Hilbert Space Methods Used in a First Course in Quantum Mechanics An Outline, With Quite a Few Details, of Pertinent Facts About Linear Spaces (Part 1)

Larry Susanka

Mathematics Program
Bellevue College

January 10, 2013

Why Are We Here?

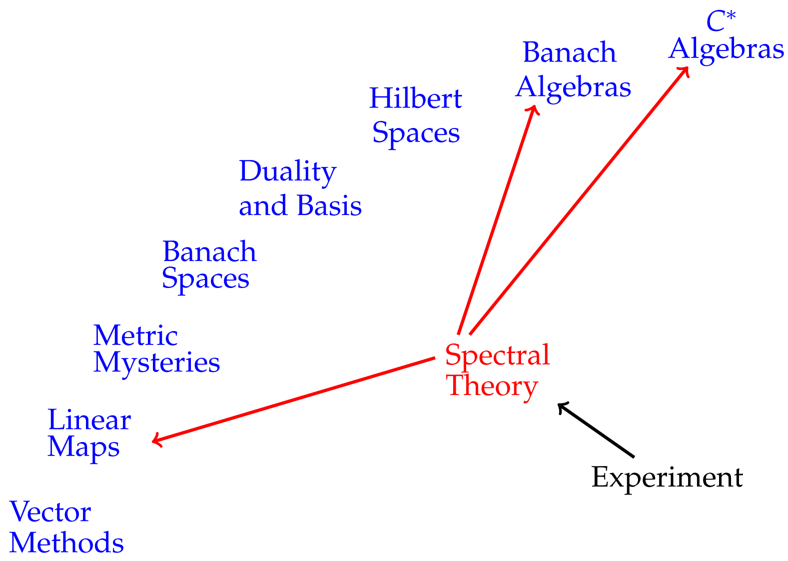
- We will examine a listing of main theorems and definitions of those purely mathematical results necessary to phrase and understand elementary quantum mechanics in the modern style.
- Frequently, there will be some feature of a proof itself which is useful to emphasize here and in that case we do give a proof, but otherwise reference to a proof will suffice.
- We will use sufficient generality so that the vector spaces used in the applications we have in mind can be seen in their natural mathematical context.
- We will strive to be very clear about what the theorems actually state: what the assumptions are, what the conclusions are ... exactly.

Quote From Wikipedia

Along the way, I personally would really really like to understand the following snippet, how it is implemented, how mathematical properties are to be interpreted physically, taken from a Wikipedia article on C^*-algebra:

> *In quantum mechanics, one typically describes a physical system using a C^*-algebra $\mathbf{A}$ with unit element; the self-adjoint elements of $\mathbf{A}$ (elements $\mathbf{x}$ with $\mathbf{x}^* = \mathbf{x}$) are thought of as the observables, the measurable quantities, of the system.*
>
> *A state of the system is defined as a positive functional ϕ defined on $\mathbf{A}$ (that is, a complex-linear functional ϕ for which $\phi(\mathbf{x}\mathbf{x}^*) \geq 0$ for all $\mathbf{x}$ in $\mathbf{A}$) such that $\phi(1) = 1$.*
>
> *The expected value of the observable $\mathbf{x}$, if the system is in state ϕ, is then $\phi(x)$.*

The Arc of Quantum Wisdom



Field Axioms

- A field is a nonempty set $\mathbb{F}$ ...
- together with two commutative and associative binary operations $\cdot$ and $+$ ...
- both of which have identities, denoted 0 and 1, respectively.
- Further, each nonzero element of $\mathbb{F}$ has a multiplicative inverse, and every element has an additive inverse.
- Finally, the usual distributive property of multiplication over addition holds.

A field is an object that acts very much like the real numbers $\mathbb{R}$ with ordinary arithmetic, and we will use the customs of order-of-operations, exponentiation and so on from arithmetic in $\mathbb{R}$.

Examples of Fields

- The rational numbers $\mathbb{Q}$.
- $\mathbb{Q}(\sqrt{2})$, which consists of all numbers of the form $a + b\sqrt{2}$ for rational a and b, is a field.
- $\mathbb{C}$, the complex numbers with complex arithmetic, is a field.
- If p is a prime, the set $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ with "mod p arithmetic" (that is, remainder after division by p) is a finite field whose order (that is, cardinality or size) is p.

Our Focus

- We will limit our *numbers* to be either the real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$.
- When we need to refer to a generic field we use the symbol $\mathbb{F}$. It will be one of these two.
- We will *not* use skew fields, finite fields, the field of rational numbers or other objects such as rings as a source for *numbers* here.

Vector Space Axioms I

A **vector space over a field** $\mathbb{F}$ is a nonempty set V together with two operations called **vector addition** and **scalar multiplication** that satisfy a collection of **ten properties.** Vector addition *acts on* pairs of members of V. Scalar multiplication *acts on* a pair one of which is a number and the other of which is a member of V.

1 and 2 We require both vector addition and scalar multiplication to be **closed in V**, and by that we mean that the result of applying these operations on members of V or numbers *always* produces a member of V. You cannot leave V by doing these operations to members of V and numbers.

3 There must be a member of V, always denoted $\mathbf{0}$ and called **the zero vector**, for which $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all members $\mathbf{v}$ of V. You distinguish this member of V from the number 0 by context.

Vector Space Axioms II

4 For each $\mathbf{v}$ in V there must be a member $\mathbf{u}$ of V for which $\mathbf{v} + \mathbf{u} = \mathbf{0}$. $\mathbf{u}$ can be denoted $-\mathbf{v}$, and is called the negative or opposite of $\mathbf{v}$.

5 and 6 Vector addition must be commutative and associative:

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \quad \text{and} \quad (\mathbf{v} + \mathbf{w}) + \mathbf{u} = \mathbf{v} + (\mathbf{w} + \mathbf{u})$$

for any members $\mathbf{v}$, $\mathbf{u}$ and $\mathbf{w}$ of V.

Vector Space Axioms III

7 and 8 The two distributive laws must hold:

$$(r + s)\mathbf{v} = r\mathbf{v} + s\mathbf{v} \quad \text{and} \quad r(\mathbf{v} + \mathbf{w}) = r\mathbf{v} + r\mathbf{w}$$

for all numbers r and s and any members $\mathbf{v}$ and $\mathbf{w}$ of V.

9 and 10 Last, $(rs)\mathbf{v} = r(s\mathbf{v})$ and $1\mathbf{v} = \mathbf{v}$ for all $r, s \in \mathbb{F}$ and all $\mathbf{v}$ in V.

Linear Combination and Span

- If V is a vector space over field $\mathbb{F}$ a **linear combination** of vectors is a *finite* sum (of any length) $a^1 v_1 + \cdots + a^n v_n$ where all the a^i are from $\mathbb{F}$ and all the v_i are from V.

  If there is more than one field around, it might be necessary to specify which one is intended to supply the coefficients: viz. $\mathbb{R}$-*linear combination.*
- A set S of vectors from the vector space V is said **to span** (this is a verb) V if every member of V can be written as a linear combination of members of S.
- If S is a set of vectors from V, the **span of S**, denoted $\mathbf{Span}(\mathbf{S})$ (this is a noun), is the set of all *finite* linear combination of members of S.

Independence

- A nonempty set S of vector space V is called **linearly independent** if whenever $s_1, \ldots, s_n$ is a *finite* list of distinct members of S, a sum

$$a^1 s_1 + \cdots + a^n s_n$$

where all the a^i are from $\mathbb{F}$ can never equal the zero vector unless *all* the a^i are 0.
- This is equivalent to saying that no member of S can be written as a linear combination of *other* members of S. It is also equivalent to saying that every member of $Span(S)$ can be written in only *one* way as a linear combination of members of S.
- If nonempty S is *not* linearly independent it is called **linearly dependent**.

Subspace and Basis

- A nonempty subset S of V is called a **subspace** of V if it is itself a vector space over the same field with the operations inherited from V.
- *To check that a subset is a subspace we need only demonstrate closure of the two vector operations in S.* The other eight properties will follow from the (already checked, presumably) fact that V itself is a vector space.
- A linearly independent set of vectors that spans the vector space V is called a **basis** for V.

Ordered Basis

- Sometimes a basis as defined here is called a **Hamel basis** to distinguish it from a different notion of basis we will use later in the context of Hilbert Spaces.
- Very often it is useful to *line up* the members of a basis in a specific order, and one then speaks of an **ordered basis**: a basis with some explicit order prescribed. It follows from one of the foundational axioms of mathematics (the axiom of choice) that any set can be ordered, lined up as the members of the natural numbers $\mathbb{N}$ are, so that every nonempty subset has a least member. In practice there is no need to invoke the axiom: bases are usually presented with members in easily-ordered form.

The Simplest Examples

- The set containing just the zero vector is a subspace of every vector space. Any one-element vector space is called a **trivial** vector space.
- Also (obvious but worth mentioning) every vector space is a subspace of itself. A **proper subspace** is a subspace smaller than this.
- $\mathbb{F}$ itself is a very simple vector space over $\mathbb{F}$. Also $\mathbb{C}$ is a vector space over $\mathbb{R}$.

The Product Space

- Many vector spaces are given as function spaces, and we require a compact notation to specify name of function, domain and range.

  We indicate that F is a function whose domain is the (nonempty) set V and whose range is *contained in* set W by

$$\mathbf{F}\colon \mathbf{V} \to \mathbf{W}.$$

- Suppose M is any nonempty index set and let V_m be any vector space over $\mathbb{F}$ for each $m \in M$.

  $\prod_{\mathbf{m}\in\mathbf{M}} \mathbf{V_m}$ is defined to be the set of all functions

$$f\colon M \to \bigcup_{m\in M} V_m \ \text{ for which } \ f_m \in V_m \ \text{ for every } m \in M.$$

The Pointwise Operations

We define **pointwise addition** and **pointwise scalar multiplication** on this set in the obvious way:

If $f, g \in \prod_{m\in M} V_m$ and $c \in \mathbb{F}$

define $f+g$ to be the function $(f+g)_m = f_m + g_m$

and $(cf)_m = cf_m$ for all $m \in M$.

$\prod_{m\in M} V_m$ is a vector space with these pointwise operations.

This space is called a **product vector space.**

All of the vector spaces you will actually use for something, both here and almost anywhere else, are given explicitly as subspaces of product spaces.

Ordered Products

If V_1 and V_2 are two vector spaces over field $\mathbb{F}$ the product space

$$\prod_{m\in\{1,2\}} V_m \ \text{ is normally denoted } \ V_1 \times V_2.$$

There is, however, just a little more going on in the second notation. There is a specific order on this particular index set, and we honor that fact by listing the vector space factors in subscript order. Many factor spaces have no obvious or natural (or intended) order on their indices, while in others an order or some other structure on the indices is crucial.

Product With Constant Range Space

Suppose M is any nonempty index set and V is any vector space over $\mathbb{F}$. Define $\mathbf{V^M}$ to be the set of functions with domain M and range in V. This is a special case of the example above, where $V_m = V$ for all m.

V^M is a vector space with pointwise operations.

The vector space $\mathbb{R}^{\{1,2,\ldots,n\}} = \mathbb{R} \times \cdots \times \mathbb{R}$ is an example of this type of product, normally denoted $\mathbb{R}^n$.

$\mathbb{R}^{\mathbb{R}}$ and $\mathbb{R}^{[0,1]}$ are two more: the real valued functions defined on the real line and the unit interval, respectively.

The continuous functions, the differentiable functions, the integrable functions and polynomials (with appropriate domain) are subspaces of these vector spaces.

Matrix Spaces

- The set of matrices of a specific shape with entries in $\mathbb{F}$ with usual operations forms a vector space. We denote these $m \times n$ matrices by $\mathbb{M}_{m\times n}(\mathbb{F})$. Matrices are the set of $\mathbb{F}$-valued functions with domain consisting of the block of integer pairs

$$\{ (a, b) \mid a, b \text{ are integers with } 1 \le a \le m \text{ and } 1 \le b \le m \}.$$

  The big rectangular symbol used to denote this function is convenient but is nothing more than a visualization of this by placing the matrix entries on grid points labeled by these pairs of integers.
- Matrices populate *many* important subspaces.

Sequences

- A real sequence is a function $f : \mathbb{N} \to \mathbb{R}$ and the set $\mathbb{R}^{\mathbb{N}}$ of all such sequences is a real vector space with pointwise operations. The set of convergent real sequences is a subspace.
- More generally, if V is a vector space of functions then $V^{\mathbb{N}}$, the set of *sequences* of functions, is too. If there is a concept of convergence you have in mind for these sequences of functions, the set of convergent sequences can form a subspace of $V^{\mathbb{N}}$.

Linear Maps

Suppose F is a function whose domain is the vector space V and whose range is contained in vector space W, both over the same field. This function is called **linear** (or a **linear transformation**) if for v, u in V and field member k we have

$$F(v + ku) = F(v) + kF(u).$$

When a function F is linear and no confusion can result the notations $F(v)$ and $F\,v$ will be used interchangeably.

Linear Maps and Bases

Every linear function $F : V \to W$ is determined by what it does to a basis of its domain. Specifically, if $v = v^1\, b_1 + \cdots + v^n\, b_n$ for numbers v^i and basis vectors b_i of V then

$$F(v) = F\left(\sum_{i=1}^{n} v^i\, b_i\right) = \sum_{i=1}^{n} v^i\, F(b_i).$$

If you know all the $F(b_i)$ then you know any $F(v)$.

Conversely, you can define a function in any way you like on a basis of V and extend by linearity to all of V.

We will see, however, that in infinite dimensional spaces with more structure the resulting functions may lack the important property of *continuity*. Some choices in such situations are forbidden if we require this property.

Types of Linear Maps

A generic linear transformation is also called a (vector space) **homomorphism.** The set of linear transformations from $\mathbb{F}$-vector space V to $\mathbb{F}$-vector space W is denoted $\mathbf{Hom}_{\mathbb{F}}(\mathbf{V},\mathbf{W})$. It is a vector space too, a (small) subspace of the $\mathbb{F}$-vector space W^V.

When $W = V$ a linear transformation is often called a **linear operator**, though it should be pointed out that this is by no means universal: in many sources the terms linear transformation and linear operator are synonymous.

When W is just the number space a linear transformation is also called a **linear functional.**

The collection of all these functionals is called the **algebraic dual** of V and denoted $\mathbf{V}^*$. If it is necessary to specify the field, the notation $\mathbf{V}^*_{\mathbb{F}}$ will be used.

Isomorphism and Identification

Linear $F\colon V \to W$ is called an **isomorphism**, and the two spaces are called **isomorphic**, if F is invertible as a function. The inverse of an isomorphism is also linear, and so is itself an isomorphism.

Two vector spaces over the same field are isomorphic exactly when they have the same dimension. Explicit isomorphisms are sometimes used to encourage identification of the two spaces involved: for raw “vector space” purposes they are, essentially, identical. Usually there must be more than shared dimension to make an identification worth the effort it takes to establish it.

Domain and Kernel and Image

For linear $F\colon V \to W$ we define $\mathcal{K}\mathbf{er}(\mathbf{F})$ to be the set of those $v \in V$ for which $Fv = 0$. It is called the **kernel** or **nullspace** of F and is a subspace of the domain space V.

We will have occasion to consider functions whose domains are various subspaces of V and not all of V. In that context we will refer to the domain of definition of a linear function F by $\mathcal{D}\mathbf{om}(\mathbf{F})$.

We define $\mathcal{R}\mathbf{an}(\mathbf{F})$ to be the set of all $F(v)$ for $v \in V$. This set is called the **image** of F and is also a subspace of W.

Rank and Nullity

$\dim(\mathcal{K}er(F))$ is called the **nullity** of F and this cardinal number is denoted $\mathbf{nullity}(\mathbf{F})$.

$\dim(\mathcal{R}an(F))$ is called the **rank** of F, and denoted $\mathbf{rank}(\mathbf{F})$.

It is a fact that

$$rank(F) + nullity(F) = \dim(\mathcal{D}om(F))$$

which can be quite useful, particularly when all these spaces have finite dimension.

THE EVALUATION MAP

The algebraic dual V^* of V is, itself, a vector space. So *it too* has an algebraic dual, $(V^*)^*$.

Define the **evaluation map** $E\colon V \to (V^*)^*$ by

$$E(x)(f) = f(x) \quad \text{for each } x \in V \text{ and } f \in V^*.$$

E is linear, and also one-to-one: that is, $E(x) = E(y)$ exactly when $x = y$. So if E is onto then it is invertible and an isomorphism. In that case, $(V^*)^*$ can be identified with (i.e. it *is*) V.

If V is finite dimensional, V and V^* have the same dimension. And it follows that $(V^*)^*$ has the same dimension as does V. So E must be onto. The infinite dimensional case is much more delicate and we will consider the extent to which we can recover this important identification later.

EIGENVALUES

A linear operator $G\colon V \to V$ can have **eigenvalues** and associated **eigenvectors.** A nonzero vector v is called an eigenvector of G for eigenvalue λ if $G(v) = \lambda v$.

Eigenvectors for **different** eigenvalues cannot be dependent: a set consisting of a finite number of eigenvectors for **different** eigenvalues is linearly independent.

QUOTIENT SPACES

If N is a subspace of vector space V and $v \in V$ we define $v + N$ to be the "translates" of N by v. Specifically,

$$v + N = \{\, v + n \mid n \in N \,\}.$$

We now define the **quotient space** of V by N to be the set of all these translates of N. We use the notation $\mathbf{V}/\mathbf{N}$ to denote this set of translates.

V/N has the structure of a vector space with operations, for all $a \in \mathbb{F},\ v, w \in V$, given by

$$a(v + N) = (av) + N \quad \text{and} \quad (v + N) + (w + N) = (v + w) + N.$$

The cardinal numbers of the dimensions combine as

$$\dim(V/N) + \dim(N) = \dim(V).$$

FUNCTIONS INDUCED ON A QUOTIENT SPACES

If $f\colon V \to W$ is linear and if subspace N of V is contained in $\mathcal{K}er(f)$ then f induces a linear function

$$\tilde{f}\colon V/N \to W$$

given by $\tilde{f}(v + N) = f(v)$.

Sum and Direct Sum

If Y and Z are two subspaces of vector space X, we write $\mathbf{Y+Z}$ to denote the vector subspace of X spanned by the vectors in Y and Z. We call this vector space **the sum of Y and Z.**

If, in addition, $Y \cap Z = \{0\}$ we write $\mathbf{Y \bigoplus Z}$ and call this **the direct sum of Y and Z.** The importance of direct sum is that any vector in $Y \bigoplus Z$ can be written in a unique way as $y+z$ where $y \in Y$ and $z \in Z$.

A basis of $Y \bigoplus Z$ can be given that is a disjoint union of a basis for Y and a basis for Z.

We let $\mathbf{X-Y}$ denote the members of X which are not in Y. If $w \in X - Y$ we define $\mathbb{F}\mathbf{w}$ to be the one dimensional subspace of X generated by w. Then $Y \cap \mathbb{F}w = \{0\}$ and so the sum $Y + \mathbb{F}w$ is a direct sum $Y \bigoplus \mathbb{F}w$.

Restriction and Extension

If $A \subset B$ for nonempty set A and $F\colon B \to C$ we will use the notation $\mathbf{F|_A}$ to denote the function whose domain is A and which agrees with F for all members of A.

If $G = F|_A$ for some nonempty set A then F is called **an extension of G** and G is called **the restriction of F to A.**

Any linear $F\colon Y \bigoplus Z \to W$ defines, by restriction, unique linear maps $F|_Y\colon Y \to W$ and $F|_Z\colon Z \to W$.

Conversely, any two linear functions $H\colon Y \to W$ and $K\colon Z \to W$ can be used to create a function $F\colon Y \bigoplus Z \to W$ by $F(x) = H(y) + K(z)$ where $x = y + z$ for $y \in Y$ and $z \in Z$.

Summary

Fields

$\mathbb{R}$ and $\mathbb{C}$ (Generic: $\mathbb{F}$)

$F\colon A \to B$

Vector Space and Subspaces

$\prod_{m\in M} V_m$
with *pointwise operations*

Examples:
V^n $V^{\mathbb{N}}$ $V_1 \times V_2$ $\mathbb{M}_{m\times n}(\mathbb{F})$
$\mathbb{R}^{\mathbb{R}}$ $\mathbb{R}^{[0,1]}$

Basis and dimension

Linear Maps: homomorphism
$\mathcal{K}\mathbf{er}(\mathbf{F})$ $\mathcal{D}\mathbf{om}(\mathbf{F})$ $\mathcal{R}\mathbf{an}(\mathbf{F})$

$Hom_{\mathbb{F}}(V, W)$
(small) subspace of W^V
operator, isomorphism

Functionals: *algebraic* dual V^*

Evaluation $E\colon V \to (V^*)^*$

Eigenvalues

Quotient and Direct Sum