Some important theorems in QM

Wagner L. Truppel

September 6, 2005

Brief review of basic linear algebra

Recall the definition of a **linear vector space**. An ordered pair, $(\mathcal{V}, \mathcal{S})$, where $\mathcal{V}$ is a set and $\mathcal{S}$ is a *field*, forms a linear vector space if it satisfies the following axioms:

1. An operation of *addition* is defined between elements of $\mathcal{V}$ (called *vectors*) such that:

   (a) the addition of any two elements of $\mathcal{V}$ results in an element of $\mathcal{V}$. This is called the *closure* condition: $\forall \vec{u}, \vec{v} \in \mathcal{V} \Rightarrow \vec{u} + \vec{v} \in \mathcal{V}$.

   (b) the operation is *commutative*: $\forall \vec{u}, \vec{v} \in \mathcal{V} \Rightarrow \vec{u} + \vec{v} = \vec{v} + \vec{u}$.

   (c) the operation is *associative*: $\forall \vec{u}, \vec{v}, \vec{w} \in \mathcal{V} \Rightarrow (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.

   (d) $\mathcal{V}$ has a *null* element: $\exists \vec{0} \mid \forall \vec{u} \in \mathcal{V} \Rightarrow \vec{u} + \vec{0} = \vec{u}$.

   (e) every element has an *additive inverse*: $\forall \vec{u} \in \mathcal{V}, \exists \vec{v} \in \mathcal{V} \mid \vec{u} + \vec{v} = \vec{0}$.

   The axioms above are those of a *group*. Since the operation is also commutative, the group is called an *Abelian group*. $\mathcal{V}$ is therefore an Abelian group under the operation in question. It's possible to show that the null vector is unique and that each vector has a unique additive inverse.

2. An operation of *multiplication by scalars* is defined between elements of $\mathcal{V}$ and elements of $\mathcal{S}$ (called *scalars*) such that:

   (a) $\forall a \in \mathcal{S}, \forall \vec{u}, \vec{v} \in \mathcal{V} \Rightarrow a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$.

   (b) $\forall a, b \in \mathcal{S}, \forall \vec{u} \in \mathcal{V} \Rightarrow (a + b)\vec{u} = a\vec{u} + b\vec{u}$.

   (c) $\forall a, b \in \mathcal{S}, \forall \vec{u} \in \mathcal{V} \Rightarrow a(b\vec{u}) = (ab)\vec{u}$.

   The first two require that the multiplication by scalars and the vector addition be *distributive*. The set $\mathcal{S}$ needs to be a *field* so that it has its own addition and multiplication operations: note that in $(a\vec{u} + b\vec{u})$, the addition is that of vectors while in

$(a+b)\vec{u}$, the addition is that of scalars. It's possible to show that these axioms imply: $1\vec{u} = \vec{u}$ and $0u = \vec{0}$, for any $\vec{u}$ in $\mathcal{V}$, where 1 and 0 are the elements of $\mathcal{S}$ such that $\forall a \in \mathcal{S} \Rightarrow 1a = a$ (1 is called the *multiplicative identity*) and $\forall a \in \mathcal{S}, \exists b \in \mathcal{S} \mid a+b = 0$ (0 is also called the *additive identity*, since $a + 0 = a$ for all a in $\mathcal{S}$).

A *linear combination* of vectors is an expression of the form $\sum_i a_i \vec{u}_i$ where the a_i are elements of $\mathcal{S}$ and the $\vec{u}_i$ are elements of $\mathcal{V}$. A set of vectors $\{\vec{u}_i\}$ is said to be *linearly independent* if $\sum_i a_i \vec{u}_i = \vec{0}$ implies necessarily that every scalar in the linear combination vanishes, $a_i = 0$. The *dimension* of a linear vector space is the largest number of linearly independent vectors in the space.

Any subset of $\mathcal{V}$ comprised of n linearly independent vectors, where n is the dimension of $\mathcal{V}$ (assumed finite for the moment), forms a *basis* for $\mathcal{V}$ and is said to *span* the whole vector space. Any vector in $\mathcal{V}$ is then a linear combination of basis vectors. This is easy to prove: since the n basis vectors are the largest subset of $\mathcal{V}$ that is linearly independent, then any linear combination having the basis vectors and any other vector cannot form a linearly independent subset. Thus, $\sum_i a_i \vec{e}_i + b\vec{v} = \vec{0}$ (where the $\vec{e}_i$'s are the chosen basis vectors) does not imply that all a_i's and b be all equal to zero. In particular, b cannot be equal to zero or else the expression would form a linearly independent combination (because the basis vectors already do). Thus, $\vec{v} = (-1/b) \sum_i a_i \vec{e}_i$ and we've shown that any vector can be written as a linear combination of basis vectors.

Recall also the definition of an *operator*: a function from $\mathcal{V}$ into $\mathcal{V}$. A *linear* operator T is such that $T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v})$. Since any vector can be written as a linear combination of basis vectors, the action of any **linear** operator in $\mathcal{V}$ is uniquely determined by its action on the basis vectors alone:

$$T(\vec{u}) = T(\sum_i a_i \vec{e}_i) = \sum_i a_i T(\vec{e}_i)\,.$$

Since, however, $T(\vec{e}_i)$ is itself a vector it, too, can be written as a linear combination of the basis vectors:

$$T(\vec{e}_i) = \sum_k T_{ik} \vec{e}_k\,.$$

The set of n^2 elements T_{ik} of $\mathcal{S}$ form the *matrix representation* of the operator T in the basis chosen. The problem of determining, for a given operator T, the set of vectors $\vec{u}_i$ and scalars a_i such that $T(\vec{u}_i) = a_i \vec{u}_i$ is called the *eigenvalue problem* of T. There are generally n solutions $(a_i, \vec{u}_i)$.

Brief introduction to advanced linear algebra

An *inner product* $\langle \ldots | \ldots \rangle$ is a function from $\mathcal{V} \times \mathcal{V}$ into the complex field $\mathcal{C}$ satisfying:

1. $\forall \vec{u}, \vec{v}, \vec{w} \in \mathcal{V} \Rightarrow \langle \vec{u} + \vec{v} | \vec{w} \rangle = \langle \vec{u} | \vec{w} \rangle + \langle \vec{v} | \vec{w} \rangle$
   (distributive with respect to vector addition).
2. $\forall a \in \mathcal{S}, \forall \vec{u}, \vec{v} \in \mathcal{V} \Rightarrow \langle \vec{u} | a\vec{v} \rangle = a \langle \vec{u} | \vec{v} \rangle$
3. $\forall \vec{u}, \vec{v} \in \mathcal{V} \Rightarrow \langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle^*$. Note that $\langle \vec{u} | \vec{u} \rangle$ is a real number.
4. $\forall \vec{u} \in \mathcal{V} \Rightarrow \langle \vec{u} | \vec{u} \rangle \geq 0$.
5. $\langle \vec{u} | \vec{u} \rangle = 0 \Rightarrow \vec{u} = \vec{0}$.

An inner product allows us to define a notion of *orthogonality*: two non-zero vectors are mutually orthogonal if their inner product is zero. An inner product thus lets us define a basis whose vectors are mutually orthogonal. This is done by employing the *Gram-Schmidt* orthogonalization method on any already-given basis.

A *normed* (linear) vector space is a (linear) vector space where an additional function $\| \ldots \|$ is defined, from vectors into the real numbers, such that:

1. $\forall \vec{u} \in \mathcal{V}, \|\vec{u}\| \geq 0$.
2. $\|\vec{u}\| = 0 \Rightarrow \vec{u} = \vec{0}$.

By defininining a norm, we're essentially endowing vectors with a notion of length. Vector spaces for which an inner product has been defined have a natural norm: $\|\vec{u}\| \equiv \langle \vec{u} | \vec{u} \rangle$ (or its square-root or some other similar function of $\langle \vec{u} | \vec{u} \rangle$ preserving the essential nature of norms). A normed vector space then may be endowed with an *orthonormal* basis, where $\langle \vec{e}_i | \vec{e}_k \rangle = \delta_{ik}$.

A *Hilbert space* is a linear vector space endowed with an inner product such that the natural norm is $\|\vec{u}\| \equiv \sqrt{\langle \vec{u} | \vec{u} \rangle}$ and in such a way that certain convergence conditions for sequences of vectors are met. Essentially, those conditions are that a sequence of vectors must converge to a vector in the space whenever the distance between successive vectors in the sequence tends to zero as we move along the sequence. The distance is defined as the norm of the difference between two vectors.

If $\mathcal{V}$ is a Hilbert space over the complex numbers, then an operator T in $\mathcal{V}$ is a *Hermitian* operator if $\langle \vec{u} | T\vec{v} \rangle = \langle T\vec{u} | \vec{v} \rangle$, for all $\vec{u}$ and $\vec{v}$ in $\mathcal{V}$. Choosing mutually orthonormal basis vectors, it's easy to show that $\langle \vec{e}_i | T\vec{e}_k \rangle = \langle T\vec{e}_i | \vec{e}_k \rangle$ implies $T_{ik} = (T_{ki})^*$.

Let $\mathcal{H}$ be a hermitian operator in a Hilbert space. The following results are of fundamental importance to quantum mechanics. I'll switch to Dirac's notation since it's easier to use.

1. The eigenvalues of a Hermitian operator are real:

$$\mathcal{H}|u\rangle = E|u\rangle \Rightarrow E = \frac{\langle u|\mathcal{H}|u\rangle}{\langle u|u\rangle} \Rightarrow E^* = \frac{\langle u|\mathcal{H}^\dagger|u\rangle}{\langle u|u\rangle} = \frac{\langle u|\mathcal{H}|u\rangle}{\langle u|u\rangle} = E\,.$$

2. Eigenvectors with distinct eigenvalues are mutually orthogonal:

$$\left.\begin{array}{rcl} \mathcal{H}|u_i\rangle &=& E_i|u_i\rangle \\ \langle u_k|\mathcal{H} &=& E_k\langle u_k| \end{array}\right\} \Rightarrow \langle u_k|\mathcal{H}|u_i\rangle - \langle u_k|\mathcal{H}|u_i\rangle = E_i\langle u_k|u_i\rangle - E_k\langle u_k|u_i\rangle\,.$$

Thus, $(E_i - E_k)\langle u_k|u_i\rangle = 0$. So, if $E_i \neq E_k$, it follows that $\langle u_k|u_i\rangle = 0$.

Moreover, if $\mathcal{H}$ is bounded from below, that is, if there exists a real number c such that $\langle u|\mathcal{H}|u\rangle \geq c\,\langle u|u\rangle$ for all vectors $|u\rangle$, then two very important results follow. First, though, because $\mathcal{H}$ is assumed bounded from below, we can always re-index the eigenvectors and eigenvalues of $\mathcal{H}$ such that $\mathcal{H}|u_n\rangle = E_n|u_n\rangle$, where $E_0 \leq E_1 \leq \cdots \leq E_n \leq \cdots$. Now for the two very important results.

1. Let $|a\rangle$ be *any* vector. Then, the minimum of $E \equiv \dfrac{\langle a|\mathcal{H}|a\rangle}{\langle a|a\rangle}$ is:

   (a) E_0, if $|a\rangle$ can be any vector.

   (b) E_1, if $|a\rangle$ can be any vector orthogonal to $|0\rangle$.

   (c) E_n, if $|a\rangle$ can be any vector orthogonal to $|k\rangle, 0 \leq k \leq n-1$.

   Proof: An infinitesimal change in the vector $|a\rangle$, $\delta|a\rangle$, causes a corresponding change in E:

$$\begin{aligned} \delta E &= \frac{(\delta\langle a|)\,\mathcal{H}\,|a\rangle + \langle a|\,\mathcal{H}\,(\delta|a\rangle)}{\langle a|a\rangle} - \underbrace{\frac{\langle a|\mathcal{H}|a\rangle}{\langle a|a\rangle}}_{E}\,\frac{[(\delta\langle a|)\,|a\rangle + \langle a|\,(\delta|a\rangle)]}{\langle a|a\rangle} \\ &= \frac{1}{\langle a|a\rangle}\,[(\delta\langle a|)\,(\mathcal{H}-E)\,|a\rangle + \langle a|\,(\mathcal{H}-E)\,(\delta|a\rangle)] \\ &= \frac{2}{\langle a|a\rangle}\,\Re[(\delta\langle a|)\,(\mathcal{H}-E)\,|a\rangle]\,. \end{aligned}$$

   For E to be a minimum, δE must vanish when we perturb the vector $|a\rangle$ by an *arbitrary* (but small) amount. Setting δE to zero gives us that $(\mathcal{H}-E)\,|a\rangle = 0$, which means that $E = E_0$ since that is the minimum possible eigenvalue. It also means that $|a\rangle = |0\rangle$, the corresponding eigenvector. To prove (b), we follow the same argument and conclude that $(\mathcal{H}-E)\,|a\rangle = 0$, with the proviso that $|a\rangle$ be orthogonal to $|0\rangle$.

Well, all other eigenvectors are orthogonal (or can be made orthogonal[1]) to $|0\rangle$, so the minimum can now only occur for $E = E_1$, that is, for $|a\rangle = |1\rangle$. A similar argument allows us to prove (c).

2. If $\mathcal{H}$ is bounded from below but not from above ($\mathcal{H}$ is bounded from above if there exists a real number c such that $\langle u|\mathcal{H}|u\rangle \leq c\,\langle u|u\rangle$ for all vectors $|u\rangle$), then the set of all its eigenvectors is a *complete* set (that is, they form a basis for the Hilbert space in question).

   Proof: A set of basis vectors in a *finite*-dimensional linear vector space is, as proved earlier, *always* complete. Not necessarily so in the case of an infinite-dimensional space, in which case a set of basis vectors $\{|e_k\rangle\}$ (in a complex-based vector space) is said to be complete if there exists a set of (complex) numbers $\{C_n\}$ such that, for *any* vector $|a\rangle$, $\lim\limits_{n\to\infty}\langle R_n|R_n\rangle = 0$, where $|R_n\rangle \equiv |a\rangle - \sum_{k=0}^{n} C_k|e_k\rangle$. This is where those pesky convergence conditions come in. The result above is essentially saying that the sequence of vectors $\sum_{k=0}^{n} C_k|e_k\rangle$, as $n \to \infty$, converges to a vector of zero norm (that is, the null vector). That convergence condition is always satisfied for a Hilbert space, by definition, so a set of basis vectors in a Hilbert space is guaranteed to be complete. What we need to show, though, is that the eigenvectors of a Hermitian operator satisfying the conditions of this theorem form such a basis.

   First, we can re-index the eigenvalues and eigenvectors of $\mathcal{H}$ as before so that energies increase as indices increase. Next, because $\mathcal{H}$ is assumed bounded from below, we can always make $E_0 > 0$ by adding a constant to $\mathcal{H}$. Thus, we may always set things up so that $E_{n+1} \geq E_n$ and $E_0 > 0$.

   What we need to do now is to show that there exists a set of complex numbers $\{C_n\}$ such that, for *any* vector $|a\rangle$, $\lim\limits_{n\to\infty}\langle R_n|R_n\rangle = 0$, where $|R_n\rangle \equiv |a\rangle - \sum_{k=0}^{n} C_k|k\rangle$ and where the $|k\rangle$'s are the eigenvectors of $\mathcal{H}$. We'll show that the choice $C_k \equiv \langle k|a\rangle$ fits the bill.

   From the previous theorem, we know that
$$\frac{\langle R_n|\mathcal{H}|R_n\rangle}{\langle R_n|R_n\rangle} \geq E_{n+1}\,.$$
   That's because $|R_n\rangle$ is orthogonal to every eigenvector of $\mathcal{H}$ with energy below E_{n+1}: $\langle m|R_n\rangle = \langle m|a\rangle - \sum_{k=0}^{n} C_k\langle m|k\rangle = \langle m|a\rangle - C_m = 0$ if $m \leq n$ (note that the sum extends only up to $k = n$). Thus, since $E_{n+1} \geq E_n$, it follows that
$$\langle R_n|R_n\rangle \leq \frac{\langle R_n|\mathcal{H}|R_n\rangle}{E_n}\,.$$

[1] By a previous result, eigenvectors with distinct eigenvalues are orthogonal. Degenerate eigenvectors (those with equal eigenvalues) can be made orthogonal to one another by using the Gram-Schmidt orthogonalization method.

Since $E_n \to \infty$ as $n \to \infty$, we'll have shown that $\langle R_n | R_n \rangle \to 0$ in the same limit if we can show that $\langle R_n | \mathcal{H} | R_n \rangle$ has a finite limit.

$$\begin{aligned}
\langle R_n | \mathcal{H} | R_n \rangle &= \Big[\langle a | - \sum_{k=0}^{n} C_k^* \langle k | \Big] \mathcal{H} \Big[| a \rangle - \sum_{k=0}^{n} C_k | k \rangle \Big] \\
&= \langle a | \mathcal{H} | a \rangle - \sum_{k=0}^{n} C_k E_k \langle a | k \rangle - \sum_{k=0}^{n} C_k^* E_k \langle k | a \rangle \\
&\qquad + \sum_{i=0}^{n} \sum_{k=0}^{n} C_i^* C_k E_k \underbrace{\langle i | k \rangle}_{\delta_{ik}} \\
&= \langle a | \mathcal{H} | a \rangle - \sum_{k=0}^{n} C_k C_k^* E_k - \sum_{k=0}^{n} C_k^* C_k E_k + \sum_{i=0}^{n} C_k^* C_k E_k \\
&= \langle a | \mathcal{H} | a \rangle - \sum_{k=0}^{n} C_k C_k^* E_k \,.
\end{aligned}$$

Since all the energies are positive, the last term is positive, so $\langle R_n | \mathcal{H} | R_n \rangle \leq \langle a | \mathcal{H} | a \rangle$. But $\langle a | \mathcal{H} | a \rangle$ does not depend on n and $\langle R_n | \mathcal{H} | R_n \rangle \geq 0$. Thus, $0 \leq \langle R_n | \mathcal{H} | R_n \rangle \leq$ some constant, and the result of the theorem follows.

■