Hilbert Space for Quantum Mechanics
The Finite-Dimensional Spectral Theorem
and The Functional Calculus for Self-Adjoint Operators

Larry Susanka
Mathematics Department
Bellevue College

November 6, 2014

I accept the division.
I even accept the question.
But I have trouble remembering which side of the glass I’m on.

Larry Susanka, Fall 2014

Table of Contents
Counting Measures (Part I)
Adjoint in $\mathbb{C}^n$
$L^2(\mu)$ and the operator $T$
Handling Multiplicity

THE ARC OF QUANTUM WISDOM
$\mathbb{C}^n$ can be conceived of as the set of ordered $n$-tuples $v = (v^1, v^2, \ldots, v^n)$ of complex numbers made into a vector space with the familiar properties of addition and scalar multiplication.

For typographical convenience we represent these as rows, but for compatibility with matrix operations and standard function notation they are, actually, columns.

$\mathbb{C}^n$ AND $\langle v, w \rangle_c$

Let the vector $e_i$ denote the member of $\mathbb{C}^n$ that has a 1 in the $i$th row and zeroes elsewhere.

Then we have 

$$v = (v^1, v^2, \ldots, v^n) = v^1 e_1 + v^2 e_2 + \cdots + v^n e_n.$$ 

$\mathbb{C}^n$ is a Hilbert space with inner product

$$\langle v, w \rangle_c = v^1 \overline{w^1} + v^2 \overline{w^2} + \cdots + v^n \overline{w^n}.$$ 

$\mathbb{C}^n$ AND "REALITY" (PART I)

We identify $\mathbb{C}^n$ with some fragment of "reality" by identifying each $e_i$ with a specific displacement of something measurable and "vector-like" in the "reality" fragment displacements.

The "reality" displacements we consider will be called Reality.

In the structure we build we will use a concept of "angle" so in Reality there should be some concept analogous to "orthogonal."

From the outset, we should have good reason to believe that Reality looks like an $n$-dimensional inner product (i.e. Hilbert) space.

$\mathbb{C}^n$ AND "REALITY" (PART II)

We will also be interested in operators $T$: Reality $\rightarrow$ Reality and when we write $T(v) = w$ for certain $v, w \in \mathbb{C}^n$ what we really mean is that there are items $A$ and $B$ in Reality for which $T(A) = B$ and $w$ is the member of $\mathbb{C}^n$ which we have associated with $B$ and $v$ is the member of $\mathbb{C}^n$ which we have associated with $A$.

It is important that we make this clear: without an explicit, up-front "legend" associating specific, measurable "reality" displacements with the $e_i$ the equation $T(v) = w$ is meaningless.
\textbf{\(\mathbb{C}^n\) and Matrix Representations of \(T\)}

If \(T : \mathbb{C}^n \to \mathbb{C}^n\) is any linear operator it can be represented using left multiplication by matrix \(M = (T^j_i)\) where \(T^j_i\) is the coefficient on \(e_i\) in \(T(e_j)\). So matrix \(M\) is the \(n \times n\) matrix

\[M = (T(e_1) \ T(e_2) \cdots T(e_n))\]

and it follows by linearity and direct calculation that

\[T(v) = Mv = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} T^j_i v^j \right) e_i.\]

\textbf{Normal Operators on \(\mathbb{C}^n\) (Part I)}

If \(\lambda_1, \lambda_2, \ldots, \lambda_n\) are complex numbers we define

\[\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\]

to be the \(n \times n\) matrix with \(\lambda_1, \lambda_2, \ldots, \lambda_n\), in order, along the diagonal and zeroes elsewhere. Such matrices are called \textit{diagonal matrices}.

A square matrix \(M\) is said to be \textit{diagonalizable} if there is an invertible matrix \(P\) and diagonal matrix \(D\) for which

\[P^{-1}MP = D \quad \text{or, equivalently,} \quad MP = PD.\]

If \(M\) is diagonalizable and \(D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\) the numbers \(\lambda_1, \lambda_2, \ldots, \lambda_n\) are unique except for order.

\(P\) is said to \textit{diagonalize} \(M\).

\textbf{Normal Operators on \(\mathbb{C}^n\) (Part II)}

Operators that commute with their adjoint are called \textbf{normal} and those that are their own adjoint are called \textbf{self-adjoint}.

An invertible isometry on \(\mathbb{C}^n\) is called \textbf{unitary}.

The vocabulary is also applied to the matrices that represent these operators.

If \(P\) is a unitary matrix, then \(P^{-1} = P^\dagger\), and this property characterizes this type of matrix. Unitary matrices are normal.

It is a standard theorem of Linear Algebra that normal matrices are exactly those which are \textit{diagonalizable} by a unitary \textit{matrix of transition}. 

\textbf{Adjoint in \(\mathbb{C}^n\)}

The adjoint of an operator \(T\) is an operator \(T^\dagger\) defined by

\[\langle Tv, w \rangle_c = \langle v, T^\dagger w \rangle_c.\]

The matrix of \(T^\dagger\) can be determined by applying this formula to the \(e_i\). Then

\[\langle T e_i, e_j \rangle_c = \langle e_i, T^\dagger e_j \rangle_c\]

producing the equality \((T^\dagger)^j_i = \overline{T^j_i}\).

In other words the matrix of \(T^\dagger\) is \(M^\text{transpose}\).

Defining the conjugate transpose of any square matrix \(M\) to be \(M^\dagger\) we have (by this definition) that

\[M \leftrightarrow T \quad \text{if and only if} \quad M^\dagger \leftrightarrow T^\dagger.\]
Normal Operators on $\mathbb{C}^n$ (Part III)

Suppose $M$ is normal $MP = PD$ for diagonal $D$ and unitary $P$ as above. Let $P_i$ denote the $i$th column of $P$. We have

$$MP = (MP_1 MP_2 \cdots MP_n) = (\lambda_1 P_1 \lambda_2 P_2 \cdots \lambda_n P_n) = PD.$$  

So the columns of $P$ form a linearly independent set of eigenvectors for $M$ and, of course, the operator $T$ corresponding to $M$.

By virtue of the equation $P^{-1}MP = P^\dagger MP = D$ we say that any normal matrix (or operator) is unitarily equivalent to a diagonal matrix (or operator.)

Self-adjoint Operators on $\mathbb{C}^n$ (Part I)

Normal $M$ is self-adjoint exactly when all the $\lambda_i$ are real.

When $M$ is self-adjoint, we will insist (by permuting the columns of $P$ if necessary) that the $\lambda_i$ are listed along the diagonal in decreasing order. Of course some of these numbers may coincide.

For each $\lambda_i$ the eigenspace for $\lambda_i$, which is the span of the columns of $P$ corresponding to this eigenvalue, will be denoted $E_{\lambda_i}$.

The sum of the dimensions of the eigenspaces for distinct eigenvalues is $n$.

Self-adjoint Operators on $\mathbb{C}^n$ (Part II)

Let’s see how this looks working directly with the operator $T$.

Suppose we have self-adjoint $T$ and diagonalizing unitary $P$ so $MP = PD$ for diagonal $D$. Then $T(P_i) = \lambda_i P_i$ for $i = 1, \ldots, n$.

Any vector $v$ in $\mathbb{C}^n$ has a unique representation as

$$v = a^1 P_1 + a^2 P_2 + \cdots + a^n P_n$$  

where $a = P^{-1}v$ because $Pa = v$.

Note that

$$\|v\|^2 = \sum_{i=1}^{n} v^i \overline{v^i} = \sum_{i=1}^{n} a^i \overline{a^i}.$$  

The unitary operator given by matrix multiplication

$P: \mathbb{C}^n \to \mathbb{C}^n$ is an isometry: that is, for every $v, w \in \mathbb{C}^n$

$$\langle v, w \rangle_c = \langle Pv, Pw \rangle_c.$$  

Self-adjoint Operators on $\mathbb{C}^n$ (Part III)

$$T(v) = \lambda_1 a^1 P_1 + \lambda_2 a^2 P_2 + \cdots + \lambda_n a^n P_n$$  

so representing vectors in terms of this basis makes $T$ trivial to calculate.

Should you have mild regret concerning your initial choice of $e_1, e_2, \ldots, e_n$, you can use $P$ to tell you how to “change your mind” consistently.

Had you chosen $e_1$ to be the “reality fragment” displacement represented by $P_i$ for each $i$, as you very well could have, the matrix of $T$ would have been diagonal from the outset.

$P$ is your ”new choice to original choice” translator!
**COUNTING MEASURES**

\( \mu : \mathbb{P}(\mathbb{R}) \to [0, \infty] \) is called a **counting measure** on \( \mathbb{R} \) if there is a countable subset \( S \) of \( \mathbb{R} \), called the **support** of \( \mu \), so that \( \mu(A) \) is the cardinality of \( A \cap S \) for every \( A \in \mathbb{P}(\mathbb{R}) \).

For instance, if \( S = \{1, 2, 5\} \) then \( \mu(\mathbb{Q}) = 3 \) and \( \mu([0, 3]) = 2 \) and \( \mu([10, 20]) = 0 \).

Or if \( S = \mathbb{Q} \) then \( \mu([0, 3]) = \infty \) and \( \mu(\{\sqrt{2}, 7/3, e\}) = 1 \).

For each \( \lambda \in \mathbb{R} \) we define \( \delta_\lambda \) to be counting measure on the one-point set \( \{\lambda\} \).

This is called (by mathematicians) a **point mass at \( \lambda \)**. Sometimes these measures are called **Dirac delta functions**.

Note that \( \mu = \sum_{s \in S} \delta_s \) where we apply a sum of point masses to a set in the obvious way.

---

**L^1(\mu) (PART I)**

If \( f : \mathbb{R} \to \mathbb{C} \) and if \( \sum_{s \in S} |f(s)| < \infty \) we define

\[
\int_{\mathbb{R}} f \, d\mu = \sum_{s \in S} f(s) \mu(\{s\}) = \sum_{s \in S} f(s).
\]

The functions for which we have defined \( \int_{\mathbb{R}} f \, d\mu \) are called **integrable** or **summable**.

---

**L^1(\mu) (PART II)**

If \( \sum_{s \in S} |f(s)| = 0 \) we call \( f \) a **null function**, or, specifically, a **\( \mu \)-null function**.

The set of equivalence classes of integrable functions, where \( f \) is equivalent to \( g \) when \( f - g \) is a null function, is denoted \( L^1(\mu) \).

When \( f - g \) is a null function we write

\[ f = g \text{ almost everywhere,} \]

or \( f = g \text{ a.e.} \) as an abbreviation. If there is more than one measure around we may need to use the more specific

\[ f = g \mu\text{-a.e.} \]
If \( f : \mathbb{R} \to \mathbb{C} \) and if \( \sum_{s \in S} |f(s)|^2 < \infty \) we call \( f \) **square integrable** or **square summable**.

\( L^2(\mu) \) consists of equivalence classes of square integrable functions, where equivalent functions are those which are equal almost everywhere.

For reasons of custom and convenience, we conflate a function with its equivalence class.

This is harmless in our definitions and calculations because \( \mu \) (and the inner product defined by \( \mu \)) has no way to distinguish among functions in the same class.

For any subset \( A \) of \( \mathbb{R} \) let

\[
\chi_A : \mathbb{R} \to \{ 0, 1 \}
\]

denote the function that evaluates to 1 on members of \( A \) and 0 otherwise. \( \chi_A \) is called the **characteristic function for \( A \)**.

So if \( f \in L^2(\mu) \) then \( f = \sum_{s \in S} f(s) \chi_{\{s\}} \).

In fact, the set of characteristic functions \( \{ \chi_{\{s\}} | s \in S \} \) forms an orthonormal basis for \( L^2(\mu) \).

Select a member \( g : \mathbb{R} \to \mathbb{C} \) and, for \( f \in L^2(\mu) \), consider the a.e.-defined function \( K_g(f) = gf \).

If \( g \) is in \( L^2(\mu) \) then \( K_g(f) \in L^1(\mu) \) by the BCS inequality.

If \( g \) is bounded on the support of \( \mu \) then \( K_g(f) \in L^2(\mu) \cap L^1(\mu) \).

\( K_g \) is called a **multiplication operator**.

An important example is when \( g \) is the constant function 1. \( K_1(f) = f \) for all \( f \).
Another multiplication operator, which will be key for us, is that corresponding to the identity function

\[ \text{Id}: \mathbb{R} \to \mathbb{R} \quad \text{given by} \quad \text{Id}(x) = x. \]

If \( S \) is a bounded set, then the multiplication operator \( K_{\text{Id}} \) is bounded, and

\[ K_{\text{Id}}(f) = \sum_{s \in S} s f(s) \chi_{\{s\}}. \]

For obvious reasons, it is customary to use \( K_{\text{Id}}(f) = xf(x) \) to denote the output function, which is fine as long as you don’t forget what it means.

Define \( U: L^2(\mu) \to \mathbb{C}^3 \) by

\[ U(f) = f(7)P_1 + f(5)P_2 + f(3)P_3. \]

\( U \) is an invertible isometry between these two Hilbert spaces, and satisfies (and could have been defined by)

\[ U(\chi_{\{7\}}) = P_1, \quad U(\chi_{\{5\}}) = P_2, \quad U(\chi_{\{3\}}) = P_3 \]

and extending to all of \( L^2(\mu) \) by linearity.

The isometry \( U \) lets us associate \( x \) (or rather multiplication by \( x \)) down in \( L^2(\mu) \) with operator \( T \). All operations involving \( \mathbb{C}^3 \) or \( T \) can be carried out in \( L^2(\mu) \) using \( K_{\text{Id}} \).
**Finite-Dimensional Spectral Theorems**

\[ L^2(\mu) \leftrightarrow T \text{ (PART IV)} \]

![Diagram](image)

Then by continuity of \( U \) we can define \( f(T) \) for any \( f \in L^2(\mu) \).

---

**Multiplicity (PART I)**

The previous considerations are a complete rephrasing of the Spectral Theorem and the Functional Calculus for self-adjoint operators on a three dimensional space when the eigenspaces have dimension 1.

The theory of spectral multiplicity for these operators has an easy translation here to deal with possible eigenspace of dimension greater than one.

---

**Multiplicity (PART II)**

If \( \mu \) is a counting measure on \( \mathbb{R} \) with support \( S \) and \( n \geq 1 \) define 
\( L^1(\mu, \mathbb{C}^n) \) to be the set of those functions \( f: \mathbb{R} \to \mathbb{C}^n \) for which

\[ \int_{\mathbb{R}} \|f\| \, d\mu = \sum_{s \in S} \|f(s)\| < \infty. \]

For these functions we define \( \int_{\mathbb{R}} f \, d\mu \) to be \( \sum_{s \in S} f(s) \).

We define \( L^2(\mu, \mathbb{C}^n) \) to be the set of those functions \( f: \mathbb{R} \to \mathbb{C}^n \) for which

\[ \int_{\mathbb{R}} \|f\|^2 \, d\mu = \sum_{s \in S} \|f(s)\|^2 < \infty. \]

---

**Multiplicity (PART III)**

Note that \( L^2(\mu, \mathbb{C}^1) \) was the space we previously called \( L^2(\mu) \).

For each Dirac delta measure, or “function,” the space \( L^2(\delta_\lambda, \mathbb{C}^n) \) has dimension \( n \).

A specific basis consists of the \( n \) functions \( b_{\lambda,k}: \mathbb{R} \to \mathbb{C}^n \) which are 0 except at \( \lambda \) and for which \( b_{\lambda,k}(\lambda) = e_k \), for \( 1 \leq k \leq n \).

Any function \( f: \mathbb{R} \to \mathbb{C}^n \) is defined as a member of \( L^2(\delta_\lambda, \mathbb{C}^n) \) by the single vector

\[ f(\lambda) = v^1 e_1 + \cdots + v^n e_n. \]

Then, again as members of \( L^2(\delta_\lambda, \mathbb{C}^n) \), we have

\[ f = v^1 b_{\lambda,1} + \cdots + v^n b_{\lambda,n}. \]
**Multiplicity (Part III)**

$L^2(\mu, \mathbb{C}^n)$ is a Hilbert space with inner product

$$\langle f, h \rangle = \int_{\mathbb{R}} \langle f(s), h(s) \rangle \, d\mu = \sum_{s \in S} \langle f(s), h(s) \rangle$$

$$= \sum_{s \in S} \sum_{j=1}^n f_j(s) \overline{h_j(s)}.$$  

**Multiplicity (Part IV)**

Now suppose $T : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ is a self-adjoint operator, as before, but which only has two distinct eigenvalues 5 and 3.

Suppose the eigenspace for 5 has dimension 1, so the eigenspace for 3 has dimension 2.

Let $P$ be the unitary change of basis matrix with columns $P_1$, $P_2$ and $P_3$ where $P_1$ is an eigenvector for eigenvalue 5 and the remaining columns two span the eigenspace for eigenvalue 3.

$$\mathbb{C}^3 = \mathbb{C}P_1 \oplus (\mathbb{C}P_2 \oplus \mathbb{C}P_3).$$

**Multiplicity (Part V)**

Let $L = L^2(\delta_5, \mathbb{C}^1) \times L^2(\delta_3, \mathbb{C}^2)$.

$L$ consists of ordered pairs of functions $(f, g)$ where $f$ is complex valued and $g$ has values in $\mathbb{C}^2$.

It has a basis consisting of the three function-pairs

$$(\chi_5, 0) \text{ and } (0, b_{3,1}) \text{ and } (0, b_{3,2})$$

$L$ is a Hilbert space of dimension 3 with inner product:

$$\langle (f, g), (h, k) \rangle = \langle f, h \rangle + \langle g, k \rangle$$

$$= f(5)\overline{h(5)} + g^1(3)\overline{k^1(3)} + g^2(3)\overline{k^2(3)}.$$  

**Multiplicity (Part VI)**

Define $U : L \rightarrow \mathbb{C}^3$ by

$$U(\chi_5, 0) = P_1 \text{ and } U(0, b_{3,1}) = P_2 \text{ and } U(0, b_{3,2}) = P_3$$

extending to all of $L$ by linearity.

$U$ is an isometry onto $\mathbb{C}^3$, just as before.

If $h : \mathbb{R} \rightarrow \mathbb{C}$ define, for each $(f, g) \in L$ the pair

$$K_h(f, g) = (hf, hg).$$

$K_h$ is called a multiplication operator on $L$, and once again the multiplication operator $K_{Id}$ of the identity map $Id(x) = x$ corresponds to the operator $T$, and this allows us to define $h(T)$, for a wide class of functions $h$, as the operator $h(T)$ that corresponds to the multiplication operator $K_h$. 