Notes for the undergraduate lecture by David Adams. (These are the notes I would write if I was teaching a course on this topic. I have included more material than I will cover in the 45 minute lecture; the intention is to show how I would cover this topic in its entirety. In practice I would give out an exercise sheet in connection with this material, but I will skip that here.)

Spectral Theorem for Self-adjoint Linear Operators

(finite-dimensional case)

Let V be a finite-dimensional vector space, either real or complex, and equipped with an inner product $\langle\cdot\,,\cdot\rangle$. Let $A: V \to V$ be a linear operator. Recall that the *adjoint* of A is the linear operator $A^*: V \to V$ characterized by

$$\langle A^* v, w\rangle = \langle v, Aw\rangle \qquad \forall v, w \in V \tag{0.1}$$

A is called *self-adjoint* (or *Hermitian*) when $A^* = A$.

Spectral Theorem. If A is self-adjoint then there is an orthonormal basis (o.n.b.) of V consisting of eigenvectors of A. Each eigenvalue is real.

Before proceeding to the proof, let us note why this theorem is important. Recall that, after making a choice of of basis $v_1, \dots, v_n$ for V, a linear operator A corresponds to an $n \times n$ matrix $\mathbf{A} = (A_{ij})$ where $n = \dim V$. The correspondence is given by

$$Av_j = \sum_{i=1}^{n} A_{ij} v_i \qquad\qquad \forall j = 1, \dots, n \tag{0.2}$$

The spectral theorem tells us that if A is self-adjoint then V admits an eigenvector basis, $Av_j = \lambda_j v_j$, in which case the matrix for A is *diagonal*:

$$\mathbf{A} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{0.3}$$

Thus A admits a very simple representation in this case. As an example of the useful consequences of this, we can evaluate A^p from the fact that

$$\mathbf{A}^p = \begin{pmatrix} \lambda_1^p & & 0 \\ & \ddots & \\ 0 & & \lambda_n^p \end{pmatrix} \tag{0.4}$$

The practical significance of of the eigenvector basis being *orthonormal* is that it is easier to determine in explicit calculations. In fact, a basis of eigenvectors usually exists for non-self-adjoint linear operators as well.[1] However, it is usually much more difficult to explicitly determine the basis in that case.

---

[1] This can be shown when the eigenvalues of the operator are all distinct. However, when the eigenvalues are not all distinct there are cases where a basis of eigenvectors does not exist.

A specific example: For $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, regarded as a linear operator on $\mathbf{R}^2$, one can derive the formula

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^p = \begin{pmatrix} a(p) & a(p+1) \\ a(p+1) & a(p+2) \end{pmatrix} \qquad a(p) = \lambda_+^p + \lambda_-^p \,, \qquad \lambda_\pm = \frac{1}{2}(1 \pm \sqrt{5}) \tag{0.5}$$

where $\lambda_\pm$ are the eigenvalues of $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. The derivation of this is given as an exercise (Exercise 1 in the exercise section; hints are provided).

Proof of the Spectral Theorem. We use induction on the dimension of V. First we show that the theorem holds when $\dim V = 1$. Then we assume that it holds when the dimension is $n-1$ and show that it holds when the dimension is n.

When $\dim V = 1$ all linear operators are multiplication by scalars: $Av = \lambda v$ $\forall v \in V$. The scalar λ is seen to be real as follows: For arbitrary $v \in V$ we have

$$\lambda |v|^2 = \langle Av, v \rangle = \langle v, Av \rangle = \overline{\langle Av, v \rangle} = \overline{\lambda |v|^2} = \overline{\lambda} |v|^2 \tag{0.6}$$

hence $\lambda = \overline{\lambda}$. Thus the theorem is verified in this case.

Assuming now that the theorem holds when the vector space has dimension $n-1$, we show that it holds when the dimension is n. There are two main steps:

Step 1. Show that A has at least one eigenvector, v_1, with real eigenvalue λ_1. This is the "hard part". We do it further below.

Step 2. Define W to be the vector subspace orthogonal to v_1 in V:

$$W := \{\, v \in V \mid \langle v, v_1 \rangle = 0 \,\} \tag{0.7}$$

Show that $A(W) \subseteq W$, i.e., W is *invariant* under A.

Step 2 is easily done by by noting that

$$\langle v, v_1 \rangle = 0 \;\Rightarrow\; \langle Av, v_1 \rangle = \langle v, Av_1 \rangle = \langle v, \lambda_1 v_1 \rangle = \lambda_1 \langle v, v_1 \rangle = 0 \tag{0.8}$$

where we used $Av_1 = \lambda_1 v_1$ from Step 1. This shows that if $v \in W$ then $Av \in W$, so $A(W) \subseteq W$ as claimed.

The result of Step 2 implies that A restricts to a self-adjoint linear operator on W:

$$A : W \to W \,. \tag{0.9}$$

Since $\dim W = \dim V - 1 = n - 1$ the induction hypothesis implies that W has an o.n.b. $\{v_2, \ldots, v_n\}$ of eigenvectors for A with real eigenvalues $\lambda_2, \ldots, \lambda_n$. Since these eigenvectors are all contained in W they are all orthogonal to v_1, so $\{v_1, v_2, \ldots v_n\}$ is an o.n.b. for V. (We normalized v_1 so that $|v_1| = 1$.) Thus we obtain an o.n.b. for V of eigenvectors for A with real eigenvaues, and the Spectral Theorem is established.

To summarize, we have shown that the Spectral Theorem holds if we can do Step 1. There are several ways to do this step. The quickest way is to use the "fundamental theorem of algebra" (f.t.a.) as follows. The f.t.a. asserts that any non-constant complex polynomial $p(z)$ vanishes at at least one value of the complex variable z. We apply this to the determinant

$$p(z) = \det(A - z\mathbf{1}) \tag{0.10}$$

where $\mathbf{1}$ is the identity map on V. The f.t.a. asserts that there exists a z_0 such that $\det(A - z_0\mathbf{1}) = 0$. But this implies that there is a non-zero solution v to

$$(A - z_0\mathbf{1})v = 0 \tag{0.11}$$

The solution is an eigenvector: $Av = z_0 v$. The eigenvalue z_0 is seen to be real by noting that $z_0|v|^2 = \langle Av, v\rangle$ and performing the same calculation as in (0.6) to find $z_0 = \overline{z_0}$. Setting $v_1 = v$ and $\lambda = z_0$ completes Step 1, and with this we have completed the proof of the Spectral Theorem.

An alternative way to do Step 1.:

We now present another way to do Step 1. It is longer than the preceding one, but more elementary: It does not require the fundamental theorem of algebra but only relies on a basic property of continuous functions.

Define the function $f : V \to \mathbf{R}$ by

$$f(v) = \langle Av, v\rangle \tag{0.12}$$

Note that f is real-valued since $\overline{\langle Av, v\rangle} = \langle v, Av\rangle = \langle Av, v\rangle$. Let S denote the *unit sphere* in V:

$$S := \{\, v \in V \mid |v| = 1 \,\} \tag{0.13}$$

Note that f is a continuous function and S is *compact* topological subspace of V. It follows from a basic theorem on continuous functions that f has a finite maximum on S. I.e., there is a $v_1 \in S$ such that

$$f(v) \;\le\; f(v_1) \qquad\qquad \forall v \in S \tag{0.14}$$

We now show that v_1 is an eigenvector for A and its eigenvalue is $\lambda = f(v_1)$. For arbitrary $v \in V$ and $t \in \mathbf{R}$ the inequality (0.14) implies

$$f\Big(\frac{v_1 + tv}{|v_1 + tv|}\Big) \;\le\; f(v_1)\,, \tag{0.15}$$

which can be re-expressed as

$$\langle A(v_1 + tv), v_1 + tv\rangle - f(v_1)\langle v_1 + tv, v_1 + tv\rangle \;\le\; 0\,. \tag{0.16}$$

Defining $g_v(t)$ to be the function on the left-hand side, we have

$$g_v(t) \ \leq \ 0 \qquad \text{and} \qquad g_v(0) = 0 \qquad \forall v \in V,\ t \in \mathbf{R} \tag{0.17}$$

It follows that $g_v'(0) = 0$.[2] This gives

$$0 = g_v'(0) = \langle Av, v_1 \rangle + \langle Av_1, v \rangle - f(v_1)\langle v, v_1 \rangle - f(v_1)\langle v_1, v \rangle \,. \tag{0.18}$$

When v is orthogonal to v_1 (i.e., $\langle v, v_1 \rangle = 0$) this becomes

$$\begin{aligned} 0 &= \langle Av, v_1 \rangle + \langle Av_1, v \rangle = \overline{\langle Av_1, v \rangle} + \langle Av_1, v \rangle \\ &= 2\mathrm{Re}\langle Av_1, v \rangle \end{aligned} \tag{0.19}$$

where we have used $\langle Av, v_1 \rangle = \langle v, Av_1 \rangle = \overline{\langle Av_1, v \rangle}$. If the vector space V is complex, then by replacing v by iv in this argument we find from (0.19) that the imaginary part of $\langle Av_1, v \rangle$ also vanishes. Thus we have shown that

$$\langle Av_1, v \rangle = 0 \quad \forall v \in V \ \text{ with } \ \langle v, v_1 \rangle = 0 \,. \tag{0.20}$$

This implies that Av_1 has no component orthogonal to v_1 and therefore must be proportional to v_1, i.e.,

$$Av_1 = \lambda v_1 \tag{0.21}$$

for some scalar λ. Thus v_1 is an eigenvalue for A as claimed. To see that its eigenvalue λ is real we note that, since $|v_1| = 1$,

$$\lambda = \langle Av_1, v_1 \rangle = f(v_1) \tag{0.22}$$

which, as noted earlier, is real. This completes the alternative way to do Step 1.

Spectral Theorem for Hermitian matrices

We now specialize to the case where the linear operator is a matrix $\mathbf{A} = (A_{ij})$ on the complex vector space $\mathbf{C}^n$ (or real vector space $\mathbf{R}^n$) equipped with the standard inner product.

Recall that the *adjoint* of the matrix $\mathbf{A}$ (also called the *conjugate transpose* or *Hermitian transpose*) is $\mathbf{A}^* = (A^*_{ij})$ where

$$A^*_{ij} = \overline{A_{ji}} \tag{0.23}$$

The matrix is called *Hermitian* when it is its own adjoint, i.e., when

$$A_{ji} = \overline{A_{ij}} \qquad \forall i, j = 1, \ldots, n. \tag{0.24}$$

---

[2] This is easy to see from (0.17) after recalling that $g'(0)$ gives the tangent to the graph of $g(t)$ at $t = 0$.

In the case of a real matrix this becomes $A_{ji} = A_{ij}$ in which case the matrix is called *symmetric.*

The spectral theorem for self-adjoint linear operators corresponds to the following theorem for complex Hermitian (or real symmetric) matrices. It is very useful in both pure and applied mathematics, having important and sometimes surprising applications to a wide range of problems. (Several of these will be given in the Exercises.)

Spectral theorem for complex Hermitian (or real symmetric) matrices: Every complex Hermitian (or real symmetric) matrix $\mathbf{A}$ can be diagonalized by a unitary transformation. Specifically,

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D} \tag{0.25}$$

where $\mathbf{D}$ is a diagonal matrix with real entries and $\mathbf{U}$ is a *unitary* matrix, i.e., $\mathbf{U}^* = \mathbf{U}^{-1}$. The matrices $\mathbf{U}$ and $\mathbf{D}$ are obtained as follows. Let $v_1, \ldots, v_n$ be an o.n.b. of eigenvectors for $\mathbf{A}$ (whose existence is guaranteed by the previous Spectral Theorem), and $\lambda_1, \ldots, \lambda_n$ the corresponding eigenvalues. Then (0.25) holds with $\mathbf{U}$ being the matrix whose j'th column is v_j , and $\mathbf{D}$ the diagonal matrix whose j'th diagonal entry is λ_j for all $j = 1, \ldots, n$.

Proof. From the definition of $\mathbf{U}$ it is straightforward to show that

$$v_j = \mathbf{U}e_j \qquad \text{and} \qquad e_j = \mathbf{U}^* v_j \,. \tag{0.26}$$

From this it readily follows that

$$\mathbf{U}^*\mathbf{U} = \mathbf{1}, \tag{0.27}$$

showing that $\mathbf{U}$ is unitary as claimed. (Exercise: verify (0.26) and (0.27).) Now, using the notation $\mathbf{U} = (v_1, \ldots, v_n)$ to represent that the j'th column of $\mathbf{U}$ is v_j, we have

$$\mathbf{A}\mathbf{U} = (\mathbf{A}v_1, \ldots, \mathbf{A}v_n) = (\lambda_1 v_1, \ldots, \lambda_n v_n) \tag{0.28}$$

and hence

$$\begin{aligned} \mathbf{U}^*\mathbf{A}\mathbf{U} &= (\lambda_1 \mathbf{U}^* v_1, \ldots, \lambda_n \mathbf{U}^* v_n) \\ &= (\lambda_1 e_1, \ldots, \lambda_n e_n) \\ &= \mathbf{D} \end{aligned} \tag{0.29}$$

where (0.26) was used to obtain the second equality. This completes the proof.

Note that the matrix spectral theorem can be re-expressed as

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^* \,. \tag{0.30}$$

Since $\mathbf{U}$ is unitary it follows that

$$\mathbf{A}^p = \mathbf{U}\mathbf{D}^p\mathbf{U}^* \tag{0.31}$$

for any power p. (Exercise: show this.)