Linear Transformations

Lecture 3

Physics 342
Quantum Mechanics I

Friday, February 1st, 2008

We finish up the linear algebra section by making some observations about transformations (matrices) and decompositions (diagonalization) that are particularly useful in their generalized form. Quantum mechanics can be described by a few flavors of linear algebra: Infinite dimensional function spaces (both discrete and continuous) make up one vector space of interest, finite dimensional ones are another. To the extent that we are interested in quantum mechanics, then, we are interested in vector spaces, and the notation and ideas are common to all applications (trickier are issues of interpretation and existence).

In the end, the first object of study, the wave function in position space, can be represented by a vector in Hilbert space (the vector space of square integrable functions). Operations like "measuring the energy of the system" take the form of Hermitian "matrices" (differential operators). As we shall see, this is a natural interpretation for such operators, since their spectra are real, corresponding to allowable, measurable quantities.

3.1 Linear Transformations

We left off last time by defining the operator $\mathbf{1}$:

$$\mathbf{1} = \sum_{i=1}^{N} |e_i\rangle \langle e_i| \tag{3.1}$$

where its name comes from the fact that $\mathbf{1}\,|\alpha\rangle = |\alpha\rangle$. Notice that this operator takes a ket and returns a ket, but it can also take a bra and return a bra:

$$\langle\beta|\,\mathbf{1} = \langle\beta|\left(\sum_{i=1}^{N}|e_i\rangle\,\langle e_i|\right) = \sum_{i=1}^{N}\langle\beta|e_i\rangle\,\langle e_i| = \langle\beta|\,. \tag{3.2}$$

This somewhat trivial operator is a linear operator – in general, a linear operator $\hat{T}$ (following Griffiths, we will put hats on operators to distinguish them from complex numbers) takes kets to kets (bras to bras) and is linear:

$$\hat{T}\,(u\,|\alpha\rangle + v\,|\beta\rangle) = u\,\hat{T}\,|\alpha\rangle + v\,\hat{T}\,|\beta\rangle \tag{3.3}$$

for $u, v \in \mathbb{C}$.

Then, to define a linear operator, it suffices to define its action on the basis kets, since by linearity, for

$$|\alpha\rangle = \sum_{i=1}^{N} a_i\,|e_i\rangle\,, \tag{3.4}$$

we have:

$$\hat{T}\,|\alpha\rangle = \sum_{i=1}^{N} a_i\,\hat{T}\,|e_i\rangle\,, \tag{3.5}$$

so if we knew what $\hat{T}\,|e_i\rangle$ was, for all i, we could describe the effect of $\hat{T}$ on an arbitrary ket.

In order to represent $\hat{T}$ in a basis, we use two applications of the $\mathbf{1}$ operator:

$$\begin{aligned}\mathbf{1}\,\hat{T}\,\mathbf{1} &= \left(\sum_{i=1}^{N}|e_i\rangle\,\langle e_i|\right)\,\hat{T}\,\left(\sum_{j=1}^{N}|e_j\rangle\,\langle e_j|\right)\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\langle e_i|\,\hat{T}\,|e_j\rangle\;|e_i\rangle\,\langle e_j|\,,\end{aligned} \tag{3.6}$$

then we see that the ket $|\beta\rangle = \hat{T}\,|\alpha\rangle$ can be written in component form:

$$\begin{aligned}|\beta\rangle = \mathbf{1}\,\hat{T}\,\mathbf{1}\;|\alpha\rangle &= \sum_{i=1}^{N}\sum_{j=1}^{N}\underbrace{\langle e_i|\,\hat{T}\,|e_j\rangle}_{\equiv T_{ij}}\;|e_i\rangle\,\underbrace{\langle e_j|\alpha\rangle}_{=a_j}\\ &= \sum_{i=1}^{N}\left(\sum_{j=1}^{N}T_{ij}\,a_j\right)\;|e_i\rangle\,,\end{aligned} \tag{3.7}$$

so we identify the components of $|\beta\rangle = \sum_{i=1}^{N} b_i \; |e_i\rangle$ as the sum over j above – in column vector representation:

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N} T_{1j}\, a_j \\ \sum_{j=1}^{N} T_{2j}\, a_j \\ \sum_{j=1}^{N} T_{3j}\, a_j \\ \vdots \end{pmatrix}, \tag{3.8}$$

and this is precisely the relation we would have if we viewed $\hat{T}$ as a matrix with entries $T_{ij} = \langle e_i|\; \hat{T}\; |e_j\rangle$ and $|a\rangle$ as a column vector:

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \dots \\ T_{21} & T_{22} & T_{23} & \dots \\ T_{31} & T_{32} & T_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}. \tag{3.9}$$

So there is an obvious connection between an "abstract linear transformation" and a matrix, at least in the finite dimensional case. The association relies on a representation for the basis vectors, and our ability to write an arbitrary vector in the space as a linear combination of the basis vectors (a defining property of basis vectors).

3.2 Hermitian Linear Transformations

From the basis representation, a ket $|\alpha\rangle$ is naturally associated with a bra $\langle\alpha|$ via the conjugate transpose – that makes the inner product work out in its usual row-column vector form[1] . We are suggesting that if $|\alpha\rangle$ can be

[1] The conjugate transpose of a column vector is a row vector with conjugated components – this is the finite-dimensional distinction between $|\alpha\rangle$ and $\langle\alpha|$ – concretely, if we have a column vector **a** with entries:

$$\mathbf{a} \doteq \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}, \tag{3.10}$$

then

$$\mathbf{a}^\dagger = \begin{pmatrix} a_1^* & a_2^* & \dots & a_N^* \end{pmatrix}. \tag{3.11}$$

represented through its components as:

$$|\alpha\rangle \doteq \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \tag{3.12}$$

then $\langle\alpha|$ should be a row vector with the complex conjugate of the components as its entries:

$$\langle\alpha| \doteq \begin{pmatrix} a_1^* & a_2^* & a_3^* & \dots \end{pmatrix}. \tag{3.13}$$

Now the question is: For what linear transformation $\hat{H}$ is $\langle\alpha|\ \hat{H}$ the complex conjugate of $\hat{H}\ |\alpha\rangle$? Viewed as matrix-vector multiplication, $\hat{H}\ |\alpha\rangle$ would have $(\hat{H}\ |\alpha\rangle)^\dagger \sim \langle\alpha|\ \hat{H}^\dagger$ [2] – then such a transformation has the property that $\hat{H} = \hat{H}^\dagger$, and is called a Hermitian operator. This is all pretty straightforward on the finite side, where we can really take transposes and conjugate matrices operationally, but it all holds (sort of) on the continuous, infinite dimensional side as well.

Example

For matrices, the operation of taking the conjugate-transpose is easy to perform. Consider all two dimensional complex matrices, with entries specified by $h_{ij} = u_{ij} + i\, v_{ij}$ for $u_{ij}, v_{ij} \in \mathbb{R}$

$$\mathbb{H} \doteq \begin{pmatrix} u_{11} + i\, v_{11} & u_{12} + i\, v_{12} \\ u_{21} + i\, v_{21} & u_{22} + i\, v_{22} \end{pmatrix} \tag{3.14}$$

then

$$\mathbb{H}^\dagger \equiv \left(\mathbb{H}^T\right)^* = \begin{pmatrix} u_{11} - i\, v_{11} & u_{21} - i\, v_{21} \\ u_{12} - i\, v_{12} & u_{22} - i\, v_{22} \end{pmatrix} \tag{3.15}$$

To make $\mathbb{H}$ Hermitian: $\mathbb{H}^\dagger = \mathbb{H}$, we need to set: $v_{11} = v_{22} = 0$, $u_{12} = u_{21}$ and $v_{21} = -v_{12}$. Then it is the case that any vector $\mathbf{a}$ (with components

[2] The notation $\hat{H}^\dagger$ means take the "conjugate transpose", well-defined on the matrix side.

a_1 and a_2 complex) can be multiplied by $\mathbb{H}$ in the usual way:

$$\begin{pmatrix} u_{11} & u_{12} + i\, v_{12} \\ u_{12} - i\, v_{12} & u_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1\, u_{11} + a_2\,(u_{12} + i\, v_{12}) \\ a_1\,(u_{12} - i\, v_{12}) + a_2\, u_{22} \end{pmatrix}, \tag{3.16}$$

and has $(\mathbb{H}\,\mathbf{a})^\dagger$ equal to

$$(\mathbb{H}\,\mathbf{a})^\dagger \doteq \begin{pmatrix} a_1^*\, u_{11} + a_2^*\,(u_{12} - i\, v_{12}) & a_1^*\,(u_{12} + i\, v_{12}) + a_2^*\, u_{22} \end{pmatrix}, \tag{3.17}$$

which is the same as:

$$\mathbf{a}^\dagger\,\mathbb{H} = \begin{pmatrix} a_1^*\, u_{11} + a_2^*\,(u_{12} - i\, v_{12}) & a_1^*\,(u_{12} + i\, v_{12}) + a_2^*\, u_{22} \end{pmatrix}. \tag{3.18}$$

This is all very much the analogue of symmetric matrices for real vector spaces like $\mathbb{R}^N$. There, the question is: Under what restrictions on $\mathbb{O}$ is $\mathbf{a}^T\,\mathbb{O} = (\mathbb{O}\,\mathbf{a})^T$ for $\mathbf{a} \in \mathbb{R}^N$? This gives $\mathbb{O} = \mathbb{O}^T$. We are interested in the matrices $\mathbb{H}$ and $\mathbb{O}$ because of their eigenvector decomposition properties.

3.2.1 Eigenvalues and Eigenvectors

Hermitian operators are of primary interest to us, and part of the reason is their decomposition. Recall from linear algebra the definition of eigenvectors and eigenvalues: For a generic matrix $\mathbb{A}$, we call λ (complex) and $\mathbf{v}$ an eigenvalue, and associated eigenvector if

$$\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}, \tag{3.19}$$

i.e. $\mathbf{v}$ is a special vector, specific to the matrix $\mathbb{A}$ which, when acted on by $\mathbb{A}$ gets scaled by a constant, but does not change direction. If we take the set of all eigenvectors (there will be N of them for $\mathbb{A} \in \mathbb{C}^{N\times N}$), and make them the columns of a matrix $\mathbb{V}$:

$$\mathbb{V} \doteq \left[\ \mathbf{v}_1 \ \middle|\ \mathbf{v}_2 \ \middle|\ \ldots \ \middle|\ \mathbf{v}_N\ \right], \tag{3.20}$$

then we can write the defining equation for eigenvectors and eigenvalues as:

$$\mathbb{A}\,\mathbb{V} = \mathbb{V}\,\mathbb{S} \tag{3.21}$$

and $\mathbb{S}$ is a diagonal matrix with the eigenvalues of $\mathbb{A}$ as its diagonal elements. If $\mathbb{V}$ is invertible, then, we can write

$$\mathbb{V}^{-1}\,\mathbb{A}\,\mathbb{V} = \mathbb{S} \tag{3.22}$$

the diagonal form of $\mathbb{A}$. There are various special forms that allow the above to be simplified – for example, when $\mathbb{V}^{-1} = \mathbb{V}^T$, we call $\mathbb{V}$ an "orthogonal" matrix, and these are associated with symmetric, real $\mathbb{A}$ (i.e. $\mathbb{A} = \mathbb{A}^T$ with real entries). For a complex matrix, if the conjugate-transpose is equal to the inverse: $\mathbb{V}^{-1} = \left(\mathbb{V}^T\right)^* \equiv \mathbb{V}^\dagger$, then the matrix is said to be "unitary", and these come from the decomposition of Hermitian $\mathbb{A}$.

Hermitian operators $\mathbb{A}^\dagger = \mathbb{A}$ have two important properties that are the motivation for their role in quantum theory:

- The eigenvalues of $\mathbb{A} = \mathbb{A}^\dagger$ are real: For an eigenvector $\mathbf{v}$ with $\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}$, we can take the Hermitian conjugate (and use $\mathbb{A} = \mathbb{A}^\dagger$) to write:

$$\mathbf{v}^\dagger\,\mathbb{A} = \lambda^*\,\mathbf{v}^\dagger, \tag{3.23}$$

  and if we multiply this on the right by $\mathbf{v}$, we get

$$\mathbf{v}^\dagger\,\mathbb{A}\,\mathbf{v} = \lambda^*\,\mathbf{v}^\dagger\,\mathbf{v}. \tag{3.24}$$

  On the other hand, if we multiply $\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}$ on the left by $\mathbf{v}^\dagger$, then:

$$\mathbf{v}^\dagger\,\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}^\dagger\,\mathbf{v}, \tag{3.25}$$

  and the right hand sides of (3.24) and (3.25) can only be equal (as they must) if $\lambda = \lambda^*$.

- The eigenvectors of $\mathbb{A} = \mathbb{A}^\dagger$ are orthogonal. We'll show that eigenvectors associated with different eigenvalues are orthogonal – take $\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}$ and $\mathbb{A}\,\mathbf{w} = \sigma\,\mathbf{w}$ with $\sigma \neq \lambda$, and both σ and λ real by the above.

  Then we can take the defining equation for $\mathbf{v}$: $\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{v}$, and multiply on the left by $\mathbf{w}^\dagger$:

$$\mathbf{w}^\dagger\,\mathbb{A}\,\mathbf{v} = \lambda\,\mathbf{w}^\dagger\,\mathbf{v}. \tag{3.26}$$

  For the conjugate-transpose of the defining equation for $\mathbf{w}$: $\mathbf{w}^\dagger\,\mathbb{A} = \sigma\,\mathbf{w}^\dagger$ (using $\mathbb{A} = \mathbb{A}^\dagger$ and σ real), we can multiply on the right by $\mathbf{v}$ to obtain:

$$\mathbf{w}^\dagger\,\mathbb{A}\,\mathbf{v} = \sigma\,\mathbf{w}^\dagger\,\mathbf{v}, \tag{3.27}$$

and comparing (3.26) and (3.27), we end up with

$$(\lambda - \sigma)\ \mathbf{w}^\dagger\,\mathbf{v} = 0 \tag{3.28}$$

which can only be true if $\mathbf{w}^\dagger\,\mathbf{v} = 0$ since $\lambda \neq \sigma$ by assumption.

Hermitian Operators Eigenvalues in Bra-Ket Notation

We reproduce the two arguments above using operators and bras and kets, just as another example of the bra-ket notation.

- The eigenvalues are all real: Take an eigenket $|\alpha\rangle \neq 0$ with eigenvalue α, then

$$\hat{A}\,|\alpha\rangle = \alpha\,|\alpha\rangle \qquad \langle\alpha|\,\hat{A} = \langle\alpha|\,\alpha^* \tag{3.29}$$

  but, we can hit the ket on the left in the above with $\langle\alpha|$, and the bra on the right with $|\alpha\rangle$, and these must be equal:

$$\langle\alpha|\ \alpha\ |\alpha\rangle = \langle\alpha|\ \alpha^*\ |\alpha\rangle \longrightarrow \langle\alpha|\alpha\rangle\ (\alpha) = \langle\alpha|\alpha\rangle\ \alpha^*, \tag{3.30}$$

  and $\langle\alpha|\alpha\rangle \neq 0$ by assumption, so $\alpha = \alpha^*$ and $\alpha \in \mathbb{R}$.

- Eigenkets associated with different eigenvalues are orthogonal – take $\hat{A}\ |\alpha\rangle = \alpha\ |\alpha\rangle$ and $\langle\beta|\ \hat{A} = \langle\beta|\ \beta^* = \langle\beta|\ \beta$ (since the eigenvalues of $\hat{A}$ are real), then as before, we can hit the $|\alpha\rangle$ equation with $\langle\beta|$ and the $\langle\beta|$ equation with $|\alpha\rangle$:

$$\langle\beta|\ \hat{A}\ |\alpha\rangle = \langle\beta|\ \beta\ |\alpha\rangle = \langle\beta|\ \alpha\ |\alpha\rangle \tag{3.31}$$

  so, using the right-most equality, we have

$$(\beta - \alpha)\ \langle\beta|\alpha\rangle = 0, \tag{3.32}$$

  and if $\beta \neq \alpha$, then $\langle\beta|\alpha\rangle = 0$, the vectors are orthogonal.

As it turns out, even if eigenvalues are equal, we can construct orthogonal vectors – the eigenvectors span a subspace (eigenvectors of $\mathbb{A}$ with the same eigenvalue) – any linear combination of these vectors remains in the space (i.e. is also an eigenvector of $\mathbb{A}$ with the same eigenvalue) – and we can choose a set of spanning vectors that are orthonormal.

All of this is to say that the eigenvectors of a Hermitian matrix (or more generally, a Hermitian operator) span the original space – they form a linearly independent set (which can be normalized as well), and, if we form a matrix $\mathbb{U}$ with the eigenvectors as its columns, then $\mathbb{U}^\dagger \mathbb{U} = \mathbb{I}$. Hermitian matrices are diagonlized by unitary matrices. In matrix form, if $\mathbb{H}^\dagger = \mathbb{H}$, then:

$$\mathbb{H}\,\mathbb{U} = \mathbb{U}\,\mathbb{S} \longrightarrow \mathbb{U}^\dagger\,\mathbb{H}\,\mathbb{U} = \mathbb{S}. \tag{3.33}$$

This should be familiar – it is the complex version of the statement: Real, symmetric matrices are diagonalized by orthogonal matrices. We are relying heavily on the finite matrix form, but we will consider operator versions of these ideas.

3.3 Commutators

We define the operation [,], the commutator of two matrices $\mathbb{A}$ and $\mathbb{B}$ as:

$$[\mathbb{A}, \mathbb{B}] \equiv \mathbb{A}\,\mathbb{B} - \mathbb{B}\,\mathbb{A}, \tag{3.34}$$

The commutator of two operators (in the more general setting of linear operators not easily representable as matrices) tells us a lot – for example, take Hermitian $\mathbb{A}$ and $\mathbb{B}$, and suppose these can be diagonalized by the same unitary matrix $\mathbb{U}$:

$$\mathbb{U}^\dagger\,\mathbb{A}\,\mathbb{U} = \mathbb{P} \qquad \mathbb{U}^\dagger\,\mathbb{B}\,\mathbb{U} = \mathbb{Q}, \tag{3.35}$$

then, noting that for diagonal matrices $\mathbb{P}$ and $\mathbb{Q}$, $[\mathbb{P}, \mathbb{Q}] = 0$, we have

$$0 = \mathbb{P}\,\mathbb{Q} - \mathbb{Q}\,\mathbb{P} = \mathbb{U}^\dagger\,\mathbb{A}\,\mathbb{U}\,\mathbb{U}^\dagger\,\mathbb{B}\,\mathbb{U} - \mathbb{U}^\dagger\,\mathbb{B}\,\mathbb{U}\,\mathbb{U}^\dagger\,\mathbb{A}\,\mathbb{U} = \mathbb{U}^\dagger\,(\mathbb{A}\,\mathbb{B} - \mathbb{B}\,\mathbb{A})\,\mathbb{U}. \tag{3.36}$$

Since $\mathbb{U}^\dagger$ is invertible, it has no null space, and we must have $[\mathbb{A}, \mathbb{B}] = 0$. The other direction works as well, so that *if* two Hermitian operators have vanishing commutator, they can be simultaneously diagonalized by $\mathbb{U}$.

Homework

Reading: Griffiths, pp. 441–446, 449–456.

Problem 3.1

In this problem, we will generate the units for the Schrödinger equation. Use the notation $|f|$ to denote the "units of f", and use M, T, E, and L to refer to "mass", "time", "energy" and "length". So, for example, if we have $x(t) = v\,t$ (describing an object moving with a constant speed), we would write: $|t| = T$, $|v| = L/T$ and $|x| = |v|\,|t| = (L/T)\,T = L$.

a. Given a function $f(x,t)$, with units defined by $|f| = F$, what are the units of:

$$\frac{\partial^2 f(x,t)}{\partial x^2} \quad \text{and} \quad \frac{\partial f(x,t)}{\partial t} \tag{3.37}$$

b. Using the above result, find the units of $\hbar$ in Schrödinger's equation in one spatial dimension:

$$-\frac{\hbar^2}{2\,m}\,\frac{\partial^2 \Psi(x,t)}{\partial x^2} + U(x)\,\Psi(x,t) = i\,\hbar\,\frac{\partial \Psi(x,t)}{\partial t} \tag{3.38}$$

given $|m| = M$, and $|U(x)| = E$.

Problem 3.2

Use a multiplicative separation ansatz $\Psi(x,t) = \psi(x)\,\phi(t)$ and separate Schrödinger's equation (in one spatial dimension with $U(x)$ a function only of x, as in (3.38))– write all the x-dependence on the left-hand side, all the t-dependence on the right.

a. Set both sides equal to a separation constant, call it E – what are the units of this constant?

b. Solve for $\phi(t)$ in terms of E and write down the ODE that governs $\psi(x)$.

Problem 3.3

We know that for a finite-dimensional, real vector space (like $\mathbf{R}^3$), if we have a symmetric matrix $\mathbb{A}$ ($\mathbb{A}^T = \mathbb{A}$), with eigenvectors $\mathbf{v}_i$ and eigenvalues λ_i, then for $\lambda_i \neq \lambda_j$, we must have $\mathbf{v}_i \cdot \mathbf{v}_j = 0$. This problem is meant to indicate that even when $\lambda_i = \lambda_j$, we can *choose* $\mathbf{v}_i \cdot \mathbf{v}_j = 0$.

a. Suppose we have only two eigenvectors that share a common eigenvalue:

$$\mathbb{A}\,\mathbf{v}_1 = \lambda\,\mathbf{v}_1 \qquad \mathbb{A}\,\mathbf{v}_2 = \lambda\,\mathbf{v}_2 \tag{3.39}$$

but that all the rest of the eigenvectors of $\mathbb{A}$ have: $\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}$, and $\mathbf{v}_1 \cdot \mathbf{v}_i = \mathbf{v}_2 \cdot \mathbf{v}_i = 0$ for all i (that is: All other eigenvalues are distinct and not equal to λ). Show that the generic linear combination: $\mathbf{p} \equiv \alpha\,\mathbf{v}_1 + \beta\,\mathbf{v}_2$ (with α, β real) is also an eigenvector of $\mathbb{A}$ with eigenvalue λ.

b. From above, then, we know that $\mathbf{v}_1$ and $\mathbf{p}$ are two eigenvectors of $\mathbb{A}$ with eigenvalue λ for any α, β. Suppose $\mathbf{v}_1 \cdot \mathbf{v}_2 = \gamma \neq 0$ so that $\mathbf{v}_1$ and $\mathbf{v}_2$ are not orthogonal. Our goal is to replace $\mathbf{v}_2$ with a vector that *is* orthogonal to $\mathbf{v}_1$: Find a combination of α and β so that

$$\mathbf{v}_1 \cdot \mathbf{p} = 0. \tag{3.40}$$

We now have a complete orthogonal set of eigenvectors.