Hilbert Space Primer

(Evolving notes. Don't print!)

Hilbert spaces in Dirac notation

Let's recall some basic facts from linear algebra.

A Hilbert space $\mathcal{H}$ is a complex vector space with a *sesquilinear* inner product. Sesquilinearity means that for all

$$\alpha, \beta, \gamma \in \mathcal{H}, \quad z \in \mathbb{C}$$

we have

$$\langle \alpha | \beta + z\gamma \rangle = \langle \alpha | \beta \rangle + z \langle \alpha | \gamma \rangle,$$

as well as

$$\langle \alpha | \beta \rangle = \overline{\langle \beta | \alpha \rangle}.$$

From this, it follows that

$$\langle \alpha + z\beta | \gamma \rangle = \langle \alpha | \gamma \rangle + \bar{z} \langle \beta | \gamma \rangle,$$

i.e. the inner product is anti-linear w.r.t. the first entry and linear w.r.t. the second one.

> Beware that mathematicians usually employ the opposite convention, where the sesquilinear inner product is linear in the first entry!

The *norm* of a vector is the square root of its inner product with itself:

$$\|\alpha\|_2 = \langle \alpha | \alpha \rangle^{1/2}.$$

Physicists often use notational aids to delinate vector-valued quantities from scalars. For example, the notations $\underline{x}$ for a set of coordinates in classical mechanics, or $\vec{E}$ for the electric field are common. In quantum mechanics, the suggestive *Dirac* notation (or "bra-ket" notation) is usually employed. Here, a vector $\alpha \in \mathcal{H}$ is written as $|\alpha\rangle$. This is called a *ket*, for reasons that will be obvious momentarily. With every *ket* $|\alpha\rangle$, we associate a *bra* $\langle \alpha|$. Bras are *dual vectors*, i.e. linear functions $\mathcal{H} \to \mathbb{C}$ defined by

$$\langle \alpha | : |\beta\rangle \mapsto \langle \alpha | \beta \rangle.$$

The genius of this notation is that one doesn't need to expend any thoughts on concepts like "dual vectors" or "linear functionals". The formalism almost forces one to use the mathematical object correctly: When a bra $\langle \alpha|$ meets a ket $|\beta\rangle$, they form a – you guessed it – braket $\langle \alpha | \beta \rangle$. We'll see more examples of the Dirac notation in action below.

Let $\{|e_i\rangle\}_{i=1}^d$ be an ortho-normal basis (ONB) of $\mathcal{H}$:

$$\langle e_i | e_j \rangle = \delta_{i,j}.$$

Then any vector $\alpha \in \mathcal{H}$ can be expanded as

$$|\alpha\rangle = \sum_{i=1}^{d} \alpha_i |e_i\rangle,$$

with coefficients

$$\alpha_i = \langle e_i|\alpha\rangle.$$

Once we have agreed on a basis, we can identify the Hilbert space $\mathcal{H}$ with $\mathbb{C}^d$, the set of complex column vectors of size d by mapping an element $|\alpha\rangle \in \mathcal{H}$ to the vector of its expansion coefficients:

$$|\alpha\rangle \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}.$$

Such column vectors are particularly useful for computer calculations, as we will see. Bras, in turn, are mapped to *row* vectors with conjugate coefficients: If

$$|\beta\rangle = \sum_{i=1}^{d} \beta_i |e_i\rangle,$$

then

$$\langle\beta| = \sum_{i=1}^{d} \bar{\beta}_i \langle e_i| \mapsto (\bar{\beta}_1, \dots, \bar{\beta}_d).$$

The usual rules of matrix multpiplication then give

$$(\bar{\beta}_1, \dots, \bar{\beta}_d) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = \sum_{i=1}^{d} \bar{\beta}_i \alpha_i = \langle\beta|\alpha\rangle,$$

so the two picture are compatible. In coordinates, the norm reads

$$\||\alpha\rangle\| = \left(\sum_i |\alpha_i|^2 \right)^{1/2}.$$

The “ket”-notation allows one to save a bit of ink when working with one fixed ONB. Say we have agreed to work with $\{|e_i\rangle\}_i$. Then quantum physicists (and no-one else…) commonly drop the symbol e and just put the index into the ket:

$$|i\rangle := |e_i\rangle.$$

This shorthand notation can sometimes lead to confusion, but is too common and convenient to ignore.

Having acquainted ourselves with vectors, let’s move to *operators* or *linear maps* on $\mathcal{H}$. An operator A is a linear function $\mathcal{H} \to \mathcal{H}$, i.e. it satisfies

$$A(|\alpha\rangle + z|\beta\rangle) = A|\alpha\rangle + zA|\beta\rangle.$$

An important class of operators are *outer products* or “ket-bra’s”. Indeed, if $|\alpha\rangle, |\beta\rangle$ are vectors in $\mathcal{H}$, then we define

$$|\alpha\rangle\langle\beta|$$

as the operator that sends $|\gamma\rangle \in \mathcal{H}$ to

$$|\alpha\rangle\langle\beta|\gamma\rangle = \underbrace{|\alpha\rangle}_{\in\mathcal{H}} \underbrace{\big(\langle\beta|\gamma\rangle\big)}_{\in\mathbb{C}} \in \mathcal{H}.$$

The definition thus implies the "associativity" rule

$$(|\alpha\rangle\langle\beta|)\,|\gamma\rangle = |\alpha\rangle(\langle\beta|\gamma\rangle).$$

This associativity will be used extensively when doing calculations in the brak-ket notation. For example, it gives the following important *completeness relation*: If $\{|i\rangle\}_{i=1}^{d}$ is an ONB, then

$$\sum_{i=1}^{d} |i\rangle\langle i| = \mathbb{1},$$

the identity matrix. To prove it, calculate for an arbitrary $|\alpha\rangle = \sum_i \alpha_i |i\rangle$,

$$\left(\sum_i |i\rangle\langle i|\right)|\alpha\rangle = \left(\sum_i |i\rangle\langle i|\right)\left(\sum_j \alpha_j |j\rangle\right) = \sum_{i,j} a_j |i\rangle \underbrace{\langle i|j\rangle}_{\delta_{i,j}} = \sum_i a_i |i\rangle = |\alpha\rangle.$$

Sandwiching an operator A between two completeness relation (and using associativity) gives the important decomposition

$$\left(\sum_i |i\rangle\langle i|\right) A \left(\sum_j |j\rangle\langle j|\right) = \sum_{i,j} A_{i,j} |i\rangle\langle j|$$

of an operator in terms of the *matrix elements*

$$A_{i,j} = \langle i|A|j\rangle.$$

The matrix elements link A to its representation in the column vector picutre. Indeed, applying two completeness relations to

$$|\beta\rangle = A|\alpha\rangle \tag{1}$$

gives

$$\sum_i |i\rangle\beta_i = \sum_{i,j} |i\rangle A_{i,j}\alpha_j$$

which is equivalent to

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_d\end{pmatrix} \begin{pmatrix}A_{1,1} & \dots & A_{1,d}\\ \vdots & & \vdots\\ A_{d,1} & \dots & A_{d,d}\end{pmatrix} \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_d\end{pmatrix}. \tag{2}$$

Thus (1) and (2) link the brak-ket and the matrix picture of the action of operators.

The trace

The *trace* of an operator A is

$$\operatorname{tr} A = \sum_i \langle i|A|i\rangle = \sum_i A_{i,i}. \tag{3}$$

Anyone should be able to prove the following elementary properties:

Lemma 1. *The following holds:*

1. *The trace is a linear map*
$$\operatorname{tr}(\alpha A + \beta B) = \alpha \operatorname{tr} A + \beta \operatorname{tr} B.$$

2. *Cyclic invariance*
$$\operatorname{tr} ABC = \operatorname{tr} BCA,$$
and symmetry
$$\operatorname{tr} AB = \operatorname{tr} BA$$

3. *Transpose and adjoint:*
$$\operatorname{tr} A^T = \operatorname{tr} A, \qquad \operatorname{tr} A^\dagger = \bar{\operatorname{tr}} A.$$

4. *Trace of an outer product is the inner product*
$$\operatorname{tr} |\alpha\rangle\langle\beta| = \langle\beta|\alpha\rangle.$$

5. *The trace does not depend on the ONB chosen in (3). What is more, the trace equals the sum of the eigenvalues (counted with algebraic multiplicity).*

6. *Hilbert-Schmidt inner product:*
$$\operatorname{tr} A^\dagger B = \sum_{i,j} \bar{A}_{i,j} B_{i,j}.$$

Tensor products

Let $\mathcal{H}_1, \mathcal{H}_2$ be two Hilbert spaces, with respective bases $\{|e_i\rangle\}_{i=1}^{d_1}, \{|f_j\rangle\}_{j=1}^{d_2}$. Their *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the space of linear combinations of the symbols $|e_i\rangle \otimes |f_j\rangle$, where $i \in \{1, \dots, d_1\}, j \in \{1, \dots, d_2\}$.

> At this points, these are purely formal expressions. There are various "basis-free" definitions of the tensor product that are slightly more complicated to state, but have the conceptual advantage that all elements that go into the construction have clear geometric interpretations. We won't touch on this here.

Thus,
$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \Big\{ \sum_{i,j} c_{i,j} |e_i\rangle \otimes |f_j\rangle \,\Big|\, c_{i,j} \in \mathbb{C} \Big\}.$$

As usual, there are various short-hand notations used in physics:
$$|e_i\rangle \otimes |f_j\rangle = |e_i\rangle|f_j\rangle = |e_i, f_j\rangle = |i, j\rangle.$$

If $|\alpha\rangle \in \mathcal{H}_1, |\beta\rangle \in \mathcal{H}_2$, then their tensor product is defined as
$$|\alpha\rangle \otimes |\beta\rangle = \sum_{i,j} \alpha_i \beta_j |i, j\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{4}$$

Such elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ are called *product vectors*. A quick parameter counting argument shows that product vectors depend on $d_1 + d_2 - 1$ complex parameters (why the "-1"?), while the tensor product Hilbert space has dimension $d_1 d_2$.
$$\dim \mathcal{H}_1 \otimes \mathcal{H}_2 = d_1 d_2.$$

This suggests (correctly) that "most" elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ are not of product form. In QM, such vectors are called *entangled*, and we'll have much more to say about this. On the other hand, the tensor product space is (by definition) *spanned* by product vectors. Thus, if we want to define an operator on it, it suffices to specify how it acts on products. We will use this repeatedly, starting two paragraphs below.

Exercise: Work out the correspondence between $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $L(\mathcal{H}_1, \mathcal{H}_2)$.

The definition (4) implies the "distributive laws"

$$(|\alpha\rangle + |\beta\rangle) \otimes (|\gamma\rangle + |\delta\rangle) = |\alpha\rangle \otimes |\gamma\rangle + |\alpha\rangle \otimes |\delta\rangle + |\beta\rangle \otimes |\gamma\rangle + |\beta\rangle \otimes |\delta\rangle \qquad (5)$$

and

$$(c|\alpha\rangle) \otimes |\beta\rangle = |\alpha\rangle \otimes (c|\beta\rangle) = c\,(|\alpha\rangle \otimes |\beta\rangle). \qquad (6)$$

Now let A be an operator on $\mathcal{H}_1$, and B on $\mathcal{H}_2$. Then we define an operator $A \otimes B$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ via its action on product vectors

$$(A \otimes B)(|\alpha\rangle \otimes |\beta\rangle) = (A|\alpha\rangle) \otimes (B|\beta\rangle).$$

Frequently, we will want to have an operator "acto on just one of the factors, leaving the other ones alone". This is done by tensoring with the identity operator $\mathbb{1}$: $A \otimes \mathbb{1}$ or $\mathbb{1} \otimes B$. The $\mathbb{1}$'s are often suppresed, and the factor on which a given operator acts non-trivially is indicated by a sub- or a superscript:

$$A_1 = A^{(1)} = A \otimes \mathbb{1}.$$

To be continued…