Supplementary note on self-adjoint maps

Recall that for an inner product space $(V, < \cdot, \cdot >)$, a linear map

$$L : V \rightarrow V$$

is *self-adjoint* if

$$< Lv, w >=< v, Lw >$$

for all $v, w, \in V$.

Here is another proof of the

Theorem 0.1 *If L is self-adjoint, then all eigenvalues of L are real.*

Proof.

Let λ be an eigenvalue. Recall that there is at least one eigenvector v corresponding to λ because, by definition of eigenvalues,

$$\det(\lambda I - L) = 0$$

and hence, $\lambda I - L$, being non-invertible, must have a vector v in its kernel. Now note that

$$< Lv, v >=< \lambda v, v >= \lambda < v, v >$$

while

$$< v, Lv >=< v, \lambda v >= \bar{\lambda} < v, v >$$

Since L is self-adjoint, the two are equal

$$\lambda < v, v >= \bar{\lambda} < v, v >$$

and hence, $\lambda = \bar{\lambda}$ (since $< v, v >\neq 0$). □

Another nice fact is that

Proposition 0.2 *Let L be self-adjoint, v an eigenvector of eigenvalue λ and w and eigenvector of eigenvalue μ. If $\lambda \neq \mu$, the v and w are orthogonal.*

Proof.

$< Lv, w >=< \lambda v, w >= \lambda < v, w >$ and $< v, Lw >=< v, \mu v >= \mu < v, w >$, where we have used the fact that μ is real. Since the two quantities are equal and $\lambda \neq \mu$, we must have $< v, w >= 0$. □

A related notion is that of an *anti-self-adjoint* map. This is a linear map

$$L : V \rightarrow V$$

such that

$$< Lv, w >= - < v, Lw >$$

for all v, w. By reasoning entirely analogous to that above, it is easy to see that

all eigenvalues of an anti-self-adjoiint map are purely imaginary.

So if one thinks of linear maps as 'generalized numbers,' then self-adjoint maps are generalized real numbers, while anti-self-adoint maps are generalized purely-imaginary numbers. Thus, as expected, we can write any linear map L as

$$L = (L + L^*)/2 + (L - L^*)/2$$

a sum of a self-adjoint map and an anti-self-adjoint map, an extension of the fact that a complex number can be written as a sum of a real number and a purely imaginary number.

Also, L is self-adjoint if and only if iL is anti-self-adjoint, as can be seen (in one direction) from the fact that

$$< iLv, w >= i < Lv, w >= i < v, Lw >= -\bar{i} < v, Lw >= - < v, iLw >,$$

if L is self-adjoint. Therefore, the two kinds of linear maps tend to go together, and can be studied in conjunction. However, in quantum physics, it is the self-adjoint linear maps that literally represent physical quantities, since the eigenvalues become interpreted as the possible results of measurement, which must be real.

In the notes, it is shown that any self-ajoint map admits an orthonormal basis of eigenvectors. By the reasoning above, so does any anti-self-adjiont map. Thus, it is in the theory of (anti-)self-adjoint maps that linear maps and inner products become well-harmonized.

Here is an example of finding an orthonormal basis in the self-adjoint case.

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

An easy computaiton shows that the charactersitic polynomial is $ch(x) = (x-2)^2(x-8)$.

We find the eigenvectors corresponding to 2 by computing the kernel of

$$A - 2I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

which yields the equation $v_1 + v_2 + v_3 = 0$ for an eigenvector $v = (v_1, v_2, v_3)^t$. So we get that $\{(1, -1, 0)^t, (1, 0, -1)^t\}$ is a basis for the eigenspace corresponding to 2. For the eigenvalue 8, we need to find the kernel of

$$A - 8I = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$$

Applying an obvious sequence of row operations yields

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

giving us $(1, 1, 1)^t$ as a basis for this eigenspace. As expected, this basis element is already orthogonal to the vectors in the 2-eigenspace. But the natural basis we chose for the 2-eigenspace is not orthogonal. So we orthogonalize by putting $b'_1 = (1, -1, 0)^t$ and

$$b'_2 = (1, 0, -1)^t - \frac{< (1, 0, -1)^t, (1, -1, 0)^t >}{< (1, -1, 0)^t, (1, -1, 0)^t >}(1, -1, 0)^t$$

$$= (1, 0, -1)^t - (1/2)(1, -1, 0)^t = (1/2, 1/2, -1)$$

We see that b'_2 is now indeed orthogonal to b'_1. Furthermore, we should check our computation by summing up the coordinates $1/2 + 1/2 - 1 = 0$ to see that we are still in the right eigenspace. Finally, to *orthonormalize*, we divide each vector by its length to arrive at

$$\{(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)^t, (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{\sqrt{2}}{\sqrt{3}})^t, (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^t\}$$

as an orthonormal basis of eigenvectors.

This example is unfortunately formal, because I couldn't think of an elementary example that leads to a really striking orthonormal basis with interesting interpretations. Perhaps I'll come up with one later and post an addendum. At the moment, the really good ones I can think of start with an infinite-dimensional space. If you do want to see nice actual formulas for such things, do a Google search with some term like 'spherical harmonics.'

I can't resist mentioning one last fact, also important in quantum mechanics. Some of you have encountered *matrix exponentials*

$$\exp(A) := I + A + A^2/2! + A^3/3! + \cdots$$

that always converges to a well-defined matrix for any given A. When A is diagonalizable, so that $A = PDP^{-1}$ for a diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, then we can alternatively define

$$\exp(A) = P \exp(D) P^{-1}$$

where

$$\exp(D) = \mathrm{diag}(\exp(\lambda_1), \exp(\lambda_2), \ldots, \exp(\lambda_n))$$

(How does this relate to the other definition?) Then a nice fact is that

Proposition 0.3 *If A is anti-self-adjoint, then* $\exp(A)$ *is an isometry, i.e.,*

$$< \exp(A)v, \exp(A)w >=< v, w >$$

for all v, w.

It is a rather pleasant exercise to prove this for yourself.

An equivalent statement is that if H is self-adjoint, then

$$\exp(iH)$$

is an isometry.

In applications, one puts in a variable t, standing for time, and considers

$$U(t) = \exp(itH)$$

a one-parameter family of matrices. You should convince yourself that it makes sense to *differentiate* this family (for example, by using diagonalization), and get

$$(d/dt)U(t) = iHU(t)$$

so that

$$[(d/dt)U(t)|t = 0] = iH$$

The resulting fact is that $U(t)$ is a *family of isometries* when $-i(d/dt)U(t)|t = 0$ is self-adjoint.

For any vector v in the vector space, if we put

$$v(t) := U(t)v$$

(thought of as v evolving in time), then we deduce the differential equation

$$-i(d/dt)v(t) = Hv(t)$$

known as the *Schrödinger equation* for the time-evolution of v, in the situation where H represents the *energy* of a physical system. This is all very abstract, and the significant work is to know in real-world examples what H and its eigenvectors look like, and to interprete the calculations.