# SOME BASIC STUFF ABOUT LINEAR OPERATORS AND VECTOR SPACES

*Robert H. Stolt*

*Contents*

### *Vector Spaces*

In the seismic experiment we deal with fields that are functions of space and time. These fields may be scalar (e.g. pressure) or vector (e.g. displacement), so, depending on how many dimensions are being ignored, we may find functions of as little as one continuous variable (e.g. x) or as many as four continuous variables (x,y,z,t) and one discrete variable (e.g. j, where j = 1 indicates displacement along the x-axis, j = 2 the y, and j = 3 the z). Often the continuous variables are discretized by sampling the functions at discrete intervals and restricting attention to a finite space-time region, in which case we deal with functions of up to five discrete variables.

Whatever the exact form of our functions [say $\psi_j(x,t)$] it is convenient to think of them as ***vectors***. If we call $\psi$ the vector corresponding to the function $\psi_j(x,t)$, then $\psi_j(x,t)$ is the component of $\psi$ at the particular point $(j,x,t)$. Each distinct choice of values for j, x, and t gives a different component of $\psi$, and the collection of all such components is the vector itself.

Our vector $\psi$ will normally reside in a ***vector space***, which is a collection of functions of the same variables. A ***linear vector space*** has the nice property that if $\psi$ and $\phi$ qualify for membership, then any linear combination of them[1] (e.g. $a\psi + b\phi$, where a and b are complex numbers) does too.

### *Hilbert Spaces*

Normally, the vectors we like will live in a ***Hilbert space***, which is a vector space with a few additional properties:

(1) In a Hilbert space, an ***inner product*** is definable for any pair of vectors. For example, we may define the inner product $\langle\phi|\psi\rangle$ between the pair of vectors $\psi$, $\phi$ to be

$$\langle\phi|\psi\rangle = \sum_j \int d^3x \int dt\ \phi_j^*(x,t)\ \psi_j(x,t)$$

[1] With our definition of a vector, addition and scalar multiplication should be clearly definable operations.

(* indicates complex conjugate.) That's not the only possible way to define an inner product; for example, we might want to throw some weight function $w_j(x,t)$ into the integral; but however we define it, the inner product should have the property

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$$

As in Euclidian space, the inner product measures the degree to which two vectors point in the same direction.

(2) A Hilbert space has a definable ***norm*** which measures the size of vectors. If we define the norm $||\psi||$ of $\psi$ as

$$||\psi|| = (\langle\psi|\psi\rangle)^{\frac{1}{2}}$$

$$= \left[ \sum_j \int d^3x \int dt \; |\psi_j(x,t)|^2 \right]^{\frac{1}{2}}$$

then for a vector $\psi$ to qualify for membership it must be square integrable.[2]

(3) A Hilbert space (a ***separable*** one, anyway) has at least one countable orthonormal ***basis***. This means that there is a set of vectors $\psi_i$ which are all orthogonal to each other and have unit norm

$$\langle\psi_i|\psi_i\rangle = \delta_{ij}$$

in terms of which any vector $\psi$ in our space may be expanded:

$$\psi = \sum_{i=1}^{\infty} \alpha_i \psi_i = \sum_i \psi_i \langle\psi_i|\psi\rangle$$

---

[2]The requirement that vectors in a Hilbert space ***have a finite norm*** actually excludes some of the most useful vectors, namely those representing plane waves. Consequently, we often allow them in the back door to do our dirty work for us, then when they are done, ask them to leave quietly.

The coefficients $\alpha_i$, which are just the inner product of $\psi$ with $\psi_i$, are the component of $\psi$ at each "point" $\psi_i$.

It may be a bit surprising that $\psi$ should have a countable (infinite) number of components in any basis, since, for continuous functions, the number of space-time points is uncountably infinite. By restricting our space to vectors of finite norm, however, we have effectively reduced its dimension to a countable infinity.

(If the dimension of our space is finite, then the property of separability holds by definition. Mathematicians generally like to think of Hilbert spaces as infinite dimensional, but we have no need to be that exclusive.)

(4) Hilbert spaces satisfy the Cauchy convergence criterion. Given an infinite sequence of vectors $\phi_m$, $m=1,2,3,\ldots$, then $||\phi_m - \phi|| \to 0$ as $m \to \infty$ iff $||\phi_m - \phi_n|| \to 0$ as $m,n \to \infty$ independently. This nice property allows us to take limits of vectors more or less with impunity, which at least partially justifies being so snobbish. Once the unnormalizable vectors are let in, one has to be more careful about taking limits.

### *Subspaces*

Our Hilbert space will contain many subsets which are themselves vector spaces. These subsets, called ***subspaces***, require only that if any two vectors belong to one, so must all linear combinations of them. In particular, an orthogonal basis forms a set of little one-dimensional subspaces, each containing all the vectors proportional to one of the basis vectors. One sometimes speaks of a Hilbert space being the ***direct sum*** of these basis subspaces, since any vector in the Hilbert space may be expressed as a sum of vectors, one from each basis subspace.

*Dirac Notation (bras + kets)*

It is often convenient to write vectors in the form $|\psi\rangle$. The first reason for doing this is notational: distinguishing between vectors and numbers becomes a snap. The second reason relates to inner products. If $|\psi\rangle$ is an ordinary column vector (sometimes called a ***ket*** ), we can define $\langle\psi|$ to be the corresponding complex conjugated row vector (or ***bra***). An inner product between $|\psi\rangle$ and $|\phi\rangle$ is just $\langle\psi|\phi\rangle$ (or $\langle\psi|\cdot|\phi\rangle$ - get it?) A third reason is convenience: the vector $\psi_i$ can be represented by the shorthand $|i\rangle$ as easily as $|\psi_i\rangle$ without loss of meaning. For example, the plane wave function $[e^{i(p\cdot x-\omega t)}]/[(2\pi)^2]$ can be represented by $|p,\omega\rangle$ as well as by anything else. [The $(2\pi)^2$ was put in the denominator of the plane wave function to give the corresponding vector a delta function normalization:

$$\langle p',\omega'|p,\omega\rangle = \delta(p-p')\,\delta(\omega-\omega')]$$

Fourth, components of vectors are naturally viewed as inner products with this notation. To form the function $\psi_j(x,t)$ from $|\psi\rangle$, the temptation is irresistible to define a set of vectors $|j,x,t\rangle$ which point to a specific space-time location and direction. They are defined by the requirement that $\langle j,x,t|\psi\rangle = \psi_j(x,t)$. The normalization of these vectors is easily seen to be

$$\langle j',x',t'|j,x,t\rangle = \delta_{jj'}\,\delta(x-x')\,\delta(t-t')$$

Like the plane wave vectors, they have infinite norms, so cannot belong to Hilbert space. (The Hilbert space snobs have taken to calling such vectors ***improper***, as if they didn't really exist at all). Strangely enough, they do form a basis (though uncountably infinite) of Hilbert, in that every vector $|\psi\rangle$ in Hilbert can be decomposed into its elements $\langle j,x,t|\psi\rangle$. The parameters $(j,x,t)$ in the bra of $\langle j,x,t|\psi\rangle$ indicate which element of $|\psi\rangle$ we are looking at, so are analogous to the index in a finite dimensional vector.

The principle objections to Dirac notation are (1) it tends to make "improper" vectors look legitimate; and (2) there is absolutely no need to introduce row vectors (bras) into vector space formalism. (The idea of an

inner product as a row vector times a column vector is cute, but that's about all.) Bringing unnecessary objects into the formalism, so the argument goes, can only serve to complicate and confuse it.

Since I am indifferent to both objections, Dirac notation will be used frequently from this point onward.

### *Vector Representations*

A vector $|\psi\rangle$ can be considered an abstraction, or we may identify it with the set of its components [e.g. $\psi_j(x,t)$]. In general there will be many ways to decompose $|\psi\rangle$ (one for every basis), and we refer to each possible decomposition as a ***representation*** of $|\psi\rangle$. For example the function $\psi_j(x,t)$ is called the space-time representation of $|\psi\rangle$. The decomposition of $|\psi\rangle$ into its plane-wave Fourier components[3] $\psi_j(p,\omega)$ or $\langle j,p,\omega|\psi\rangle$ is referred to loosely as the frequency-wavenumber or f-k representation, or perversely as the wavenumber-frequency or k-f representation. The decomposition of $\psi$ into its components $a_i = \langle i|\psi\rangle$ with respect to the as yet unidentified countable basis which all nice, separable Hilbert spaces have, might be referred to as the i representation.

It is interesting to note that our scalar product, which was originally defined in the space-time representation, could just as easily have been defined in any other. For example,

---

[3] $\psi_j(p,\omega)$ is the quadruple Fourier transform of $\psi_j(x,t)$:

$$\psi_j(p,\omega) = \frac{1}{(2\pi)^2} \int d^3x \int dt\, e^{-i(p\cdot x-\omega t)} \psi_j(x,t)$$

or, in Dirac notation,

$$\langle j,p,\omega|\psi\rangle = \int d^3x \int dt\, \langle p,\omega|x,t\rangle \langle j,x,t|\psi\rangle$$

$$\langle\phi|\psi\rangle = \sum_j \int d^3p \int d\omega\, \phi_j^*(p,\omega)\, \psi_j(p,\omega)$$

$$= \sum_j \int d^3p \int d\omega\, \langle\phi|j,p,\omega\rangle\, \langle j,p,\omega|\psi\rangle$$

or, for that matter

$$\langle\phi|\psi\rangle = \sum_i \langle\phi|i\rangle\, \langle i|\psi\rangle$$

all represent the ***same*** inner product. Thus $\langle\phi|\psi\rangle$ depends on $|\psi\rangle$ and $|\phi\rangle$ only, and is not tied to any particular decomposition of $|\psi\rangle$ and $|\phi\rangle$, space-time, or whatever.

### *Linear Operators*

An ***operator*** is just something which changes one vector into another. To describe a process which changes the vector $|\psi\rangle$ into a new vector, say $|\phi\rangle$, we might define an operator $\hat{A}$ such that

$$\hat{A}|\psi\rangle = |\phi\rangle$$

To read this equation we should say, "$\hat{A}$, acting on $|\psi\rangle$, produces $|\phi\rangle$."

Examples of operators are easily found by looking at vectors in the space-time representation. One operator might correspond to time-differentiating the space-time function; another to multiplying it by $e^t$; and still another to squaring it. Many other operators exist which have no simple space-time representation, but all can be expressed in the abstract formalism.

Many operators cannot be defined for every vector in Hilbert. The operator $e^t$, for example, will give some vectors infinite norm, and differential operators are defined only on differentiable functions. The set of vectors on which an operator is defined is its ***domain***, and the set of vectors it can produce is its ***range***.

A ***linear operator*** has the following nice property: For $|\psi\rangle$, $|\phi\rangle$ in the domain of $\hat{A}$, for any two complex numbers a and b,

$$\hat{A}\,[a|\psi\rangle + b|\phi\rangle] \;=\; a\hat{A}|\psi\rangle + b\hat{A}|\phi\rangle$$

Differential operators and operators like $e^t$ which multiply a vector's space-time function by a fixed space-time function are good examples of linear operators. The operator which squares the space-time function of a vector may be a nice operator, but it is not linear. A particularly relevant linear operator is the differential operator

$$\nabla \frac{1}{\rho(x)} \nabla \;-\; \frac{1}{K(x)} \frac{\partial^2}{\partial t^2}$$

which is found in the scalar wave equation. Neither the domain nor the range of a linear operator need be the whole Hilbert space, but both the domain and range qualify as subspaces.

***What on earth is a ket-bra?*** One particularly nice (or horrid, depending on your point of view) feature of Dirac formalism is the ease in which one can define funny little operators which look like this:

$$|\psi\rangle\,\langle\phi|$$

Since it may not be perfectly clear what this creature is, let me explain. In a finite dimensional space, $|\psi\rangle\,\langle\phi|$ would be a dyad or outer product of a column vector times a row vector. In Hilbert space, its meaning becomes clear as soon as we act on some vector with it:

$$[|\psi\rangle\langle\phi|]\;|\gamma\rangle \;=\; |\psi\rangle\langle\phi|\gamma\rangle$$

Its output is proportional to the vector $|\psi\rangle$, the constant of proportionality being the inner product between $|\phi\rangle$ and the input vector. Operators like $|\psi\rangle\langle\phi|$ are manifestly linear, and in fact are very well behaved as operators go.

### *Miscellaneous Properties of Linear Operators*

If $\hat{A}$ and $\hat{B}$ are both defined on a vector $|\psi\rangle$, then we can define their sum $(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle$. From this definition it is obvious that $\hat{A} + \hat{B} = \hat{B} + \hat{A}$.

The product $\hat{A}\hat{B}$ of two operators is easily definable as $\hat{A}\hat{B}|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle)$, from which it is clear that $\hat{A}\hat{B}\hat{C} = \hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$ and so on. It is ***not*** true in general, however, that $\hat{A}\hat{B} = \hat{B}\hat{A}$. [For example, $(\partial/\partial x)f(x) \neq f(x)(\partial/\partial x)$.] Operators which do obey this property are said to ***commute***. It should be kept constantly in mind that most operators do not commute and that in any product of operators the operator on the right is to be applied first.

### *Elements of Operators*

Like vectors, operators may be decomposed into a set of numbers or elements. If $\hat{A}$ acting on $|\psi\rangle$ yields a vector $\hat{A}|\psi\rangle$, then we may, for any vector $|\phi\rangle$, define the inner product $\langle\phi|\hat{A}|\psi\rangle$. (Dirac notation becomes a bit confusing at this point, since the operator $\hat{A}$ is sandwiched symmetrically between the two vectors, looking equally anxious to act on either one. One simply has to remember that it acts on the vector to its right.) Once $\langle\phi|\hat{A}|\psi\rangle$ has been evaluated for all possible $|\psi\rangle$ and $|\phi\rangle$, everything there is to know about $\hat{A}$ has been found. We say that $\hat{A}$ is completely determined by its elements $\langle\phi|A|\psi\rangle$.

### *Operator Representations*

There are a lot of vectors in Hilbert Space, so finding $\langle\phi|\hat{A}|\psi\rangle$ for all $|\phi\rangle$, $|\psi\rangle$ would appear to be a tedious business. Fortunately, it is not necessary to evaluate all the elements of an operator in order to determine it. In fact, we need only know the elements with respect to a single orthonormal basis. Suppose that the set of vectors $|i\rangle$, i=1,2,... are known, having the properties of orthonormality

$$\langle i|i'\rangle = \delta_{ii'}$$

and completeness

$$|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$$

for any $|\psi\rangle$. Then if $\langle i|\hat{A}|i'\rangle \equiv A_{ii'}$ is known for every $i$, $i'$, *any* matrix element of $\hat{A}$ may be constructed:

$$\langle\phi|\hat{A}|\psi\rangle = \sum_{i,i'} \langle\phi|i\rangle \langle i|\hat{A}|i'\rangle \langle i'|\psi\rangle$$

We call the collection of elements $\langle i|\hat{A}|i'\rangle$ to be the ***representation*** of $\hat{A}$ in the $i$-basis.

The space-time representation of an arbitrary operator $\hat{A}$ is the collection of elements

$$\langle j,x,t|\hat{A}|j',x',t'\rangle \equiv A(j,x,t|j',x',t')$$

As for any other basis, once these elements are known the operator is completely specified.

***The unit operator***. The nicest operator around is the unit operator $\hat{I}$, distinguished by the fact that it does absolutely nothing:

$$\hat{I}|\psi\rangle = |\psi\rangle \qquad \text{for any } |\psi\rangle$$

Its representation in the $i$-basis is

$$\langle i|\hat{I}|i'\rangle = \delta_{ii'}$$

Its space-time representation is

$$\langle j,x,t|\hat{I}|j',x',t'\rangle = \delta_{jj'}\,\delta(x - x')\,\delta(t - t')$$

***Local Operators and Differential Operators***

A class of operators which behave very nicely in spacial representations are the so-called ***local*** operators, which amount to multiplication by some function of position. Suppose $\hat{f}$ is the operator which maps space-time function $\psi_j(x,t)$ onto $f(x)\,\psi_j(x,t)$. The space-time representation of $\hat{f}$ would usually be said to be $f(x)$. In Dirac notation, we would write:

$$\langle j,x,t|\hat{f}|\psi\rangle \equiv f(x)\,\langle j,x,t|\psi\rangle \equiv f(x)\,\psi_j(x,t)$$

Strictly speaking, however, the space-time representation of $\hat{f}$ is the matrix of elements

$$\langle j,x,t|\hat{f}|j',x',t'\rangle \equiv f(x)\,\langle j,x,t|j',x',t'\rangle$$

$$\equiv f(x)\,\delta_{jj'}\,\delta(x-x')\,\delta(t-t')$$

The values of the function $f(x)$ are the ***diagonal*** elements of this matrix. Since all the other elements are zero, all the information about $\hat{f}$ is contained in the function $f(x)$ and we are justified in thinking of it as the space-time representation of $\hat{f}$.

A similar situation holds for differential operators. If $\hat{D}_t$ is the operator that effects a time derivative, we would say (loosely) that $\partial/\partial t$ is the time representation of $\hat{D}_t$, writing

$$\langle j,x,t|\hat{D}_t|\psi\rangle \equiv \frac{\partial}{\partial t}\,\langle j,x,t|\psi\rangle \equiv \frac{\partial}{\partial t}\,\psi_j(x,t)$$

More strictly, the space-time representation of $\hat{D}_t$ is the matrix of elements

$$\langle j,x,t|\hat{D}_t|j',x',t'\rangle \equiv \frac{\partial}{\partial t}\,\langle j,x,t|j',x',t'\rangle = \delta_{jj'}\,\delta(x-x')\,\frac{\partial}{\partial t}\,\delta(t-t')$$

though the former, more concise form is in practice preferable.

Differential operators are non-local in that their space-time matrices contain off-diagonal elements. All non-zero elements are very near the

diagonal, however, (the derivative of a delta function can be viewed as the limit of a bi-diagonal form), so the differential operators enjoy great simplicity of structure and form.

### *Operator Inverses*

An operator is just a mapping between two subspaces. If a particular mapping is unique (i.e., if

$$\hat{A}|\psi\rangle = \hat{A}|\phi\rangle \quad \text{only if} \quad |\psi\rangle = |\phi\rangle \ )$$

Then an inverse mapping can be defined:

$$\hat{A}^{-1}|\gamma\rangle = |\psi\rangle \quad \text{if} \quad \hat{A}|\psi\rangle = |\gamma\rangle$$

We then have

$$\hat{A}^{-1}\hat{A}|\psi\rangle = |\psi\rangle$$

for all $|\psi\rangle$ in the domain of $\hat{A}$, and

$$\hat{A}\hat{A}^{-1}|\gamma\rangle = |\gamma\rangle$$

for all $|\gamma\rangle$ in the range of $\hat{A}$. If (and only if) the domain and range of $\hat{A}$ are the whole Hilbert space, we can write

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{1}$$

Note that linear operators give unique mappings provided no non-zero vector maps onto zero so ***unique*** inverses are definable only for such operators. For the general operator we may be able to define many inverses. The inverse of a product of operators, all of which have inverses, is easily found: $(\hat{A}\hat{B})^{-1} = B^{-1}A^{-1}$.

### *Adjoints*

The ***adjoint*** of an operator is defined as its complex conjugate transpose. The adjoint $\hat{A}^\dagger$ of the operator $\hat{A}$ is most easily defined through its elements:

$$\langle\phi|\hat{A}^\dagger|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^*$$

Since the complex conjugation of an inner product flips the bra and ket, we can view the adjoint as explaining the effect of operators on bras:

$$\langle\psi|\hat{A}|\phi\rangle^* = \langle\psi|\cdot\hat{A}|\phi\rangle^* = \langle\phi|\hat{A}\cdot|\psi\rangle = \langle\phi|\cdot\hat{A}^\dagger|\psi\rangle$$

In fact, we could define the adjoint by saying that if $\langle\phi|$ is the bra corresponding to $|\phi\rangle$, then $\langle\phi|\hat{A}^\dagger$ is the bra corresponding to $\hat{A}|\phi\rangle$. That definition allows us to think of $\hat{A}$ in $\langle\psi|\hat{A}|\phi\rangle$ as operating in either direction. (It is at this point that many people condemn Dirac notation as hopelessly confusing.) As for the inverse, the adjoint of a product of operators is the product of the adjoints, ***in reverse order***: $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$.

The adjoint of a complex number is clearly its complex conjugate. A complex function of space and time has as its adjoint its complex conjugate function. To obtain the adjoint of a differential operator we integrate by parts the inner product, discovering, for example, that the adjoint of $\partial/\partial x$ is $-\partial/\partial x$.

### *Self-Adjoint (Hermitian) Operators*

An operator which is its own adjoint is said to be ***self-adjoint*** or ***Hermitian***. Examples of self-adjoint operators are real numbers, real space-time functions, and second-derivative operators. The scalar wave equation operator

$$\nabla\cdot\frac{1}{\rho(x)}\nabla-\frac{1}{K(x)}\frac{\partial^2}{\partial t^2}$$

is self-adjoint, which is fortunate, since self-adjoint operators have some

nice properties.

### *Unitary and Isometric Operators*

A ***unitary*** operator necessarily has all vectors in its domain and range, and has the additional nice property: its operation does not change the norm of vectors. That is, for $\hat{U}$ to be unitary, we must have

$$||\hat{U}\psi|| = ||\psi|| \quad \text{for all vectors } \psi$$

Since we often try to set things up so that the norm of a vector measures some conserved physical quantity, such as energy, the operators which describe physical processes are often unitary.

A unitary operator has the property $\hat{U}^{\dagger}\hat{U} = \hat{1} = \hat{U}\hat{U}^{\dagger}$.

An ***isometric*** operator also preserves the norm of vectors, and obeys $\hat{U}^{\dagger}\hat{U} = \hat{1}$. However, its range may be smaller than its domain, leading to $\hat{U}\hat{U}^{\dagger} \neq 1$. On a finite dimensional space, all isometric operators are in fact unitary.

### *Projectors*

Some of the nicest operators around are the ***projectors*** or projection operators. The projector onto a particular subspace is an operator which does nothing to the vectors in that subspace, but which wipes out (maps onto zero) any vector outside the subspace. For the general vector, which has components both inside and outside the subspace, the projector yields the component inside the subspace.

For any vector $|\phi\rangle$, we can define a projector

$$\hat{P}_{\phi} = \frac{|\phi\rangle\langle\phi|}{||\phi||^2}$$

which maps onto the little one-dimensional subspace of vectors proportional to $|\phi\rangle$. Obviously, $\hat{P}_{\phi}|\phi\rangle = |\phi\rangle$, and $\hat{P}_{\phi}|\psi\rangle = 0$ if $\langle\phi|\psi\rangle = 0$.

If the set $|i\rangle$, $i=1,2,\ldots$ is an orthonormal basis, then for each basis vector $|i\rangle$ we can define a projector

$$\hat{P}_i = |i\rangle\langle i|$$

The completeness relation for this basis

$$|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle \quad \text{for every} \quad |\psi\rangle$$

then becomes

$$|\psi\rangle = \sum_i \hat{P}_i |\psi\rangle \quad \text{for every} \quad |\psi\rangle \quad \text{or} \quad \sum_i \hat{P}_i = \hat{1}$$

Likewise, for the space-time basis $|j,x,t\rangle$ we can define projectors (albeit "improperly")

$$\hat{P}(j,x,t) = |j,x,t\rangle\langle j,x,t|$$

and note that

$$\sum_j \int d^3x \int dt\, |j,x,t\rangle\langle j,x,t| = \sum_j \int d^3x \int dt\, \hat{P}(j,x,t) = \hat{1}$$

*Eigenvectors and Eigenvalues*

A vector $|\psi\rangle$ is said to be an eigenvector of an operator $\hat{A}$ if

$$\hat{A}|\psi\rangle = a|\psi\rangle \quad (a \text{ a number})$$

i.e., if $\hat{A}$ maps $|\psi\rangle$ onto a vector proportional to $|\psi\rangle$. The number $a$ (in general, complex) is called an eigenvalue of $\hat{A}$.

***Examples***: Projectors have eigenvalues 1 and 0: their eigenvectors are vectors in or orthogonal to their subspace. The constant velocity-density wave-equation operator

$$\nabla^2 - \frac{1}{v^2}\frac{\delta}{\delta t^2}$$

has as (improper) eigenvectors plane waves $e^{i(\mathbf{p}\cdot\mathbf{x}-\omega t)}$ with corresponding eigenvalues $\omega^2/v^2 - p^2$. The solutions to the source-free wave equation are those plane waves with eigenvalue zero.

The eigenvalues of a self-adjoint operator are obviously real. Moreover, two eigenvectors of a self-adjoint operator with different eigenvalues are necessarily orthogonal[4]: If $\hat{L}|\ell\rangle = \ell|\ell\rangle$, and $\hat{L}|\ell'\rangle = \ell'|\ell'\rangle$, then $\langle\ell'|\hat{L}|\ell\rangle = \ell\langle\ell'|\ell\rangle = \ell'\langle\ell'|\ell\rangle$ which implies that $\langle\ell'|\ell\rangle = 0$ if $\ell\neq\ell'$.

This property may lead one to suspect that an orthonormal basis might be constructed from the eigenvectors of a self-adjoint operator, and in fact, with a few reservations this turns out to be the case. One reservation is that the general operator will have both proper and improper eigenvectors, so our basis will have to contain both kinds. If that is OK, another reservation has to do with ***degeneracy***: in general, there will be many eigenvectors with the same eigenvalue. For example, the eigenvectors of the constant velocity-density wave-equation operator are the plane-wave vectors $|\mathbf{p},\omega\rangle$. ***All*** the solutions to the source-free wave equation have eigenvalue zero. To specify which plane wave we are dealing with, another operator $\hat{\mathbf{p}} \sim -i\nabla$ is required. The plane waves are ***also*** eigenvectors of $\hat{\mathbf{p}}$ with eigenvalue $\mathbf{p}$.

In general, to deal with degeneracy, a "complete set" of commuting self-adjoint operators is required. The eigenvectors of $\hat{L}$ with eigenvalue $\ell$ form what is called a ***degenerate subspace*** (degenerate with respect to $\hat{L}$, that is). We can always find a basis $|\ell, k\rangle$, $k=1,2,\ldots$ which "spans" the subspace, and another operator $\hat{K}$ may be defined such that $\hat{K}|\ell,k\rangle = k|\ell,k\rangle$. The operator $\hat{K}$, once defined on all of Hilbert space, must commute with $\hat{L}$

[4] Some non-self-adjoint operators may also have orthogonal eigenvectors. The necessary requirement is that the operator commute with its adjoint (that is, be a ***normal*** operator. Don't ask me where that name came from.)

since it shares eigenvectors and together the eigenvalues of $\hat{K}$ and $\hat{L}$ determine a unique (within a scale factor) vector.

Modulo these reservations, a complete set of orthonormal basis vectors can be constructed from the eigenvectors of $\hat{L}$. In fact, we may define, for each eigenvalue $\ell$ of $\hat{L}$ a projector $\hat{P}_\ell$ onto the subspace with that eigenvalue, and write a completeness relation

$$\hat{1} = \sum_\ell \hat{P}_\ell$$

bearing in mind that some of that sum is probably actually an integral over "improper" states. A common notation is to write

$$\hat{P}_\ell = \hat{\delta}(\ell - \hat{L})$$

exploiting the similarity between this projector and a Kroniker delta (for proper vectors) or a delta function (for improper vectors).

### *Spectral Decomposition and Functions of Operators*

Once a basis has been found for a self-adjoint operator $\hat{L}$, that operator can be expressed in terms of its basis vectors as

$$\hat{L} = \sum_\ell \ell \hat{P}_\ell$$

(Think about it. For any eigenvector[5] $|\ell,\cdot\rangle$ of $\hat{L}$, $\sum_{\ell'} \ell' \hat{P}_1' |\ell,\cdot\rangle = \ell|\ell,\cdot\rangle$, so for any vector which is a linear combination of eigenvectors of $\hat{L}$, the action of $\sum_\ell \ell \hat{P}_\ell$ must be the same as $\hat{L}$.) This equation is the ***spectral decomposition***, so called, of $\hat{L}$.

The spectral decomposition of $\hat{L}$ allows an easy definition of functions (e.g. square root, inverse, square, etc.) of an operator. We write, for an

[5]We use the dot to signify any other parameters besides $\ell$ needed to uniquely specify a vector.

arbitrary function f,

$$\hat{f}(\hat{L}) = \sum_{\ell} f(\ell)\hat{P}_{\ell}$$

*Green's Functions or Resolvents*

One function of special interest is the ***Green's function*** or ***resolvent*** of an operator. Define

$$\hat{G}(z - \hat{L}) = \sum_{\ell} \frac{1}{(z - \ell)} \hat{\delta}(\ell - \hat{L})$$

$\hat{G}(z - \hat{L})$ is seen to be an inverse of the operator $z - \hat{L}$:

$$\hat{G}(z - \hat{L}) = (z - \hat{L})^{-1}$$

It is well defined for any complex number z which is not an eigenvalue of $\hat{L}$. As z approaches an eigenvalue $\ell$ of $\hat{L}$, however, $1/(z-\ell)$ blows up. Unfortunately, $\hat{G}$ is most useful when z is an eigenvalue of $\hat{L}$. If $\hat{L}$ is the constant-frequency wave-equation operator $\hat{L} \sim k^{\frac{1}{2}} \nabla \cdot \frac{1}{\rho} \nabla k^{\frac{1}{2}}$, for example, then the equation

$$(\omega^2 - \hat{L})\, \hat{G}(\omega^2 - \hat{L}) = \hat{1}$$

is the equation for an impulsive source of unit magnitude, and we see $\hat{G}$ to be an impulse response.

Fortunately, the fact that $(\omega^2 - \hat{L})$ has zero eigenvalues does not mean there is no inverse; it just means the inverse is not unique.

Here are a few possibilities for G:

Principle value Green's function:

$$\hat{G}_p(\ell - \hat{L}) = \sum_{\ell' \neq \ell} \frac{1}{\ell - \ell'} \hat{\delta}(\ell' - \hat{L})$$

Exploding and imploding Green's functions:

$$\hat{G}_{\pm}(\ell - \hat{L}) = \lim_{\epsilon \to 0} \sum_{\ell'} \frac{1}{\ell \pm i\epsilon - \ell'} \hat{\delta}(\ell' - \hat{L})$$

one can note that

$$\hat{G}_p = \frac{1}{2}(\hat{G}_+ + \hat{G}_-), \quad \text{and} \quad \hat{G}_+^{\dagger} = \hat{G}_-$$

One may also define other Green's functions. Each corresponds to a different physical situation, and it is necessary to choose which one fits a particular circumstance.

***The Lippmann-Schwinger Equation and the Born Series***

Suppose we have two linear operators $\hat{L}_1$ and $\hat{L}_2$ differing by an operator $\hat{V}$:

$$\hat{L}_2 = \hat{L}_1 + \hat{V}$$

The resolvents (Green's functions) for $L_1$ and $L_2$ are then simply related by an equation called the Lippmann-Schwinger equation:

$$\hat{G}_{\pm}(\ell - \hat{L}_2) = \hat{G}_{\pm}(\ell - \hat{L}_1) + \hat{G}_{\pm}(\ell - \hat{L}_1)\, \hat{V} G_{\pm}(\ell - L_2)$$

or, equivalently,

$$\hat{G}_{\pm}(\ell - L_2) = \hat{G}_{\pm}(L - \hat{L}_1) + \hat{G}_{\pm}(\ell - \hat{L}_2)\, \hat{V} G_{\pm}(\ell - \hat{L}_1)$$

It should be noted that in general the two G's do not commute with each other

or with $\hat{V}$, so the identity of these two equations is not a trivial observation.

The Lippmann-Schwinger equations are widely used in scattering theory as a prescription for constructing an unknown Green's function from a known one. Formally, we may solve for $\hat{G}_{\pm}(\ell-\hat{L}_2)$ as

$$\hat{G}_{\pm}(\ell-\hat{L}_2) = [1 - \hat{G}_{\pm}(\ell-\hat{L}_1)\,\hat{V}]^{-1}\,\hat{G}_{\pm}(\ell-\hat{L}_1)$$

$$= \sum_{m=0}^{\infty} \left[\hat{G}_{\pm}(\ell-\hat{L}_1)\,\hat{V}\right]^m \hat{G}_{\pm}(\ell-\hat{L}_1)$$

Of course this bit of formalism works only if the inverse of $1 - \hat{G}_{\pm}(\ell-\hat{L}_1)\,\hat{V}$ exists and is series-expandable, which depends on the potential operator $\hat{V}$ being small in some sense. This series expansion for $\hat{G}_{\pm}(\ell-\hat{L}_2)$ in terms of $\hat{G}_{\pm}(\ell-\hat{L}_1)$ is called the ***Born series***. It is used by modern physicists with very little regard as to whether it converges or not.

The ***Born approximation*** amounts to truncating this series after the second term:

$$\hat{G}_{\pm}(\ell-\hat{L}_2) \simeq \hat{G}_{\pm}(\ell-\hat{L}_1) + G_{\pm}(\ell-\hat{L}_1)\,\hat{V}\hat{G}_{\pm}(\ell-\hat{L}_1)$$

Since higher-order terms in the Born expansion become very hard to evaluate, the Born approximation is widely used. We may also write down a Lippmann-Schwinger equation for the eigenvectors of $\hat{L}_1$, and $\hat{L}_2$. Suppose for simplicity that $\hat{L}_1$ and $\hat{L}_2$ have only improper eigenvectors (no "bound" states) and that their eigenvalues cover the same range $\ell \in (0,\infty)$. If $|\ell,\cdot\rangle$ is an eigenvector of $\hat{L}_1$ with eigenvalue 1, then two eigenvectors $|\ell_{\pm},\cdot\rangle$ of $\hat{L}_2$ with the same eigenvalue are found from the equation

$$|\ell_{\pm},\cdot\rangle = |\ell,\cdot\rangle + \hat{G}_{\pm}(\ell-\hat{L}_1)\,\hat{V}|\ell_{\pm},\cdot\rangle$$

It can be shown that this equation defines a unitary mapping (isometric if "bound" states exist), so the set of all $|\ell+,\cdot\rangle$ (or $|\ell-,.\rangle$) obey the same orthogonality and (if there are no bound states) completeness relations as do

the set $|\ell,\cdot\rangle$.

### *The $\hat{T}$ Operator*

Another operator often encountered in scattering theory is the $T$ operator. It is defined from $\hat{V}$ and $\hat{G}$ as

$$\hat{T}_{\pm}(\ell) = \hat{V} + \hat{V}\, G_{\pm}(\ell - \hat{L}_2)\, \hat{V}$$

Unlike $\hat{G}$, which depends only on one operator $\hat{L}_2$, $\hat{T}$ is a function of both $\hat{L}_1$ and $\hat{L}_2$.

A useful operator identity is found by premultiplying this equation by $\hat{G}_{\pm}(\ell - L_1)$:

$$\hat{G}_{\pm}(\ell - \hat{L}_1)\, \hat{T}_{\pm}(\ell) = \left[\hat{G}_{\pm}(\ell - \hat{L}_1) + VG_{\pm}(\ell - \hat{L}_2)\right] \hat{V}$$

$$= \hat{G}_{\pm}(\ell - \hat{L}_2)\, \hat{V}$$

Postmultiplying by the same operator yields another relation

$$\hat{T}_{\pm}(\ell)\, \hat{G}_{\pm}(\ell - \hat{L}) = \hat{V}\, \hat{G}_{\pm}(\ell - \hat{L}_2)$$

Substituting these relations into the defining equation for $\hat{T}$ gives us another pair of Lippmann-Schwinger equations:

$$\hat{T}_{\pm}(\ell) = \hat{V} + \hat{V}\, \hat{G}_{\pm}(\ell - \hat{L}_1)\, \hat{T}_{\pm}(\ell)$$

$$= \hat{V} + \hat{T}_{\pm}(\ell)\, \hat{G}_{\pm}(\ell - \hat{L}_1)\, \hat{V}$$

One is at liberty to expand these equations into a Born series for $\hat{T}$ if desired.

From the defining equation for $\hat{T}$ it is clear that it may be computed if $\hat{V}$ and $\hat{G}(\ell - \hat{L}_2)$ are known. The converse can also be seen to be true. Start

from the Lippmann-Schwinger equation for G:

$$\hat{G}(\ell-\hat{L}_2) = \hat{G}(\ell-\hat{L}_1) + \hat{G}(\ell-\hat{L}_1)\,\hat{V}G(\ell-\hat{L}_2)$$

and substitute the identity

$$\hat{V}G(\ell-\hat{L}_2) = \hat{T}(\ell)\,\hat{G}(\ell-\hat{L}_1)$$

to obtain

$$\hat{G}(\ell-\hat{L}_2) = \hat{G}(\ell-\hat{L}_1) + \hat{G}(\ell-\hat{L}_1)\,\hat{T}(\ell)\,\hat{G}\,(\ell-\hat{L}_1)$$

which gives $\hat{G}(\ell-\hat{L}_2)$ once $\hat{T}(\ell)$ and $\hat{G}(\ell-\hat{L}_1)$ are known. This equation figures in the solution of the seismic inverse problem.

$\hat{T}$ is also involved in the mapping of eigenvectors of $\hat{L}_1$ onto the eigenvectors of $\hat{L}_2$. Namely,

$$\hat{T}_{\pm}(\ell)\;|\ell,\cdot\rangle = \hat{V}|\ell_{\pm},\cdot\rangle$$

as is easily seen from the defining equations for $\hat{T}$ and $|\ell_{\pm},\cdot\rangle$. In terms of matrix elements

$$\langle\ell',\cdot|\hat{T}_{\pm}(\ell)|\ell,\cdot\rangle = \langle\ell',\cdot|\hat{V}|\ell_{\pm},\cdot\rangle$$

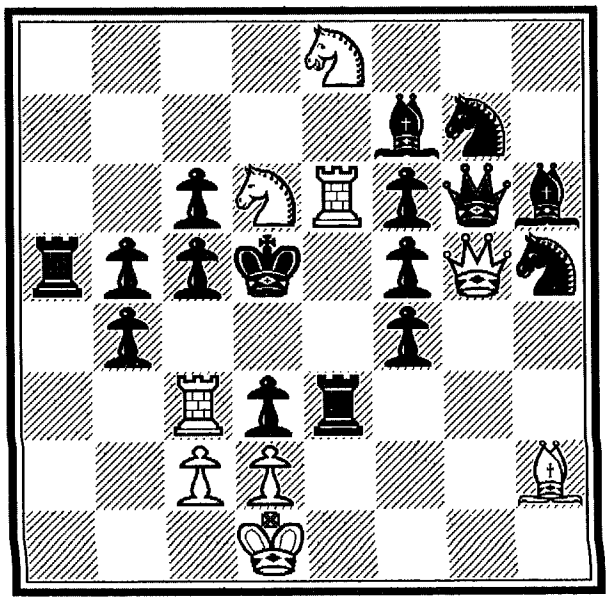
# REFERENCES

The following should be useful to someone trying to pick up some operator formalism as it relates to scattering theory. For a really thorough grounding in the subject one should descend into the mathematical literature, which I have scrupulously avoided.

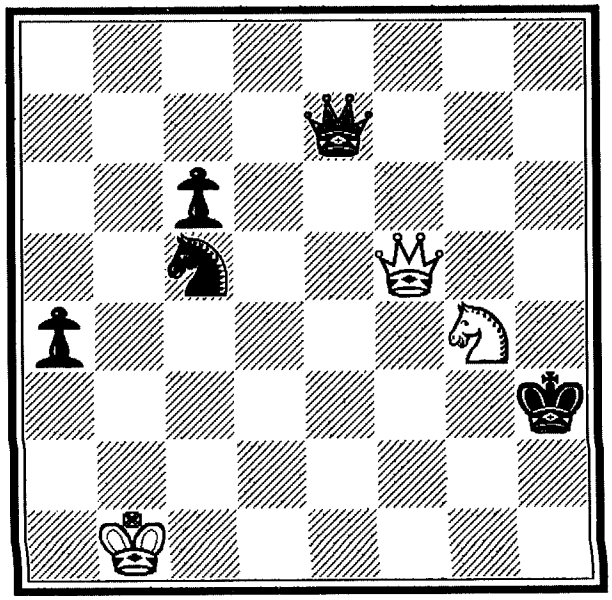
Jordon, T.F., 1969, Linear operators for quantum mechanics: New York, John Wiley & Sons, Inc. (*This book is so useful it is out of print.*)

Newton, R.G., 1966, Scattering theory of waves and particles: New York, McGraw-Hill Book Co. (*Rigorous, thorough, and comprehensive, but poorly organized and hard to read.*)

Taylor, John R., 1972, Scattering theory: New York, John Wiley & Sons, Inc. (*This provides an excellent review of the basic formalism and is probably the best introduction to scattering theory around. Consequently, it is hard to find.*)



White to play and mate in 10 moves -- or less!
(Alan Gottlieb)



White to play and win. This is a tough one.
(K.A.L. Kubbel)

Answers in next report.