3. DIAGONALISATION REVISITED

§3.1. Bases of Eigenvectors

Recall that an eigenvector \( v \) for a matrix \( A \) is a non-zero vector where \( Av = \lambda v \). The scalar \( \lambda \) is called the corresponding eigenvalue. We can extend these concepts to linear transformations.

If \( f : U \to V \) is a linear transformation between two vector spaces, over the field \( F \), and \( v \) is a non-zero element of \( U \) such that \( f(v) = \lambda v \) for some \( \lambda \in F \) then \( v \) is called an eigenvector for \( f \) and \( \lambda \) is the corresponding eigenvalue.

This agrees with our previous definition by considering the linear transformation \( f(v) = Av \). But it makes sense when there is no obvious matrix in the picture.

**Example 1:** Let \( U \) be the set of differentiable functions of one real variable \( x \) and let \( V \) be the space of all real functions of one real variable \( x \), both considered as vector spaces over \( \mathbb{R} \).

Let \( D : U \to V \) be the linear transformation \( D(f) = \frac{df}{dx} \).

An eigenvector is a function \( f \) such that \( \frac{df}{dx} = \lambda f \) for some \( \lambda \in \mathbb{R} \). Clearly the eigenvectors are the exponential functions, or more properly, the non-zero multiples of the exponential functions. For any real \( \lambda \) the functions \( f(x) = Ce^{\lambda x} \), for any \( C \neq 0 \), are eigenvectors for \( D \), with the real number \( \lambda \) being the corresponding eigenvalue.

**Example 2:** Let \( V \) be 3-dimensional Euclidean space, over \( \mathbb{R} \). Let \( \pi \) be any plane in \( V \) that passes through the origin. Then reflection in \( \pi \) is a linear transformation, \( M \). The eigenvalues are \( \pm 1 \).

For \( \lambda = 1 \) the eigenvectors are the non-zero vectors in \( \pi \) because these are fixed by the reflection. For \( \lambda = -1 \) the eigenvectors are the non-zero vectors that are perpendicular to the plane.

You shouldn’t have your thinking about eigenvalues dominated by the equation \( |\lambda I - A| = 0 \). True, if you are given a random \( n \times n \) matrix this is usually the best way to find the eigenvalues. And once you’ve found the eigenvalues you can then find the eigenvectors. But there are some matrices, as we shall see, when it is easier to find the eigenvectors first, and then the eigenvalues.

But when you have to find the eigenvalues and eigenvectors of a linear transformation that is not given by a matrix then \( |\lambda I - A| \) will not make sense. How could you use determinants in example 1 or example 2? Of course we can always put in a basis, and represent the linear transformation by a matrix, but this is often a long-winded way of going about it. Always think first of the equation \( f(v) = \lambda v \).

If you have a linear transformation \( f : V \to V \) on a finite-dimensional vector space \( V \), the nicest basis, if such a basis exists, would be a basis of eigenvectors. For, if \( \alpha = \{v_1, \ldots, v_n\} \) is a basis of eigenvectors for the \( n \times n \) matrix \( A \), with corresponding eigenvalues \( \lambda_1, \ldots, \lambda_n \) then the matrix of \( f \) relative to this basis of eigenvectors is simply the diagonal matrix \( D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \).

But if the linear transformation is \( f(v) = Av \) it will be expressed in terms of the standard basis \( \beta \). To reach the diagonal matrix we simply need a change of basis. If \( S = \left[ \begin{array}{c} \alpha \\ \beta \end{array} \right] \) then
Hence E

It follows that x

Hence x

Then A

the factors A

For each r let A

Suppose

For each r let v

Proof: Suppose

Theorem

1.

E

§3

eigenvectors only span a 1-dimensional subspace. The space of eigenvectors for λ = 5 is just 1, even though 5 is a double zero of the characteristic polynomial. The space of eigenvectors for λ = 10 also has dimension 1. Hence there is no basis of eigenvectors for \( \mathbb{R}^3 \). The eigenvectors only span a 2-dimensional subspace.

Example 3: Let 

\[
A = \begin{pmatrix} 7 & 2 & -1 \\ -3 & 2 & 4 \\ -2 & 2 & 11 \end{pmatrix}
\]

The characteristic polynomial is \((\lambda - 10)(\lambda - 5)^2\).

But

\[
A - 5I = \begin{pmatrix} 2 & 2 & -1 \\ -3 & -3 & 4 \\ -2 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -3 \\ -3 & -3 & 4 \\ 1 & 1 & -3 \end{pmatrix} \to \begin{pmatrix} 1 \ 1 \ -3 \\ 0 \ 0 \ 5 \\ 0 \ 0 \ 0 \end{pmatrix} \text{ so the}
\]

eigenvectors for λ = 5 are all scalar multiples of

\[
\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}
\]

The space of eigenvectors for the
eigenvalue 5 is just 1, even though 5 is a double zero of the characteristic polynomial. The space of eigenvectors for λ = 10 also has dimension 1. Hence there is no basis of eigenvectors for \( \mathbb{R}^3 \). The eigenvectors only span a 2-dimensional subspace.

§3.2. Eigenspaces

If λ is an eigenvalue for the matrix A then the corresponding eigenspace is:

\( E_A(\lambda) = \{ v \mid Av = \lambda v \} \). By the definition of eigenvalue the dimension of each eigenspace is at least 1. The total eigenspace is the space spanned by all the eigenvectors and is denoted by \( E_A \).

Theorem 2: The total eigenspace is the direct sum of the individual eigenspaces.

Proof: Suppose the eigenvalues of the n×n matrix A are \( \lambda_1, \ldots, \lambda_k \).

For each r let \( v_r \in E_{A}(\lambda_r) \). Then \( Av_r = \lambda_r v_r \).

Suppose \( v = x_1v_1 + \ldots + x_kv_k = 0 \).

For each r let \( A_r = (A - \lambda_1 I) \ldots (A - \lambda_{r-1} I)(A - \lambda_{r+1} I) \ldots (A - \lambda_k I) \). That is, \( A_r \) is the product of the factors \( A - \lambda_s I \) for all s except for s = r.

Then \( A_r v = x_1A_r v_1 + \ldots + x_{r-1}A_r v_{r-1} + x_r A_r v_r + \ldots + x_k A_r v_k = 0 \).

Hence \( x_r A_r v_r = 0 \) because all the other terms are zero.

But \( A_r v_r = (\lambda_1 - \lambda_r) \ldots (\lambda_1 - \lambda_r) (\lambda_1 - \lambda_r) \ldots (\lambda_1 - \lambda_r) v_r \neq 0 \).

It follows that \( x_r = 0 \).

Hence \( E_A = E_A(\lambda_1) \oplus \ldots \oplus E_A(\lambda_k) \).
Example 4: Let \( A = \begin{pmatrix} 1 & 8 & -4 \\ 8 & 1 & 4 \\ -4 & 4 & 7 \end{pmatrix} \).

\[ \chi_A(\lambda) = |\lambda I - A| = (\lambda - 9)^2(\lambda + 9) \] so the eigenvalues are \( \pm 9 \), with \( 9 \) being a repeated eigenvalue.

\( E_A(9) \) is the null space of \( A - 9I = \begin{pmatrix} -8 & 8 & -4 \\ 8 & -8 & 4 \\ -4 & 4 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \). Hence \( E_A(9) = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} \rangle \).

\( E_A(-9) \) is the null space of \( A + 9I = \begin{pmatrix} 10 & 8 & -4 \\ 8 & 10 & 4 \\ -4 & 4 & 16 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & -4 \\ 0 & 9 & 18 \\ 0 & 9 & 18 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \).

Hence \( E_A(-9) = \langle \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \rangle \).

So \( E_A = E_A(9) \oplus E_A(-9) \) and so has dimension 3.

It follows that we have a basis of eigenvectors for \( \mathbb{R}^3 \) and so \( A \) is diagonalisable.

§3.3. Diagonalisable Matrices

A square matrix is diagonalisable if and only if it has a basis of eigenvectors, or equivalently if every vector is a sum of eigenvectors. In many cases it is possible to conclude that the matrix is diagonalisable because of other properties.

**Theorem 3:** An \( n \times n \) matrix is diagonalisable if it has \( n \) distinct eigenvalues.

**Proof:** If the eigenvalues are \( \lambda_1, \ldots, \lambda_n \) then \( E_A = E_A(\lambda_1) \oplus \cdots \oplus E_A(\lambda_n) \) and so \( \dim(E_A) = n \).

A projection matrix is one that satisfies the equation \( A^2 = A \). They are the matrices of a projection onto a subspace, such as the projection of points in \( \mathbb{R}^3 \) onto a plane. The points on the plane are fixed by the projection and so projecting twice is the same as projecting just once.

**Theorem 4:** Projection matrices are diagonalisable.

**Proof:** Let \( A \) be a projection matrix. Any vector \( v \) can be expressed as

\[ v = (v - Av) + Av \]

Clearly \( v - Av \in E_A(0) \) and \( Av \in E_A(1) \), so the eigenvectors span the whole space.

A matrix \( A \) has finite order \( n \) if \( A^n = I \).

**Theorem 5:** Matrices over \( \mathbb{C} \) of finite order are diagonalisable.

**Proof:** Let \( A \) be an \( n \times n \) matrix over \( \mathbb{C} \) such that \( A^m = I \).

Let \( \theta = e^{2\pi i/m} \) and let \( v \) be any vector in \( \mathbb{C}^n \). Then \( 1 + \theta + \theta^2 + \cdots + \theta^{(m-1)} \) is the sum of a geometric progression and so, if \( \theta^t \neq 1 \), is

\[ \frac{\theta^m - 1}{\theta - 1} = \frac{(\theta^m)^r - 1}{\theta^r - 1} = 0. \]

Then \( v = \frac{1}{m} \left( v + \theta Av + \theta^2 A^2v + \cdots + \theta^{m-1} A^{m-1}v \right) + \frac{1}{m} \left( v + \theta^2 Av + \theta^4 A^2v + \cdots + \theta^{2(m-1)} A^{m-1}v \right) + \theta \cdots + \frac{1}{m} \left( v + \theta^{m-1} Av + \theta^{2(m-1)} A^2v + \cdots + \theta^{(m-1)^2} A^{m-1}v \right) + \frac{1}{m} \left( v + Av + A^2v + \cdots + A^{m-1}v \right). \)

This is because the sum of each column, except the first, is zero. The sum of the \( r \)th column is
\[(1 + \theta^r + \theta^{2r} + \ldots + \theta^{(m-1)r})A^r v = 0 \text{ when } r > 1.\]

Now \(A(v + \theta^r A v + \theta^{2r} A^2 v + \ldots + \theta^{(m-1)r} A^{m-1} v) = A v + \theta^r A^2 v + \ldots + \theta^{(m-1)r} A^{m-1} v = A v + \theta^r A^2 v + \ldots + \theta^{(m-1)r} A^{m-1} v + \theta^r v\) 
\[= \theta^r (v + \theta^r A v + \theta^{2r} A^2 v + \ldots + \theta^{(m-1)r} A^{m-1} v) = 0 \text{ for all } r, A v \in \text{span the whole space of eigenvectors for } A.\]

Hence each row in the above is an eigenvector, and so every vector is a sum of eigenvectors.

**Example 5**: If \(A^2 = I\) then \(A\) is diagonalisable. Every vector can be expressed as \(v = \frac{1}{2} (v - Av) + \frac{1}{2} (v + Av) \in E_A(-1) + E_A(1).

The following is an alternative proof of the above theorem. It is interesting that this proof uses some elementary calculus, including the Fundamental Theorem of Calculus.

The zeros of \(x^n - 1\) over \(\mathbb{C}\), are \(1, 0, \theta^2, \ldots, \theta^{n-1}\) where \(\theta = e^{2\pi i n}\). Hence \(x^n - 1\) factorises over \(\mathbb{C}\) as \(x^n - 1 = (x - 1)(x - \theta)(x - \theta^2) \ldots (x - \theta^{n-1}).\)

**Lemma**: Let \(\theta = e^{2\pi i k}\). Then \(\frac{1}{x - 1} + \frac{1}{x - \theta} + \frac{1}{x - \theta^2} + \ldots + \frac{1}{x - \theta^{n-1}} = \frac{nx^{n-1}}{x^n - 1}.\)

**Proof**: The integral of the LHS is \(\log(x - 1) + \log(x - \theta) + \log(x - \theta^2) + \ldots + \log(x - \theta^{n-1}) = \log((x - 1)(x - \theta)(x - \theta^2) \ldots (x - \theta^{n-1})) = \log(x^n - 1),\) since the zeros of \(x^n - 1\) are \(1, 0, \theta^2, \ldots, \theta^{n-1}.\)

Differentiating we get \(\frac{nx^{n-1}}{x^n - 1}\). By the Fundamental Theorem of Calculus the derivative of the integral of the LHS is the LHS itself, and so the result follows.

For \(r = 0, 1, 2, \ldots, n - 1\) define \(e r(x) = \frac{x^n - 1}{x - \theta^r}\). This has degree \(n - 1\) and is the product of the terms \(x - \theta^t\) for \(t \neq r,\)

**Lemma**: \(e_0(x) + e_1(x) + e_2(x) + \ldots + e_n(x) = nx^n - 1.\)

**Proof**: Putting \(\frac{1}{x - 1} + \frac{1}{x - \theta} + \frac{1}{x - \theta^2} + \ldots + \frac{1}{x - \theta^{n-1}}\) over a common denominator we obtain \(e_0(x) + e_1(x) + e_2(x) + \ldots + e_n(x)\)

\(\frac{nx^n - 1}{x^n - 1}.\)

Now the variable \(x\) in the proof of the lemma was a real variable. But the conclusion to lemma 2 will be valid as an equality of polynomials over \(\mathbb{C}\).

**Second Proof**: The eigenvalues of \(A\) are \(n\)-th roots of unity, \(1, 0, \theta^2, \ldots, \theta^{n-1}\) where \(\theta = e^{2\pi i n}.\)

From the second lemma we get \(e_0(A) + e_1(A) + e_2(A) + \ldots + e_{n-1}(A) = nA^{n-1}.\)

Clearly \(A\) is invertible and so \(nI = A^{1-n} [e_0(A) + e_1(A) + e_2(A) + \ldots + e_{n-1}(A)]\).

Let \(v\) be any column vector with as many components as \(A\) has rows.

Then \((A - \theta I)(A^{1-n} e_t(A) v) = A^{1-t}(A - I)(A - \theta I)(A - \theta^2 I) \ldots (A - \theta^{n-1} I) v = A^{1-t} (A^n - 1)v = 0.\)

It follows that, for all \(r, A^{1-n} e_r(A) v\) is an eigenvector for the eigenvalue \(\theta^r\), or it is zero.

Now \(n v = A^{1-n} e_0(A) v + A^{1-n} e_1(A) v + A^{1-n} e_2(A) v + \ldots + A^{1-n} e_{n-1}(A) v\) and so \(v\) is a linear combination of eigenvectors for \(A\). The eigenvectors thus span the whole space and so \(A\) is diagonalizable.

30
Example 6: Suppose $A^3 = I$. The eigenvalues can be found among the cube roots of unity, 1, $\omega$ and $\omega^2$.

$e_0(x) = (x - \omega)(x - \omega^2)$, $e_1(x) = (x - 1)(x - \omega^2)$ and $e_2(x) = (x - 1)(x - \omega)$.

So $3A^2 = (A - \omega I)(A - \omega^2 I) + (A - I)(A - \omega^2 I) + (A - I)(A - \omega I)$ which means that every vector $v$ can be expressed as

$$v = \frac{1}{3} A(A - \omega I)(A - \omega^2 I)v + \frac{1}{3} A(A - I)(A - \omega^2 I)v + \frac{1}{3} (A - I)(A - \omega I)v.$$

A cyclic matrix is one of the form

$$\begin{pmatrix}
a_1 & a_2 & \ldots & a_n \\
a_n & a_1 & \ldots & a_{n-1} \\
\ldots & \ldots & \ldots & \ldots \\
a_2 & a_3 & \ldots & a_1
\end{pmatrix}$$

Example 7: 

$$\begin{pmatrix}
1 & 2 & 3 & 4 \\
4 & 1 & 2 & 3 \\
3 & 4 & 1 & 2 \\
2 & 3 & 4 & 1
\end{pmatrix}$$

is a cyclic matrix.

Theorem 6: Cyclic matrices are diagonalisable.

Proof: Let $A = \begin{pmatrix}
a_1 & a_2 & \ldots & a_n \\
a_n & a_1 & \ldots & a_{n-1} \\
\ldots & \ldots & \ldots & \ldots \\
a_2 & a_3 & \ldots & a_1
\end{pmatrix}$ and let $\theta = e^{2\pi i/n}$.

Then for all $r$, 

$$A \begin{pmatrix}
1 \\
\theta^r \\
\theta^{2r} \\
\ldots
\end{pmatrix} = (a_1 + a_2\theta^r + \ldots + a_n\theta^{(n-1)} \begin{pmatrix}
1 \\
\theta^r \\
\theta^{2r} \\
\ldots
\end{pmatrix}$$

It follows that each 

$$\begin{pmatrix}
1 \\
\theta^r \\
\theta^{2r} \\
\ldots
\end{pmatrix}$$

is an eigenvector for $A$.

The determinant whose $r$'th column is 

$$\begin{pmatrix}
1 \\
\theta^r \\
\theta^{2r} \\
\ldots
\end{pmatrix}$$

is the Vandermonde determinant $V(1, \theta, \theta^2, \ldots, \theta^{n-1})$. Since 1, $\theta$, $\theta^2$, $\ldots$, $\theta^{n-1}$ are distinct this Vandermonde determinant is non-zero and so its columns are linearly independent. Hence there is a basis of eigenvectors.
§8.4. Normal Matrices

In terms of this adjoint we define several properties that complex matrices may have.

<table>
<thead>
<tr>
<th>NAME</th>
<th>DEFINITION</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal matrix</td>
<td>$AA^* = A^*A$</td>
<td>$\begin{pmatrix} 2 &amp; 1 \ i &amp; 1 + i \end{pmatrix}$</td>
</tr>
<tr>
<td>Hermitian matrix</td>
<td>$A^* = A$</td>
<td>$\begin{pmatrix} 3 &amp; 1 - 2i \ 1 + 2i &amp; -1 \end{pmatrix}$</td>
</tr>
<tr>
<td>Skew-Hermitian matrix</td>
<td>$A^* = -A$</td>
<td>$\begin{pmatrix} 3i &amp; 1 - 2i \ -1 - 2i &amp; i \end{pmatrix}$</td>
</tr>
<tr>
<td>Unitary matrix</td>
<td>$A^* = A^{-1}$</td>
<td>$\begin{pmatrix} 1 + i &amp; 1 \sqrt{2} \ 2 &amp; -1 \sqrt{2} \end{pmatrix}$</td>
</tr>
</tbody>
</table>

For a real matrix we use different names.

<table>
<thead>
<tr>
<th>NAME</th>
<th>DEFINITION</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real symmetric matrix</td>
<td>$A^T = A$</td>
<td>$\begin{pmatrix} 3 &amp; -2 \ -2 &amp; 1 \end{pmatrix}$</td>
</tr>
<tr>
<td>Real skew-symmetric matrix</td>
<td>$A^T = -A$</td>
<td>$\begin{pmatrix} 0 &amp; -2 \ 2 &amp; 0 \end{pmatrix}$</td>
</tr>
<tr>
<td>Orthogonal matrix</td>
<td>$A^T = A^{-1}$</td>
<td>$\begin{pmatrix} 1 &amp; \frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} &amp; -\frac{1}{\sqrt{2}} \end{pmatrix}$</td>
</tr>
</tbody>
</table>

**Theorem 7**: A square matrix over $\mathbb{C}$ is a normal matrix if and only if there exists a unitary matrix $U$ and a diagonal matrix $D$ such that $U^*AU = D$.

**Proof**: Suppose $U^*AU = D$.
Then $A = UDU^*$ and $A^* = UD^*U^*$.
So $AA^* = UDU^*UD^*U^* = UDD^*U^*$.
But $D, D^*$ are diagonal matrices so $DD^* = D^*D$.
Hence $AA^* = UD^*DU^* = (UD^U)(UDU^*) = A^*A$.

We now prove a number of properties of normal matrices.

**Theorem 8**: Suppose that $A$ is a normal matrix. Then $E_{A}(\lambda) \leq E_{A^*}(\bar{\lambda})$.

**Proof**: Let $v \in E_{A^*}(\bar{\lambda})$.
Then $\langle A^*v - \bar{\lambda}v, A^*v - \bar{\lambda}v \rangle = \langle A^*v, A^*v \rangle - \langle \bar{\lambda}v, A^*v \rangle - \langle A^*v, \bar{\lambda}v \rangle + \langle \bar{\lambda}v, \bar{\lambda}v \rangle$

$= \langle v, AA^*v \rangle - \bar{\lambda} \langle v, A^*v \rangle - \lambda \langle A^*v, v \rangle + \bar{\lambda} \bar{\lambda} \langle v, v \rangle$

$= \langle v, A^*Av \rangle - \bar{\lambda} \langle A^*v, v \rangle - \lambda \langle v, Av \rangle + \bar{\lambda} \bar{\lambda} \langle v, v \rangle$

$= \langle v, \lambda v \rangle - \bar{\lambda} \lambda \langle v, v \rangle - \lambda \bar{\lambda} \langle v, v \rangle + \bar{\lambda} \bar{\lambda} \langle v, v \rangle$

$= \bar{\lambda} \lambda \langle v, v \rangle - \bar{\lambda} \lambda \langle v, v \rangle = 0$.

Hence $A^*v - \bar{\lambda}v = 0$ and so $v \in E_{A^*}(\bar{\lambda})$. 

32
**Theorem 9:** Suppose that $A$ is a normal matrix. Then eigenvectors corresponding to different eigenvalues are orthogonal.

**Proof:** Let $v \in E_A(\lambda)$ and $w \in E_A(\mu)$ where $\lambda \neq \mu$.

Then $\lambda \langle v | w \rangle = \langle \lambda v | w \rangle$

$= \langle Av | w \rangle$

$= \langle v | A^* w \rangle$

$= \langle v | \mu w \rangle = \mu \langle v | w \rangle.$

Since $\lambda \neq \mu$ then $\langle v | w \rangle = 0.$

**Theorem 10:** Suppose that $A$ is a normal matrix. Then $E_A$ has an orthonormal basis of eigenvectors.

**Proof:** Let the distinct eigenvectors of $A$ be $\lambda_1, \ldots, \lambda_k$.

Then $E_A = E_A(\lambda_1) + \ldots + E_A(\lambda_k)$.

For each $i$ choose an orthonormal basis for $E_A(\lambda_i)$ and take the union of these bases. By Theorem 9 this basis will be an orthonormal basis of $E_A$.

**Theorem 11:** Suppose that $A$ is a normal matrix and let $v \in E_A^\perp$. Then $Av \in E_A^\perp$.

In other words, $E_A^\perp$ is invariant under multiplication by $A$.

**Proof:** Let $v \in E_A^\perp$ and let $w \in E_A$.

Then $w = v_1 + \ldots + v_k$ where each $v_i \in E_A(\lambda_i)$.

$\therefore A^*w = A^*v_1 + \ldots + A^*v_k$

$= \tilde{\lambda}_1 v_1 + \ldots + \tilde{\lambda}_k v_k$ by Theorem 8.

$\in E_A.$

Hence $\langle v | A^*w \rangle = 0$ and so $\langle Av | w \rangle = 0$.

Since this holds for all $w \in E_A$, $Av \in E_A^\perp$.

**Theorem 12:** Suppose that $A$ is a normal matrix. Then the map $f: E_A^\perp \to E_A^\perp$ defined by $f(v) = Av$ is a linear transformation.

**Proof:** It is the restriction of $v \to Av$ to $E_A^\perp$.

The only thing that is not obvious is the fact that $f(v) \in E_A^\perp$ for all $v \in E_A^\perp$, and that was proved in Theorem 11.

**Theorem 13:** Suppose that $A$ is a normal matrix. Then $E_A^\perp = 0$.

**Proof:** Suppose $\dim(E_A^\perp) \geq 1$.

Then $f$ has at least one eigenvalue $\lambda$ and a corresponding eigenvector $v$.

Clearly $v \in E_A$, a contradiction.

**Theorem 14:** Suppose that $A$ is a normal matrix. Then $A$ is unitarily diagonalisable.

**Proof:** If $A$ is an $n \times n$ matrix then we have shown that $E_A = C^n$ and hence $C^n$ has an orthonormal basis of eigenvectors $\{v_1, \ldots, v_n\}$.

If $U = (v_1, \ldots, v_n)$ is the matrix whose columns are the $v_i$ then $U$ is unitary and $U^{-1}AU = D$ for some diagonal matrix $D$. Since $U$ is unitary we can also write $U^{-1}$ as $U^*$. 

33