N$ forms a basis$^{\mathbb{P}}$.

When we add vectors, the coordinates add: if $|w\rangle = \alpha|u\rangle + \beta|v\rangle$, with $|u\rangle = \sum u_i|x_i\rangle$, $|v\rangle = \sum v_i|x_i\rangle$ and $|w\rangle = \sum w_i|x_i\rangle$, then $w_i = \alpha u_i + \beta v_i$$^{\mathbb{P}}$.

Any set of at least N vectors which includes a basis as a subset is said to **span the space**; obviously a basis spans the space.

For convenience, we will often write a basis as $\{|i\rangle\} \equiv \{|1\rangle, |2\rangle, \ldots |N\rangle\}$. Recall that what is written inside the ket is just a label. Numbers-as-labels in kets will be widely used, so it is important to remember they have no other significance. $|1\rangle + |2\rangle \neq |3\rangle$!

Representations

Shankar pp 10-11, Riley 8.3, Griffiths A.1

For a given basis $\{|x_i\rangle\}$, and a vector $|v\rangle = \sum_{i=1}^{N} v_i|x_i\rangle$, the list of components v_i **is a representation of** the *abstract* vector $|v\rangle$. We write this as a vertical list (or column vector):

$$|v\rangle \underset{x}{\longrightarrow} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}.$$

The symbol $\underset{x}{\longrightarrow}$ means "is represented by", with the x being a name or label for the basis (which will be omitted if the basis is obvious).

Note that in their own representation the basis vectors are simple:

$$|x_1\rangle \underset{x}{\longrightarrow} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |x_2\rangle \underset{x}{\longrightarrow} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad |x_N\rangle \underset{x}{\longrightarrow} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

If u_i, v_i and w_i are the components of $|u\rangle$, $|v\rangle$ and $|w\rangle$ in this basis, and $|w\rangle = \alpha|u\rangle + \beta|v\rangle$,

$$|w\rangle \underset{x}{\longrightarrow} \begin{pmatrix} \alpha u_1 + \beta v_1 \\ \alpha u_2 + \beta v_2 \\ \vdots \\ \alpha u_N + \beta v_N \end{pmatrix}$$

Hence all manipulations (addition, multiplication by a scalar) of the abstract vectors or kets are mirrored in corresponding manipulations of the column vectors. A fancy way of saying the same this is that all N-dimensional vector spaces are **homomorphic** to $\mathbb{C}^N$, and hence to one another. Practical calculations often start by specifying a basis and working with the corresponding representations of vectors in that basis. We will repeatedly find the same calculations recurring for physically different vector spaces that happen to have the same dimension.

If we have another basis, $\{|y_i\rangle\}$, $|v\rangle$ will be represented by a different column vector in this basis, and the $|x_i\rangle$ will have more than one non-zero component.

Example: given a 2D real vector space with a basis $\{|x_1\rangle, |x_2\rangle\}$ and another $\{|y_1\rangle = |x_1\rangle + |x_2\rangle,\ |y_2\rangle = |x_1\rangle - |x_2\rangle\}$, and $|v\rangle = 2|x_1\rangle + 3|x_2\rangle = \frac{5}{2}|y_1\rangle - \frac{1}{2}|y_1\rangle$, we have for instance

$$|v\rangle \underset{x}{\longrightarrow} \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad |v\rangle \underset{y}{\longrightarrow} \begin{pmatrix} 5/2 \\ -1/2 \end{pmatrix}, \qquad |y_2\rangle \underset{x}{\longrightarrow} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad |x_1\rangle \underset{y}{\longrightarrow} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}.$$

Subspaces and direct sums

Shankar pp 17-18

Given an N-D vector space $\mathbb{V}^N$, a subset of its elements that form a vector space among themselves is a **subspace**.

For examples in ordinary 3-D space:
• all vectors along the x axis are a 1-D subspace: $\mathbb{V}^1_x$
• all vectors in the xy plane which includes the origin are a 2-D subspace: $\mathbb{V}^2_{xy}$.
Note both of these contain the origin, and the inverse of any vector in the subspace.

Any n-D subset of a basis of $\mathbb{V}^N$ will span a new subspace $\mathbb{V}^{n\,\mathbb{P}}$. Of course the space contains all linear combinations of these basis vectors, not just the vectors themselves.

Given two spaces, $\mathbb{V}^N_a$ and $\mathbb{V}^M_b$, (where a and b are just labels), their so-called **direct sum**, written $\mathbb{V}^N_a \oplus \mathbb{V}^M_b$ is the set containing all elements of $\mathbb{V}^N_i$ and $\mathbb{V}^M_j$ *and* all possible linear combinations between them. This makes it closed, and so the direct sum is a new vector space. A set consisting of N basis vectors from $\mathbb{V}^N_a$ and M from $\mathbb{V}^M_b$ forms a basis in $\mathbb{V}^N_a \oplus \mathbb{V}^M_b$, which is an $N + M$ dimensional space. $\mathbb{V}^N_a$ and $\mathbb{V}^M_b$ are subspaces of this space.

Example: $\mathbb{V}^1_x \oplus V^1_y = \mathbb{V}^2_{xy}$. Bases for the two 1-D spaces are the 1-element sets $\{\mathbf{i}\}$ and $\{\mathbf{j}\}$; so $\{\mathbf{i}, \mathbf{j}\}$ is a basis on their direct sum. $\mathbb{V}^1_x$ and $\mathbb{V}^1_y$ are now subspaces of the new space $\mathbb{V}^2_{xy}$. Note that $\mathbb{V}^2_{xy}$ contains points off the x and y axes which are not in either of the component spaces, but are produced by linear combinations (e.g. $2\mathbf{i} - 10\mathbf{j}$).

Note that for this to work, the two spaces must have only the zero vector in common. The direct sum of the xy plane and the xz plane is not four-dimensional!

Product spaces

Shankar pp 248-249 (chapter 10)

A different way of combining two spaces is the "tensor direct product", denoted $\mathbb{V}^N_a \otimes \mathbb{V}^M_b$. Though important in quantum mechanics, it is hard to come up with examples from classical physics. They arise when a system has two distinct aspects, both of which are vectors, and in order to specify the state of the system both vectors have to be given.

If $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ are basis sets for the two spaces, one possible basis for the product space is formed by picking one from each—say the ith from the first set and the jth from the second set. There are $N \times M$ possibilities, so the product space has dimension $N \times M$. These states are written $|i, j\rangle \equiv |a_i\rangle \otimes |b_j\rangle$. The $\otimes$ is best regarded simply as a separator; it doesn't indicate any operation that is carried out.

Note that for $|p\rangle, |q\rangle \in \mathbb{V}_a^N$ and $|v\rangle, |w\rangle \in \mathbb{V}_b^M$, while all vectors $|p\rangle \otimes |v\rangle$ are in $\mathbb{V}_a^N \otimes \mathbb{V}_b^M$, not all vectors in the product space can be written in this way. Those that can are called **separable**, i.e. they have a specified vector in each separate space. The vector $\alpha|p\rangle \otimes |v\rangle + \beta|q\rangle \otimes |w\rangle$ is in the product state but is not separable unless $|p\rangle \propto |q\rangle$ or $|v\rangle \propto |w\rangle$.[2] This is where the distinction between classical and quantum mechanics comes in. In quantum mechanics, a non-separable state is called an **entangled** state.

Linearity and associative and distributive laws hold, eg

$$\big(\alpha|p\rangle\big) \otimes \big(\beta|v\rangle + \gamma|w\rangle\big) = \alpha\beta\big(|p\rangle \otimes |v\rangle\big) + \alpha\gamma\big(|p\rangle \otimes |w\rangle\big)$$

Note $|v\rangle \otimes |0\rangle$ and $|0\rangle \otimes |w\rangle$ are the same and equal to the null vector $^{\mathbb{P}}$.

1.3 Inner Products

Definitions

Shankar pp 7-9, Riley 8.1.2, Griffiths A.2

In applications in physics we usually want to define the length or "norm" of a vector, and the "angle" between two vectors. To be precise, we define the **inner product** of $|v\rangle$ and $|w\rangle$, written $\langle v|w\rangle$, as a complex number that obeys three rules:

(I) $\langle v|w\rangle = \langle w|v\rangle^*$. (Skew symmetry)

(II) $\langle v|v\rangle \geq 0$, with equality if and only if $|v\rangle$ is the zero vector. (Positive definiteness)

(III) $\langle v|\big(\alpha|u\rangle + \beta|w\rangle\big) = \alpha\langle v|u\rangle + \beta\langle v|w\rangle$, where $\alpha, \beta \in \mathbb{C}$. (Linearity on the right or ket side).

A vector space with an inner product is called an **inner-product space**. The term **Hilbert space** is also used; in finite-dimensional spaces at least they are equivalent for our purposes.

Examples

• For real vectors in 3-D the usual scalar product satisfies these rules $^{\mathbb{P}}$.
• So does the "sum of products" rule $\sum_i v_i w_i$ for lists of real numbers ($\mathbb{R}^N$).
• However the "sum of products" rule does NOT work for lists of complex numbers ($\mathbb{C}^N$); but $\sum_i v_i^* w_i$ does.

It follows that for vectors from a complex vector space, if $|p\rangle = \alpha|u\rangle + \beta|v\rangle$,

$$\langle p|w\rangle = \alpha^*\langle u|w\rangle + \beta^*\langle v|w\rangle :$$

i.e. inner products are "*anti*-linear" or "conjugate-linear" on the left $^{\mathbb{P}}$.

Two vectors are **orthogonal** if their inner product is zero: $\langle v|w\rangle = 0 = \langle w|v\rangle$.

We *choose* the **norm** or length of a vector $|v\rangle$ to be $|v| = \sqrt{\langle v|v\rangle}$. If $|v| = 1$, $|v\rangle$ is **normalised**.

[2] $|p\rangle \propto |q\rangle$ means that there is some scalar α such that $|p\rangle = \alpha|q\rangle$

Orthonormal bases

Shankar pp 9-12, 14-15, Riley 8.1.2, Griffiths A.2

A set of vectors in a vector space $\mathbb{V}^N$, $\{|i\rangle\} \equiv \{|1\rangle, |2\rangle, ...|n\rangle\}$, all of unit norm, and all orthogonal to each other, is called an **orthonormal set**. By definition they satisfy $\langle i|j\rangle = \delta_{ij}$ (i.e. 1 if $i = j$ and 0 otherwise).

(We could equally have denoted the basis $\{|x_i\rangle\}$. Especially if we are talking about vectors in real 3D space we might use the notation $\{|e_i\rangle\}$ instead.)

Vectors in an orthonormal set are linearly independent $^{\mathbb{P}}$, so $n \leq N$.

If there are enough vectors in the orthonormal set to make a basis (for finite-dimensional spaces, $n = N$), we call it an **orthonormal basis** or **complete orthonormal set**.

Every [finite-dimensional] vector space has an orthonormal basis $^{\mathbb{P}}$ (actually infinitely many). (This theorem is actually true even for infinite-dimensional vector spaces but the proof is hard.)

Coordinates in an orthonormal basis have very simple expressions: if $|v\rangle = \sum_i v_i|i\rangle$, then $v_i = \langle i|v\rangle$ $^{\mathbb{P}}$.

If v_i and w_i, are the coordinates of $|v\rangle$ and $|w\rangle$ respectively, $\langle v|w\rangle = \sum_i v_i^* w_i$ and $\langle v|v\rangle = \sum_i v_i^* v_i = \sum_i |v_i|^2 \geq 0$ $^{\mathbb{P}}$.

(Remember in proving these, you need to use different indices ("dummy indices") for each sum, and these in turn must be different from any "free" index, which stands for any of $1 \ldots N$. Thus for example $\langle i|v\rangle = \sum_j v_j \langle i|j\rangle$.)

Though coordinates are basis-dependent, the sums that give norms and inner products are basis-independent, as will be shown later.

Gram-Schmidt orthogonalisation can be used to construct an orthonormal basis $\{|i\rangle\}$ from a set $\{|v_i\rangle\}$ of N linearly-independent vectors. First, let $|1\rangle$ be $|v_1\rangle/|v_1|$. Then take $|v_2\rangle$, subtract off the component parallel to $|1\rangle$, and normalise:

$$|2\rangle = C_2\big(|v_2\rangle - \langle 1|v_2\rangle|1\rangle\big) \qquad \text{where } |C_2|^{-2} = \langle v_2|v_2\rangle - \langle v_2|1\rangle\langle 1|v_2\rangle$$

Continue taking the remaining $|v_j\rangle$ in turn, subtract off the component parallel to each previously constructed $|i\rangle$, normalise and call the result $|j\rangle$:

$$|j\rangle = C_j \left(|v_j\rangle - \sum_{i=1}^{j-1} \langle i|v_j\rangle|i\rangle \right)$$

where C_j is the normalisation constant. The resulting basis is not unique, because it depends on the ordering of the basis vectors, which is arbitrary; also the normalisation constants are only defined up to a phase. (This construction proves the existence of an orthonormal basis, as asserted above.)

Bras

Shankar pp 11-14, Griffiths 3.6

In Dirac notation, the inner product $\langle v|w\rangle$ is considered as a **bra** $\langle v|$ acting on ket $|w\rangle$ to form a (scalar) "bracket". Another way of saying this is that a bra $\langle v|$ is an object with the property that it can be combined with any ket $|w\rangle$ from $\mathbb{V}^N$ to give the inner product $\langle v|w\rangle$. For each ket, there is a corresponding bra and vice versa, so $\langle w|v\rangle = \langle v|w\rangle^*$ will be the result of combining the bra $\langle w|$ with the ket $|v\rangle$.

Mathematically, if the ket lives in a vector space $\mathbb{V}^N$, then the bra is an element of another vector space, called the **dual** of $\mathbb{V}^N$, but we will not need this distinction. (Students often stumble over the concept of bras when they first meet them, so the interpretation in terms of row vectors to be given below is a very useful picture.)

Given a basis $\{|i\rangle\}$, the corresponding bras $\{\langle i|\}$ span the space of bras, and an arbitrary bra can be expanded $\langle v| = \sum_i v_i^* \langle i|$, with $\langle v|i\rangle = v_i^*$. Thus the coordinates of the bra $\langle v|$ are v_i^* $^{\mathbb{P}}$. Note that if the ket $|w\rangle = \big(\alpha|u\rangle + \beta|v\rangle\big)$, the corresponding bra is $\langle w| = \big(\alpha^*\langle u| + \beta^*\langle v|\big)$.

If we represent a ket $|v\rangle$ as a column matrix of coordinates $\mathbf{v}$:

$$|v\rangle \rightarrow \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} \equiv \mathbf{v},$$

the corresponding bra is a row matrix:

$$\langle v| \rightarrow (v_1^*, v_2^*, \dots v_N^*) = (\mathbf{v}^\top)^* \equiv \mathbf{v}^\dagger.$$

and the ordinary rules of matrix multiplication make the operation of a bra on a ket give a single complex number:

$$\langle v|w\rangle \rightarrow (v_1^*, \dots v_N^*) \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = \sum_{i=1}^{N} v_i^* w_i$$

just as before.

Note that the basis kets given by $\langle 1| \rightarrow (1, 0, \dots, 0, 0)$ etc.

Inequalities

Shankar pp 16-17, Riley 8.1.3, Griffiths A.2

The Schwarz Inequality: for any vectors $|v\rangle, |w\rangle$, $|\langle v|w\rangle| \le |v|\,|w|$ $^{\mathbb{P}}$.
The equality holds only if $|v\rangle \propto |w\rangle$ $^{\mathbb{P}}$.
Notice that the same rule applies to ordinary dot products, since $|\cos\theta| \le 1$.

The triangle inequality: If $|w\rangle = |u\rangle \pm |v\rangle$, then $|w| \le |u| + |v|$ $^{\mathbb{P}}$.
Notice that this holds for the lengths of ordinary "arrow" vectors that form a triangle! By symmetry, the result is cyclic, i.e. $|v| \le |w| + |u|$ etc.

Inner products in product spaces

Let $\{|p\rangle, |q\rangle\} \in \mathbb{V}_a$ and $\{|v\rangle, |w\rangle\} \in \mathbb{V}_b$, and let an inner product be defined on each space. The inner product in the product space $\mathbb{V}_a \otimes \mathbb{V}_b$ is defined as $\big(\langle p| \otimes \langle v|\big)\big(|q\rangle \otimes |w\rangle\big) = \langle p|q\rangle\langle v|w\rangle$, which of course is a scalar.

If $\{|p_i\rangle\}$ and $\{|v_i\rangle\}$ are orthonormal bases in each space, then $\{|p_i\rangle \otimes |v_j\rangle\}$ is an orthonormal basis in the product space (there are of course others, which need not be separable).

1.4 Operators

Definition

Shankar 18-20, Riley 8.2, 7.2.1, Griffiths A.3

Operators change kets into other kets in the same vector space:

$$\hat{A}|v\rangle = |w\rangle$$

For the moment we mark operators with a hat, $\hat{}$.

Linear operators (we will not consider others) have the property that

$$\hat{A}\left(\alpha|v\rangle + \beta|w\rangle\right) = \alpha\hat{A}|v\rangle + \beta\hat{A}|w\rangle \qquad \text{and} \qquad \left(\alpha\hat{A} + \beta\hat{B}\right)|v\rangle = \alpha\hat{A}|v\rangle + \beta\hat{B}|v\rangle.$$

Hence any operator acting on the zero vector gives zero.
The **identity operator** $\hat{I}$ leaves a ket unchanged: $\hat{I}|v\rangle = |v\rangle$.

The **product of two operators**, say $\hat{A}\hat{B}$, means "apply $\hat{B}$ first and then apply $\hat{A}$ to the result". If $\hat{B}|v\rangle = |u\rangle$, $\hat{A}\hat{B}|v\rangle = \hat{A}|u\rangle$.

$\hat{A}$ and $\hat{B}$ will not in general commute, in which case this is not the same as $\hat{B}\hat{A}|v\rangle$.

If, for all kets in the space, $\hat{B}\hat{A}|v\rangle = |v\rangle$, then $\hat{B}$ is called the inverse of $\hat{A}$ and denoted $\hat{A}^{-1}$. We can write $\hat{A}^{-1}\hat{A} = \hat{I}$. For finite dimensional spaces, $\hat{A}\hat{A}^{-1} = \hat{I}$ also $^{\mathbb{P}}$.

Not all operators have inverses. However if the equation $\hat{A}|v\rangle = |0\rangle$ has no solutions except $|v\rangle = |0\rangle$, the inverse $\hat{A}^{-1}$ does exist $^{\mathbb{P}}$.

Inverse of matrix products: if $\hat{C} = \hat{A}\hat{B}$, then $\hat{C}^{-1} = \hat{B}^{-1}\hat{A}^{-1}\,^{\mathbb{P}}$.

Identity and Projection operators

Shankar 22-24, (Riley 8.4), Griffiths 3.6

The object $|a\rangle\langle b|$ is in fact an operator since, acting on any ket $|v\rangle$, it gives another ket, $\langle b|v\rangle|a\rangle$. (Whatever $|v\rangle$ we choose, the resulting ket is always proportional to $|a\rangle$.) This is termed the **outer product** of $|a\rangle$ and $|b\rangle$, and is entirely distinct from the inner product $\langle b|a\rangle$, which is a scalar.

Using an orthonormal basis $\{|i\rangle\}$, we can define **projection operators**, $\hat{P}_i = |i\rangle\langle i|$, which "pull out" only the part of a vector $|v\rangle$ which is parallel to $|i\rangle$: $\hat{P}_i|v\rangle = v_i|i\rangle$. The product of two projection operators is zero or equivalent to a single projection $^{\mathbb{P}}$: $\hat{P}_i\hat{P}_j = \delta_{ij}\hat{P}_i$.

These are examples of operators which do not have an inverse, since $\hat{P}_i|v\rangle = 0$ will be satisfied for many non-zero kets $|v\rangle$. The lack of an inverse reflects the fact that when we operate with $\hat{P}_i$ on a vector, we lose all information about components orthogonal to $|i\rangle$, and no operator can restore it.

One very useful way of writing the identity operator is as follows $^{\mathbb{P}}$:

$$\hat{I} = \sum_i \hat{P}_i = \sum_i |i\rangle\langle i|$$

This is called the **completeness relation**. The sum *must* be over projectors onto an *orthonormal* basis.

Matrix representation of operators

Shankar 20-22, 25, Riley 8.3, 7.3.1, Griffiths A.3

[*Comment on notation in Riley:* Riley uses boldface for abstract vectors where we use kets, and calligraphic letters without "hats" for operators: hence $\hat{A}|v\rangle = |u\rangle$ is written $\mathcal{A}\mathbf{v} = \mathbf{u}$. We use boldface for column vectors and matrices of components, but Riley uses a sans-serif font, so $\mathsf{A}\,\mathsf{v} = \mathsf{u}$ is a matrix equation.]

We can form the inner product of $|u\rangle \equiv \hat{A}|v\rangle$ with another vector $|w\rangle$, to get $\langle w|u\rangle = \langle w|\big(\hat{A}|v\rangle\big)$. This is called a **matrix element** of $\hat{A}$, and is more often written $\langle w|\hat{A}|v\rangle$.

If we have an orthonormal basis $\{|i\rangle\}$, we can form all possible matrix elements of $\hat{A}$ between vectors of the basis, $A_{ij} = \langle i|\hat{A}|j\rangle$; these are the coordinates of $\hat{A}$ in this basis. Then $^{\mathbb{P}}$

$$\hat{A}|v\rangle = \sum_{ij} A_{ij} v_j |i\rangle \qquad \text{and} \qquad \langle w|\hat{A}|v\rangle = \sum_{ij} w_i^* A_{ij} v_j.$$

The numbers A_{ij} can be arranged in a matrix $\mathbf{A}$, i labelling the row and j the column, which gives

$$\langle w|\hat{A}|v\rangle = (w_1^*, w_2^*, \dots w_N^*) \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \mathbf{w}^\dagger \mathbf{A} \mathbf{v} \qquad (1.1)$$

The ith column of matrix $\mathbf{A}$ is just the coordinates of $|Ai\rangle \equiv \hat{A}|i\rangle$, i.e. the transformed basis ket.

If the **determinant** of $\mathbf{A}$ vanishes, its columns are not linearly independent. That means that $\{|Ai\rangle\}$ is not a basis, and the vectors $|Av\rangle$ belong to a lower-dimensional sub-space of $\mathbb{V}^N$. Hence $\det \mathbf{A} = 0$ means that $\hat{A}^{-1}$ does not exist.

The matrix elements of the product of two operators can be found by inserting the completeness relation $\sum_k |k\rangle\langle k|$ as an identity operator in $\hat{A}\hat{B} = \hat{A}\hat{I}\hat{B}$:

$$(AB)_{ij} = \langle i|\hat{A}\hat{B}|j\rangle = \sum_k \langle i|\hat{A}|k\rangle\langle k|\hat{B}|j\rangle = \sum_k A_{ik} B_{kj}$$

i.e. the usual matrix multiplication formula.

Examples:

Identity: $I_{ij} = \langle i|\hat{I}|j\rangle = \langle i|j\rangle = \delta_{ij}$. So

$$\hat{I} \to \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \end{pmatrix}.$$

Projectors: $\langle i|\hat{P}_k|j\rangle = \langle i|k\rangle\langle k|j\rangle = \delta_{ik}\delta_{jk} = \delta_{ij}\delta_{ik}$ (note we do not use a summation convention), eg

$$\hat{P}_3 \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix}$$

i.e. 1 on the diagonal for the selected row/column

An outer product: The matrix elements of $\hat{C} = |v\rangle\langle w|$ are just $c_{ij} = v_i w_j^*$. We can obtain a square matrix from a column and a row vector if we multiply them in that order (as opposed to the opposite order which gives the inner product, a scalar):

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} (w_1^*, w_2^*, \dots w_N^*) = \begin{pmatrix} v_1 w_1^* & v_1 w_2^* & \dots & v_1 w_N^* \\ v_2 w_1^* & v_2 w_2^* & & \vdots \\ \vdots & & \ddots & \\ v_N w_1^* & \dots & & v_N w_N^* \end{pmatrix}$$

Adjoints

Shankar pp 25-27, (Riley 8.6), Griffiths A.3, A.6

An operator such as $|a\rangle\langle b|$ can clearly act on bras as well as kets: $\langle u| \big(|a\rangle\langle b|\big) = \big(\langle u|a\rangle\big) \langle b|$.

In fact all operators can act to the left on bras as well as to the right on kets. This is obvious from the matrix representation in an orthonormal basis, since a row vector can be multiplied from the right by a matrix.

Now the ket $|u\rangle = \hat{A}|v\rangle$ has a bra equivalent, $\langle u|$, but for most operators it is not the same as $\langle p| = \langle v|\hat{A}$. We define the **adjoint** of $\hat{A}$, $\hat{A}^\dagger$, as the operator that, acting to the left, gives the bra corresponding to the ket which results from $\hat{A}$, acting to the right: $\langle u| = \langle v|\hat{A}^\dagger$. Hence $\langle w|\hat{A}|v\rangle = \langle v|\hat{A}^\dagger|w\rangle^*$.

$\hat{A}^\dagger$ has matrix elements $A^\dagger_{ij} = A^*_{ji}$ $^{\mathbb{P}}$ i.e. the matrix representation of the adjoint operator is the **transposed complex conjugate** of the original matrix, also called the **Hermitian conjugate**. It follows that $(\hat{A}^\dagger)^\dagger = \hat{A}$, i.e. the adjoint of the adjoint is the original.

Adjoints of products: $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$.

Adjoints of scalars: if $\hat{B} = c\hat{A}$, $\hat{B}^\dagger = c^*\hat{A}^\dagger$. Complex numbers go to their complex conjugates in the adjoint.

The adjoint of $|a\rangle\langle b|$ is $|b\rangle\langle a|$ $^{\mathbb{P}}$.

Operators in product spaces

Let $\hat{C}_a$ be an operator in a vector space $\mathbb{V}_a$ and $\hat{D}_b$ one in $\mathbb{V}_b$. Then in the product space $\mathbb{V}_a \otimes \mathbb{V}_b$ we can form product operators $\hat{C}_a \otimes \hat{D}_b$, which act on the kets as follows:

$$\big(\hat{C}_a \otimes \hat{D}_b\big)\big(|p\rangle \otimes |v\rangle\big) = \big(\hat{C}_a|p\rangle\big) \otimes \big(\hat{D}_b|v\rangle\big).$$

Here it is particularly important to be clear that we are not multiplying $\hat{C}_a$ and $\hat{D}_b$ together; they act in different spaces. Once again $\otimes$ should be regarded as a separator, not a multiplication.

Denoting the identity operators in each space as $\hat{I}_a$ and $\hat{I}_b$ respectively, in the product space the identity operator is $\hat{I}_a \otimes \hat{I}_b$. An operator in which each additive term acts in only one space, such as $\hat{C}_a \otimes \hat{I}_b + \hat{I}_a \otimes \hat{D}_b$, is called a **separable operator**. $\hat{C}_a \otimes \hat{I}_b$ and $I_a \otimes \hat{D}_b$ commute.

The inverse of $\hat{C}_a \otimes \hat{D}_b$ is $\hat{C}_a^{-1} \otimes \hat{D}_b^{-1}$ and the adjoint, $\hat{C}_a^\dagger \otimes \hat{D}_b^\dagger$. (The order is NOT reversed, since each still has to act in the correct space.)

Matrix elements work as follows: $(\langle p| \otimes \langle v|)\left(\hat{C}_a \otimes \hat{D}_b\right)(|q\rangle \otimes |w\rangle) = \langle p|\hat{C}_a|q\rangle\langle v|\hat{D}_b|w\rangle$. (This is the arithmetic product of two scalars.)

The labels $_a$ and $_b$ are redundant since the order of the operators in the product tells us which acts in which space. Alternatively if we keep the labels, it is common to write $\hat{C}_a$ when we mean $\hat{C}_a \otimes \hat{I}_b$ and $\hat{C}_a\,\hat{D}_b$ (or even, since they commute, $\hat{D}_b\,\hat{C}_a$) when we mean $\hat{C}_a \otimes \hat{D}_b$.

1.5 Hermitian and Unitary operators

Definition and Properties of Hermitian operators

Shankar p 27, Riley 8.12.5, Griffiths A.3

An operator $\hat{H}$ is **Hermitian** if $\hat{H}^\dagger = \hat{H}$ or **anti-Hermitian** if $\hat{G}^\dagger = -\hat{G}$. Another term for Hermitian is **self-adjoint**.

In real spaces Hermitian operators are represented by symmetric matrices, $\mathbf{H}^\top = \mathbf{H}$.

For Hermitian operators, if $|u\rangle = \hat{H}|v\rangle$ and $|z\rangle = \hat{H}|w\rangle$, then $\langle z| = \langle w|\hat{H}$, and $\langle w|\hat{H}|v\rangle = \langle w|u\rangle = \langle z|v\rangle\,^{\mathbb{P}}$.
It follows that $\langle v|\hat{H}|w\rangle = \langle w|\hat{H}|v\rangle^*$ and $\langle v|\hat{H}^2|v\rangle \geq 0\,^{\mathbb{P}}$.

Definition and Properties of Unitary operators

Shankar pp 28-29, Riley 8.12.6, Griffiths A.3

An operator $\hat{U}$ is **unitary** if $\hat{U}^\dagger = \hat{U}^{-1}$. (In infinite dimensional spaces $\hat{U}\hat{U}^\dagger = \hat{I}$ *and* $\hat{U}^\dagger\hat{U} = \hat{I}$ must both be checked.)

In real spaces unitary operators are represented by orthogonal matrices, $\mathbf{U}^\top = \mathbf{U}^{-1}$.

Unitary operators preserve the inner product, i.e. if $\hat{U}|v\rangle = |v'\rangle$ and $\hat{U}|w\rangle = |w'\rangle$, then $\langle v|w\rangle = \langle v'|w'\rangle\,^{\mathbb{P}}$. (The use of a "prime", $'$, just creates a new label. It has nothing to do with differentiation!)

The columns of a unitary matrix are orthonormal vectors, as are the rows $^{\mathbb{P}}$.

Since the matrix contains N columns (or rows), where N is the dimension of the vector space, these orthonormal sets are actually complete bases.

The converse is also true: any matrix whose columns (or rows) form orthonormal vectors is guaranteed to be unitary.

The determinant of a unitary matrix is a complex number of unit modulus $^{\mathbb{P}}$.

Unitary transformations: Change of basis

Shankar pp 29-30, Riley 8.15, Griffiths A.4

Let us define two orthonormal bases in $\mathbb{V}^N$, $\{|x_i\rangle\}$ and $\{|y_i\rangle\}$. We will label components in these bases by superscripts (x) and (y), eg $v_i^{(x)} = \langle x_i|v\rangle$, $A_{ij}^{(y)} = \langle y_i|\hat{A}|y_j\rangle$.

The components in the two bases are related by the matrix $\mathbf{S}$, where $S_{ij} = \langle x_i|y_j\rangle$ (and $(\mathbf{S}^\dagger)_{ij} = \langle y_i|x_j\rangle$) as follows $^{\mathbb{P}}$:

$$v_i^{(y)} = \sum_j S_{ji}^* v_j^{(x)} \Rightarrow \mathbf{v}^{(y)} = \mathbf{S}^\dagger \mathbf{v}^{(x)}; \qquad A_{ij}^{(y)} = S_{ki}^* A_{kl} S_{lj} \Rightarrow \mathbf{A}^{(y)} = \mathbf{S}^\dagger \mathbf{A}^{(x)} \mathbf{S}.$$

A simple example of a change of basis in a two-dimensional space is given by $|y_1\rangle = \cos\theta|x_1\rangle + \sin\theta|x_2\rangle$ and $|y_2\rangle = \cos\theta|x_2\rangle - \sin\theta|x_1\rangle$. Then $\mathbf{S} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

We often use $\{|i\rangle\}$ and $\{|i'\rangle\}$ for the two bases, with $S_{ij} = \langle i|j'\rangle$, $v_i = \langle i|v\rangle$ and $v_i' = \langle i'|v\rangle$.

$\mathbf{S}$ is a unitary matrix: $(\mathbf{S}^\dagger\mathbf{S})_{ij} = \sum_k \langle y_i|x_k\rangle\langle x_k|y_j\rangle = \delta_{ij}$. Hence (as we already knew) inner products ($\langle v|w\rangle$) and matrix elements ($\langle v|\hat{A}|w\rangle$) are independent of coordinate system, even if the individual numbers we sum to get them are different.

In addition, $\mathrm{Tr}(\mathbf{A}^{(x)}) = \mathrm{Tr}(\mathbf{A}^{(y)})$ and $\det(\mathbf{A}^{(x)}) = \det(\mathbf{A}^{(y)})$, so these also are basis-independent $^{\mathbb{P}}$. For that reason we can assign these properties to the operators and talk about about $\mathrm{Tr}(\hat{A})$ and $\det(\hat{A})$.[3]

The reverse transformation, from y-basis to x-basis, is done by interchanging $\mathbf{S}^\dagger$ and $\mathbf{S}$.

Note that $\mathbf{A}^{(x)}$ and $\mathbf{A}^{(y)}$ are representations of the *same* abstract operator $\hat{A}$ in different bases (similarly $\mathbf{v}^{(x)}$, $\mathbf{v}^{(y)}$ of the abstract ket $|v\rangle$). Therefore, $\mathbf{S}$ is not an operator, since it does not change the abstract kets. We call this a **passive transformation** or **coordinate change**.

However there are also unitary operators which do change the kets. An example is a rotation of a vector in ordinary 3D (real) space (an **active transformation**), which is represented by the transpose of the (orthogonal) matrix which transforms between rotated coordinate systems.

1.6 Eigenvectors and Eigenvalues

Note that from now on, we will write the zero vector as 0. We may even use $|0\rangle$ for a non-zero vector with label 0!

Basic properties

Shankar pp 30-35, Riley 8.13, Griffiths A.5

The **eigenvalue equation** for a linear operator $\hat{\Omega}$ is

$$\hat{\Omega}|\omega\rangle = \omega|\omega\rangle.$$

The equation is solved by finding both the allowed values of the scalar number ω, the **eigenvalue**, and for each eigenvalue the corresponding ket $|\omega\rangle$, the **eigenvector** or **eigenket**.

[3] The trace of a matrix $\mathbf{A}$ is the sum of the diagonal elements: $\mathrm{Tr}(\mathbf{A}) = \sum_i A_{ii}$.

The German word "eigen" means "own" or "characteristic"— i.e. the eigenkets are a special set of vectors for each particular operator which have a very simple behaviour when operated on: no change in "direction", just a multiplication by a scalar eigenvalue. As we have done above, we habitually use the eigenvalue ("ω") to label the corresponding eigenket ("$|\omega\rangle$").

The zero vector does not count as an eigenvector.

If $|\omega\rangle$ is a solution to the eigenvalue equation, so is $\alpha|\omega\rangle$ for any $\alpha \neq 0$. All such multiples are considered to be a single eigenvector, and we usually quote the normalized value, with real elements if that is possible.

We can rewrite the eigenvalue equation as $(\hat{\Omega}-\omega\hat{I})|\omega\rangle = 0$. (We can insert the identity operator at will as it does nothing. The final zero is of course the zero *vector*.)

This is an equation that we want to solve for a non-zero $|\omega\rangle$, so $(\hat{\Omega}-\omega\hat{I})$ cannot have an inverse, and its determinant must vanish. This is the the **characteristic equation**:

$$\det(\hat{\Omega} - \omega\hat{I}) = 0.$$

In any basis this is the determinant of an $N \times N$ matrix, which is an Nth-order polynomial in ω. The fundamental theorem of algebra states that such a polynomial has N roots $\omega_1, \omega_2 \ldots \omega_N$, where some roots may be repeated and roots may be complex even if the coefficients are real. Therefore any operator on $\mathbb{V}^N$ has N eigenvalues, not necessarily all different.

The sum of all eigenvalues of $\hat{\Omega}$ (including repeated ones) is $\mathrm{Tr}(\hat{\Omega})$, and their product equals $\det(\hat{A})$ $^{\mathbb{P}}$. Thus if $\hat{\Omega}$ has any zero eigenvalues, its inverse does not exist.

For each non-repeated eigenvalue ω_i we will call the corresponding eigenvector $|\omega_i\rangle$ Working in an orthonormal basis, the equation $(\hat{\Omega} - \omega_i\hat{I})|\omega_i\rangle = 0$ will give $N-1$ linearly-independent equations for the components of $|\omega_i\rangle$, so—as we knew—we can determine $|\omega_i\rangle$ only up to a multiplicative constant.

A set of eigenvectors corresponding to distinct eigenvalues is linearly independent. $^{\mathbb{P}}$

For an eigenvalue which is repeated n times, there will be at least $N-n$ linearly-independent equations. These will have up to n linearly-independent solutions. Thus an operator with repeated eigenvalues will have up to N linearly-independent eigenvectors.

Hermitian and unitary operators

Shankar pp 35-40, Riley 8.13.2 & 18.13.3, 7.12.3, Griffiths A.6

Important results $^{\mathbb{P}}$:

I) For Hermitian operators, eigenvalues are real.

II) For unitary operators, eigenvalues have unit modulus, i.e. they can be written $e^{i\theta}$, $\theta \in \mathbb{R}$.

III) For both Hermitian and unitary operators, eigenkets with different eigenvalues are orthogonal.

IV) For all Hermitian and unitary operators, the eigenvectors span the space. (The general proof of this one is more involved, but it follows from (III) if all the eigenvalues are distinct). This is called the **Spectral Theorem**.

Suppose a Hermitian or unitary operator $\hat{\Omega}$ has a repeated eigenvalue, say $\omega_1 = \omega_2 = \ldots = \omega_n = \lambda$. By the spectral theorem there are n linearly-independent solutions $|\lambda, m\rangle$ (where $m = 1 \ldots n$

is just a label here). These eigenvectors are said to be **degenerate** (same eigenvector). Then any linear combination $\sum_{m=1}^{n} c_m|\lambda, m\rangle$ is also an eigenvector. Therefore any vector in the subspace spanned by the set $\{|\lambda, m\rangle\}$ is an eigenvector of $\hat{\Omega}$. We call this an **eigenspace**. Even if the first set of degenerate eigenvectors we found was not orthogonal, a new orthogonal basis in the sub-space can always be found (by the Gram-Schmidt method or otherwise). Thus we can always find a set of N orthonormal eigenvectors of $\hat{\Omega}$.

Any Hermitian or unitary operator can be written in terms of this orthonormal basis as

$$\hat{\Omega} = \sum_{i,m} \omega_i |w_i, m\rangle\langle w_i, m|.$$

This is called the **spectral resolution** of $\hat{\Omega}$. The first sum is over distinct eigenvalues. The second sum runs over all the states within each eigenspace; for non-degenerate eigenvalues it is not needed. We will not always write it explicitly, often just referring to the set of N vectors $\{|\omega_i\rangle\}$, but if degeneracy is present an orthogonalised basis is always meant.

Diagonalisation of Hermitian or unitary operators

Shankar pp 40-43, Riley 8.16, Griffiths A.5

To convert from some orthonormal basis $\{|x_i\rangle\}$ to the eigenvector basis $\{|\omega_i\rangle\}$ in which $\hat{\Omega}$ is diagonal, we need the unitary conversion matrix $S_{ij} = \langle x_j|\omega_i\rangle$. The columns of $\mathbf{S}$ are the eigenvectors of $\mathbf{\Omega}$ in the original basis, hence it is sometimes called the **matrix of eigenvectors**.

Using this matrix to change coordinates we get:

$$\mathbf{v}^{(\omega)} = \mathbf{S}^\dagger \mathbf{v}^{(x)}, \qquad \mathbf{\Omega}^{(\omega)} = \mathbf{S}^\dagger \mathbf{\Omega}^{(x)} \mathbf{S},$$

where superscripts in braces indicate the basis in which $|v\rangle$ and $\hat{\Omega}$ are represented.

However we do not need to perform the operation to know what we will get for $\mathbf{\Omega}^{(\omega)}$:

$$\hat{\Omega} \underset{\omega}{\longrightarrow} \begin{pmatrix} \omega_1 & & & \\ & \omega_2 & & \\ & & \ddots & \\ & & & \omega_N \end{pmatrix}$$

(all the off-diagonal elements being zero). The order is arbitrary of course, though we often choose ascending order (since they are, of course, real).

Commuting Hermitian Operators

Shankar pp 43-46, Riley 8.13.5

If the commutator $[\hat{\Omega}, \hat{\Lambda}] = 0$ (where $\hat{\Omega}$ and $\hat{\Lambda}$ are Hermitian), there is at least one basis of common eigenvectors (therefore both operators are represented by diagonal matrices in this basis).

Proof outline: by considering $[\hat{\Omega}, \hat{\Lambda}]|\omega_i\rangle = 0$ we can immediately see that $\hat{\Lambda}|\omega_i\rangle$ is also an eigenvector of $\hat{\Omega}$ with eigenvalue ω_i. In the absence of degeneracy, that can only be the case if $\hat{\Lambda}|\omega_i\rangle$ is proportional to $|\omega_i\rangle$, so the non-degenerate eigenstates of $\hat{\Omega}$ are also those of $\hat{\Lambda}$. If there is degeneracy, though, $\hat{\Lambda}|\omega_i\rangle$ only needs to be another state in the same n-dimensional eigenspace of $\hat{\Omega}$. However we know we can find n orthogonal eigenvectors of $\hat{\Lambda}$ within that subspace (i.e. we can diagonalise $\hat{\Lambda}$ within that subspace) and the resulting eigenvectors of $\hat{\Lambda}$ are an equally valid basis of degenerate eigenstates of $\hat{\Omega}$. We can now label the states $|\omega_i, \lambda_j\rangle$, and λ_j is no longer just an arbitrary label.

There may still be states that have the same ω_i and the same λ_i, but we can repeat with further commuting operators until we have a **complete set of commuting operators** defining a *unique* orthonormal basis, in which each basis ket can be labelled unambiguously by the eigenvalues $|\omega, \lambda, \gamma, \ldots\rangle$ of the operators $\{\hat{\Omega}, \hat{\Lambda}, \hat{\Gamma}, \ldots\}$.

Examples of commuting operators are those in a product space of the form $\hat{C}_a \otimes \hat{I}_b$ and $\hat{I}_a \otimes \hat{D}_b$. If an operator is separable, i.e. it can be written as $\hat{C}_a \otimes \hat{I}_b + \hat{I}_a \otimes \hat{D}_b$, then the eigenvectors are $|c_i\rangle \otimes |d_j\rangle$ with eigenvalue $c_i + d_j$. As already mentioned the operator is often written $\hat{C}_a + \hat{D}_b$, where the label makes clear which space each operator acts in; similarly the eigenstates are often written $|c_i, d_j\rangle$.

1.7 Functions of Operators

Shankar pp 54-57, Riley 8.5, Griffiths A.6

We can add operators, multiply them by scalars, and take products of them. Hence we can define a power series

$$f(\hat{\Omega}) = \sum_{n=0}^{\infty} a_n \hat{\Omega}^n.$$

This will make sense if it converges to a definite limit. In its eigenbasis a Hermitian operator is diagonal, so the power series acts on each diagonal element separately:

$$f(\hat{\Omega}) \underset{\omega}{\longrightarrow} \begin{pmatrix} f(\omega_1) & & & \\ & f(\omega_2) & & \\ & & \ddots & \\ & & & f(\omega_N) \end{pmatrix}$$

i.e. the power series converges for the operator if it converges for all its eigenvalues, and the eigenvalues of $f(\hat{\Omega})$ are just the corresponding functions of the eigenvalues of $\hat{\Omega}$.

A very important operator function is the exponential, which is *defined* though the power series

$$\mathrm{e}^{\hat{\Omega}} \equiv \sum_{n=0}^{\infty} \frac{\hat{\Omega}^n}{n!}.$$

Since the corresponding power series for e^{ω} converges for all finite numbers, this is defined for all Hermitian operators, and its eigenvalues are e^{ω_i}.

From the definition it is clear that if $\hat{\Omega}$ and $\hat{\Lambda}$ do not commute, $\mathrm{e}^{\hat{\Omega}}\mathrm{e}^{\hat{\Lambda}} \neq \mathrm{e}^{\hat{\Omega}+\hat{\Lambda}}$.

Acknowledgements

This section is based quite closely on Chapter 1 of Shankar, and owes a considerable debt to the notes prepared by Dr J P Leahy for a precursor to this course. Any mistakes however are

mine.

1.8 Summary

- A real or complex **vector space** is a **set** of **abstract vectors**, written as **kets** (e.g. $|v\rangle$), which is **closed** under both addition and multiplication by scalar real or complex numbers: all vectors you can reach by any combination of addition and scalar multiplication are **elements** of the vector space. There must be a **zero vector** $|0\rangle$ (or often, just 0) and vector have inverses: $|v\rangle + |-v\rangle = |0\rangle$.

- **Linearly-independent sets of vectors** are sets in which no member can be written as a linear sum of the others. A **basis** is a set of linearly-independent vectors big enough to allow any vector in the space to be written as a sum over the basis vectors. All bases have the same size, which is the **dimension** of the space.

- The **coordinates** of an arbitrary vector in a given basis are the factors that multiply each basis vector $|i\rangle$ in the linear sum: $|v\rangle = \sum v_i|i\rangle$. The column vector of these coordinate is the **representation** of $|v\rangle$ in this basis. The representation depends on the basis.

- In some vector spaces there exists an **inner product** of two vectors, $\langle v|w\rangle$, which give us **orthogonality**, the **norm** of each vector, and hence allows us to construct **orthonormal bases**.

- In an orthonormal basis, coordinates are given by $v_i = \langle i|v\rangle$, and from coordinates we can evaluate inner products $\langle v|w\rangle = \sum_i v_i^* w_i$ and norms of arbitrary vectors.

- We can think of the left side of inner products as **bras**, $\langle a|$, represented by row matrices if kets are column matrices (with elements that are complex conjugates, v_i^*). Inner products are then given by ordinary matrix multiplication.

- **Direct tensor product** spaces are composite spaces in which kets are obtained by taking a ket from each of two separate spaces: $|p\rangle \otimes |v\rangle$ (or taking sums of such terms). Inner products are taken in each space separately: $\big(\langle p|\otimes\langle v|\big)\big(|q\rangle\otimes|w\rangle\big) = \langle p|q\rangle\langle v|w\rangle$. A basis of the product space can be formed by taking all possible combinations of basis vectors from each subspace—$M \times N$ for the product of an M and an N-dimensional space.

- **Linear operators** change kets to kets: $\hat{A}|u\rangle = |v\rangle$, or bras to bras: $\langle u|\hat{A} = \langle w|$.

- The **adjoint** operator $\hat{A}^\dagger$ is defined by $\langle u|\hat{A}^\dagger = \langle v|$. For any $|v\rangle$ and $|x\rangle$, we have $\langle v|\hat{A}^\dagger|x\rangle = \langle x|\hat{A}|v\rangle^*$

- Operators can be multiplied: $\hat{A}\hat{B}$ means "do B then A". They may not commute.

- They may have **inverses**: $\hat{A}\hat{A}^{-1} = \hat{I} = \hat{A}^{-1}\hat{A}$.

- $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$; $(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}$

- In an orthonormal basis $\{|i\rangle\}$, $\hat{I} = \sum_i^N |i\rangle\langle i|$; this is the **completeness relation**.

- Operators in a product space have the form $\hat{A}\otimes\hat{P}$ (or sums of such terms) with $\big(\hat{A}\otimes\hat{P}\big)\big(|a\rangle\otimes|v\rangle\big) - \big(\hat{A}|a\rangle\big)\otimes\big(\hat{P}|v\rangle\big)$.

- Operators in N-dimensional vector spaces can be represented as $N \times N$ matrices.

- Operator product and inverses correspond to matrix products and inverses. The adjoint is the transposed complex conjugate matrix or **Hermitian conjugate**.
- A **Hermitian** operator satisfies $\hat{A} = \hat{A}^\dagger$ ('Self-Adjoint') and $\langle w|\hat{A}|v\rangle = \langle v|\hat{A}|w\rangle^*$.
- A **unitary** operator satisfies $\hat{U}^{-1} = \hat{U}^\dagger$; like a rotation or change of coordinates
- **Eigenvectors** (eigenkets) and **eigenvalues** satisfy $\hat{A}|a_i\rangle = a_i|a_i\rangle$.
- Eigenvectors of Hermitian and unitary operators can form an orthonormal basis (eigenbasis).
- Hermitian operators are **diagonal** in their eigenbasis $\{|\omega_i\rangle\}$, the diagonal elements are the eigenvalues and $\hat{\Omega} = \omega_i \sum_i^N |\omega_i\rangle\langle\omega_i|$.
- Given a **complete sets of commuting Hermitian operators**: each such set defines a unique eigenbasis, with each vector uniquely labelled by its eigenvalues for the operators in the set.
- **Functions of operators** are defined through power series; for Hermitian (or unitary) operators, diagonalize the matrix and apply function to each diagonal element (eigenvalue).