1. Matrices

Suppose that $F$ is a field. $F^n$ will denote the set of $n \times 1$ column vectors with coefficients in $F$, and $F^m$ will denote the set of $1 \times m$ row vectors with coefficients in $F$. $M_{m,n}(F)$ will denote the set of $m \times n$ matrices $A = (a_{ij})$ with coefficients in $F$. We will sometimes denote the zero vector of $F^n$ by $0^n$, the zero vector of $F_m$ by $0_m$ and the zero vector of $M_{m,n}$ by $0_{m,n}$. We will also sometimes abbreviate these different zeros as $\vec{0}$ or 0. We will let $e_i \in F^n$ be the column vector whose entries are all zero, except for a 1 in the $i$-th row. If $A \in M_{m,n}(F)$ then $A^t \in M_{n,m}(F)$ will denote the transpose of $A$. We will write $A = (A_1, A_2, \ldots, A_n)$ where $A_i \in F_m$ are the columns of $A$, and
definition
$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}$$
definition
where $A_j \in F_n$ are the rows of $A$. For $x = (x_1, \ldots, x_n)^t \in F^n$, we have the formula
$$Ax = x_1 A^1 + x_2 A^2 + \cdots + x_n A^n.$$  
In particular, $A e_i = A^i$ for $1 \leq i \leq n$ and $(e_i)^t A e_j = a_{ij}$ for $1 \leq i \leq m$, $1 \leq j \leq n$. We also have
$$Ax = \begin{pmatrix} A_1 x \\ A_2 x \\ \vdots \\ A_m x \end{pmatrix} = \begin{pmatrix} (A_1)^t \cdot x \\ (A_2)^t \cdot x \\ \vdots \\ (A_m)^t \cdot x \end{pmatrix}$$
definition
where for $x = (x_1, \ldots, x_n)^t, y = (y_1, \ldots, y_n) \in F^n$, $x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$ is the dot product on $F^n$.

Lemma 1.1. The following are equivalent for two matrices $A, B \in M_{m,n}(F)$.

1. $A = B$.
2. $Ax = Bx$ for all $x \in F^n$.
3. $A e_i = B e_i$ for $1 \leq i \leq n$.

$\mathbb{N}$ will denote the natural numbers $\{0,1,2\ldots\}$. $\mathbb{Z}_+$ is the set of positive integers $\{1,2,3,\ldots\}$.
In this section we suppose that \( V \) is a vector space over a field \( F \). We will denote the zero vector in \( V \) by \( 0_V \). Sometimes, we will abbreviate, and write the zero vector as \( \vec{0} \) or \( 0 \).

**Definition 2.1.** Suppose that \( S \) is a nonempty subset of \( V \). Then
\[
\text{Span}(S) = \{ c_1 v_1 + \cdots + c_n v_n \mid n \in \mathbb{Z}_+, v_1, \ldots, v_n \in S \text{ and } c_1, \ldots, c_n \in F \}.
\]
This definition is valid even when \( S \) is an infinite set. We define \( \text{Span}\{\emptyset\} = \{ \vec{0} \} \).

**Lemma 2.2.** Suppose that \( S \) is a subset of \( V \). Then \( \text{Span}(S) \) is a subspace of \( V \).

**Definition 2.3.** Suppose that \( S \) is a subset of \( V \). \( S \) is a linearly dependent set if there exists \( n \in \mathbb{Z}_+, v_1, \ldots, v_n \in S \) and \( c_1, \ldots, c_n \in F \) such that \( c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0 \). \( S \) is a linearly independent set if it is not linearly dependent.

Observe that \( \emptyset \) is a linearly independent set.

**Definition 2.4.** Suppose that \( S \) is subset of \( V \) which is a linearly independent set and \( \text{Span}(S) = V \). Then \( S \) is a basis of \( V \).

**Theorem 2.5.** Suppose that \( V \) is a vector space over a field \( F \). Then \( V \) has a basis.

We have that the empty set \( \emptyset \) is a basis of the 0 vector space \( \{ \vec{0} \} \).

**Theorem 2.6.** (Extension Theorem) Suppose that \( V \) is a vector space over a field \( F \) and \( S \) is a subset of \( V \) which is a linearly independent set. Then there exists a basis of \( V \) which contains \( S \).

**Theorem 2.7.** Suppose that \( V \) is a vector space over a field \( F \) and \( S_1 \) and \( S_2 \) are two bases of \( V \). Then \( S_1 \) and \( S_2 \) have the same cardinality.

This theorem allows us to make the following definition.

**Definition 2.8.** Suppose that \( V \) is a vector space over a field \( F \). Then the dimension of \( V \) is the cardinality of a basis of \( V \).

The dimension of \( \{ \vec{0} \} \) is 0.

**Definition 2.9.** \( V \) is called a finite dimensional vector space if \( V \) has a finite basis.

For the most part, we will consider finite dimensional vector spaces. If \( V \) is finite dimensional of dimension \( n \), a basis is considered as an ordered set \( \beta = \{ v_1, \ldots, v_n \} \).

**Lemma 2.10.** Suppose that \( V \) is a finite dimensional vector space and \( W \) is a subspace of \( V \). Then \( \dim W \leq \dim V \), and \( \dim W = \dim V \) if and only if \( V = W \).

**Lemma 2.11.** Suppose that \( V \) is a finite dimensional vector space and \( \beta = \{ v_1, v_2, \ldots, v_n \} \) is a basis of \( V \). Suppose that \( v \in V \). Then \( v \) has a unique expression
\[
v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n
\]
with \( c_1, \ldots, c_n \in F \).

**Lemma 2.12.** Let \( \{ v_1, \ldots, v_n \} \) be a set of generators of a vector space \( V \). Let \( \{ v_1, \ldots, v_r \} \) be a maximal subset of linearly independent elements. Then \( \{ v_1, \ldots, v_r \} \) is a basis of \( V \).
3. Direct Sums

Suppose that $V$ is a vector space over a field $F$, and $W_1, W_2, \ldots, W_m$ are subspaces of $V$. Then the sum

\[ W_1 + \cdots + W_m = \text{Span}(W_1 \cup \cdots \cup W_m) = \{ w_1 + w_2 + \cdots + w_m \mid w_i \in W_i \text{ for } 1 \leq i \leq m \} \]

is the subspace of $V$ spanned by $W_1, W_2, \ldots, W_m$.

The sum $W_1 + \cdots + W_m$ is called a direct sum, denoted by $W_1 \oplus W_2 \oplus \cdots \oplus W_m$, if every element $v \in W_1 + W_2 + \cdots + W_m$ has a unique expression $v = w_1 + \cdots + w_m$ with $w_i \in W_i$ for $1 \leq i \leq m$.

We have the following useful criterion.

**Lemma 3.1.** The sum $W_1 + W_2 + \cdots + W_m$ is a direct sum if and only if $0 = w_1 + w_2 + \cdots + w_m$ with $w_i \in W_i$ for $1 \leq i \leq m$ implies $w_i = 0$ for $1 \leq i \leq m$.

In the case when $V$ is finite dimensional, the direct sum is characterized by the following equivalent conditions.

**Lemma 3.2.** Suppose that $V$ is a finite dimensional vector space. Let $n = \dim (W_1 + \cdots + W_m)$ and $n_i = \dim W_i$ for $1 \leq i \leq m$. The following conditions are equivalent

1. If $v \in W_1 + W_2 + \cdots + W_m$ then $v$ has a unique expression $v = w_1 + \cdots + w_m$ with $w_i \in W_i$ for $1 \leq i \leq m$.
2. Suppose that $\beta_i = \{ w_{i,1}, \ldots, w_{i,n_i} \}$ are bases of $W_i$ for $1 \leq i \leq m$. Then $\beta = \{ w_{1,1}, \ldots, w_{n_1,1}, w_{2,1}, \ldots, w_{n_2,1}, \ldots, w_{n_m,1}, \ldots, w_{n_m,n_m} \}$ is a basis of $W_1 + \cdots + W_m$.
3. $n = n_1 + \cdots + n_m$.

The following lemma gives a useful criterion.

**Lemma 3.3.** Suppose that $U$ and $W$ are subspaces of a vector space $V$. Then $V = U \oplus W$ if and only if $V = U + W$ and $U \cap W = \{0\}$.

4. Linear Maps

In this section, we will suppose that all vector spaces are over a fixed field $F$.

We recall the definition of equality of maps, which we will use repeatedly to show that two maps are the same. Suppose that $f : V \to W$ and $g : V \to W$ are maps (functions). Then $f = g$ if and only if $f(v) = g(v)$ for all $v \in V$.

If $f : U \to V$ and $g : V \to W$, we will usually write the composition $g \circ f : U \to W$ as $gf$. If $x \in U$, we will sometimes write $fx$ for $f(x)$, and $gf$ for $(g \circ f)(x)$.

**Definition 4.1.** Suppose that $V$ and $W$ are vector spaces. A map $L : V \to W$ is linear if $L(v_1 + v_2) = L(v_1) + L(v_2)$ for $v_1, v_2 \in V$ and $L(cv) = cL(v)$ for $v \in V$ and $c \in F$.

**Lemma 4.2.** Suppose that $L : V \to W$ is a linear map. Then

1. $L(0_V) = 0_W$.
2. $L(-v) = -L(v)$ for $v \in V$.

**Lemma 4.3.** Suppose that $L : V \to W$ is a linear map and $T : W \to U$ is a linear map. Then the composition $TL : V \to U$ is a linear map.

If $S$ and $T$ are sets, a map $f : S \to T$ is 1-1 and onto if and only if there exists a map $g : T \to S$ such that $g \circ f = I_S$ and $f \circ g = I_T$, where $I_S$ is the identity map of $S$ and $I_T$ is
the identity map of $T$. When this happens, $g$ is uniquely determined. We write $g = f^{-1}$ and say that $g$ is the inverse of $f$.

**Definition 4.4.** Suppose that $L : V \to W$ is a linear map. $L$ is an isomorphism if there exists a linear map $T : W \to V$ such that $TL = I_V$ is the identity map of $V$, and $LT = I_W$ is the identity map of $W$.

The map $T$ in the above definition is unique, if it exists, by the following lemma. If $L$ is an isomorphism, we write $L^{-1} = T$, and call $T$ the inverse of $L$.

**Lemma 4.5.** Suppose that $L : V \to W$ is a map. Suppose that $T : W \to V$ and $S : W \to V$ are maps which satisfy $TL = I_V$, $LT = I_W$ and $SL = I_V$, $LS = I_W$. Then $S = T$.

**Lemma 4.6.** A linear map $L : V \to W$ is an isomorphism if and only if $L$ is 1-1 and onto.

**Lemma 4.7.** Suppose that $L : V \to W$ is a linear map. Then

1. $\text{Image } L = \{L(v) \mid v \in V\}$ is a subspace of $W$.
2. $\text{Kernel } L = \{v \in V \mid L(v) = 0\}$ is a subspace of $V$.

**Lemma 4.8.** Suppose that $L : V \to W$ is a linear map. Then $L$ is 1-1 if and only if $\text{Kernel } L = \{0\}$.

**Definition 4.9.** Suppose that $F$ is a field, and $V$, $W$ are vector spaces over $F$. Let $\mathcal{L}_F(V, W)$ be the set of linear maps from $V$ to $W$.

We will sometimes write $L(V, W) = \mathcal{L}_F(V, W)$ if the field is understood to be $F$.

**Lemma 4.10.** Suppose that $F$ is a field, and $V$, $W$ are vector spaces over $F$. Then $\mathcal{L}_F(V, W)$ is a vector space.

**Theorem 4.11.** (Dimension Theorem) Suppose that $L : V \to W$ is a homomorphism of finite dimensional vector spaces. Then

$$\dim \text{ Kernel } L + \dim \text{ Image } L = \dim V.$$  

**Proof.** Let $\{v_1, \ldots, v_r\}$ be a basis of $\text{Kernel } L$. Extend this to a basis $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ of $V$. Let $w_i = L(v_i)$ for $r + 1 \leq i \leq n$. We will show that $\{w_{r+1}, \ldots, w_n\}$ is a basis of $\text{Image } L$. Since $L$ is linear, $\{w_{r+1}, \ldots, w_n\}$ span $\text{Image } L$. We must show that $\{w_{r+1}, \ldots, w_n\}$ are linearly independent. Suppose that we have a relation

$$c_{r+1}w_{r+1} + \cdots + c_nw_n = 0$$

for some $c_{r+1}, \ldots, c_n \in F$. Then

$$L(c_{r+1}v_{r+1} + \cdots + c_nv_n) = c_{r+1}L(v_{r+1}) + \cdots + c_nL(v_n) = c_{r+1}w_{r+1} + \cdots + c_nw_n = 0.$$  

Thus $c_{r+1}v_{r+1} + \cdots + c_nv_n \in \text{Kernel } L$, so that we have an expansion

$$c_{r+1}v_{r+1} + \cdots c_nv_n = d_1v_1 + \cdots + d_r v_r$$

for some $d_1, \ldots, d_r \in F$, which we can rewrite as

$$d_1v_1 + \cdots + d_r v_r - c_{r+1}v_{r+1} - \cdots - c_nv_n = 0.$$  

Since $\{v_1, \ldots, v_r\}$ are linearly independent, we have that $d_1 = \cdots = d_r = c_{r+1} = \cdots = c_n = 0$. Thus $\{w_{r+1}, \ldots, w_n\}$ are linearly independent. \qed
Corollary 4.12. Suppose that $V$ is a finite dimensional vector space and $\Phi : V \to V$ and $\Psi : V \to V$ are linear maps such that $\Psi \Phi = I_V$. Then $\Phi$ and $\Psi$ are isomorphisms and $\Psi = \Phi^{-1}$.

Theorem 4.13. (Universal Property of Vector Spaces) Suppose that $V$ and $W$ are finite dimensional vector spaces, $\beta = \{v_1, \ldots, v_n\}$ is a basis of $V$, and $w_1, \ldots, w_n \in W$. Then there exists a unique linear map $L : V \to W$ such that $L(v_i) = w_i$ for $1 \leq i \leq n$.

Proof. We first prove existence. For $v \in V$, define $L(v) = c_1v_1 + c_2v_2 + \cdots + c_nv_n$ if $v = c_1v_1 + \cdots + c_nv_n$ with $c_1, \ldots, c_n \in F$. $L$ is a well-defined map since every $v \in V$ has a unique expression $v = c_1v_1 + \cdots + c_nv_n$ with $c_1, \ldots, c_n \in F$, as $\beta$ is a basis of $V$. We have that $L(v_i) = w_i$ for $1 \leq i \leq n$. We leave the verification that $L$ is linear to the reader.

Now we prove uniqueness. Suppose that $T : V \to W$ is a linear map such that $T(v_i) = w_i$ for $1 \leq i \leq n$. Suppose that $v \in V$. Then $v = c_1v_1 + \cdots + c_nv_n$ for some $c_1, \ldots, c_n \in F$. We have that

$$T(v) = T(c_1v_1 + \cdots + c_nv_n) = c_1T(v_1) + \cdots + c_nT(v_n) = c_1w_1 + \cdots + c_nw_n = L(v).$$

Since $T(v) = L(v)$ for all $v \in V$, we have that $T = L$. \hfill \Box

Lemma 4.14. Suppose that $A \in M_{m,n}(F)$. Then the map $L_A : F^n \to F^m$ defined by $L_A(x) = Ax$ for $x \in F^n$ is linear.

Lemma 4.15. Suppose that $A = (A^1, \ldots, A^n) \in M_{m,n}(F)$. Then

$$\text{Image } L_A = \text{Span}\{A^1, \ldots, A^n\}.$$

$$\text{Kernel } L_A = \{x \in F^n \mid Ax = 0\}$$

is the solution space to $Ax = 0$.

Lemma 4.16. Suppose that $A, B \in M_{m,n}(F)$. Then $L_{A+B} = L_A + L_B$. Suppose that $c \in F$. Then $cL_A = L_{cA}$. Suppose that $A \in M_{m,n}(F)$ and $C \in M_{l,m}$. Then the composition of maps $L_CL_A = L_{CA}$.

Lemma 4.17. The map $\Phi : M_{m,n} \to \mathcal{L}_F(F^n, F^m)$ defined by $\Phi(A) = L_A$ for $A \in M_{m,n}(F)$ is an isomorphism.

Proof. By Lemma 4.16, $\Phi$ is a linear map. We will now show that Kernel $\Phi$ is $\{0\}$. Let $A \in \text{Kernel } \Phi$. Then $L_A = 0$. Thus $Ax = 0$ for all $x \in F^n$. By Lemma 1.1, we have that $A = 0$. Thus Kernel $\Phi = \{0\}$ and $\Phi$ is 1-1. Suppose that $L \in \mathcal{L}_F(F^n, F^m)$. Let $w_i = L(e_i)$ for $1 \leq i \leq n$. Let $A = (w_1, \ldots, w_n) \in M_{m,n}(F)$. For $1 \leq i \leq n$, we have that $L_A(e_i) = Ae_i = w_i$. By Theorem 4.13, $L = L_A$. Thus $\Phi$ is onto, and $\Phi$ is an isomorphism. \hfill \Box

Definition 4.18. Suppose that $V$ is a finite dimensional vector space. Suppose that $\beta = \{v_1, \ldots, v_n\}$ is a basis of $V$. Define the coordinate vector $(v)_\beta \in F^n$ of $v$ with respect to $\beta$ by

$$(v)_\beta = (c_1, \ldots, c_n)^T,$$

where $c_1, \ldots, c_n \in F$ are the unique elements of $F$ such that $v = c_1v_1 + \cdots + c_nv_n$.

Lemma 4.19. Suppose that $V$ is a finite dimensional vector space. Suppose that $\beta = \{v_1, \ldots, v_n\}$ is a basis of $V$. Suppose that $u_1, u_2 \in V$ and $c \in F$. Then $(u_1 + u_2)_\beta = (u_1)_\beta + (u_2)_\beta$ and $(cu_1)_\beta = c(u_1)_\beta$.

Theorem 4.20. Suppose that $V$ is a finite dimensional vector space. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis of $V$. Then the map $\Phi : V \to F^n$ defined by $\Phi(v) = (v)_\beta$ is an isomorphism.
Proof. \( \Phi \) is a linear map by Lemma 4.19. Note that \( \Phi(v_i) = e_i \) for \( 1 \leq i \leq n \). By Theorem 4.13, there exists a unique linear map \( \Psi : F^n \to V \) such that \( \Psi(e_i) = v_i \) for \( 1 \leq i \leq n \) (where \( \{e_1, \ldots, e_n\} \) is the standard basis of \( F^n \)). Now \( \Psi \Phi : V \to V \) is a linear map which satisfies \( \Psi \Phi(v_i) = v_i \) for \( 1 \leq i \leq n \). By Theorem 4.13, there is a unique linear map from \( V \to V \) which takes \( v_i \) to \( v_i \) for \( 1 \leq i \leq n \). Since the identity map \( I_V \) of \( V \) has this property, we have that \( \Psi \Phi = I_V \). By a similar calculation, \( \Phi \Psi = I_{F^n} \). Thus \( \Phi \) is an isomorphism (with inverse \( \Psi \)). \( \square \)

**Definition 4.21.** Suppose that \( V \) and \( W \) are finite dimensional vector spaces. Suppose that \( \beta = \{v_1, \ldots, v_n\} \) is a basis of \( V \) and \( \beta' = \{w_1, \ldots, w_m\} \) is a basis of \( W \). Suppose that \( L : V \to W \) is a linear map. Define the matrix \( M^\beta_{\beta'}(L) \in \mathbb{M}_{m,n}(F) \) of \( L \) with respect to the bases \( \beta \) and \( \beta' \) to be the matrix

\[
M^\beta_{\beta'}(L) = ((L(v_1)_{\beta'}, (L(v_2))_{\beta'}, \ldots, (L(v_n))_{\beta'}).
\]

The matrix \( M^\beta_{\beta'}(L) \) of \( L \) with respect to \( \beta \) and \( \beta' \) has the following important property:

\[ M^\beta_{\beta'}(L)(v)_\beta = (L(v))_{\beta'} \]

for all \( v \in V \). In fact, \( M^\beta_{\beta'}(L) \) is the unique matrix \( A \) such that \( A(v)_\beta = (L(v))_{\beta'} \) for all \( v \in V \).

**Lemma 4.22.** Suppose that \( V \) and \( W \) are finite dimensional vector spaces. Suppose that \( \beta = \{v_1, \ldots, v_n\} \) is a basis of \( V \) and \( \beta' = \{w_1, \ldots, w_m\} \) is a basis of \( W \). Suppose that \( L_1, L_2 \in \mathcal{L}_F(V,W) \) and \( c \in F \). Then \( M^\beta_{\beta'}(L_1 + L_2) = M^\beta_{\beta'}(L_1) + M^\beta_{\beta'}(L_2) \) and \( M^\beta_{\beta'}(cL_1) = cM^\beta_{\beta'}(L_1) \).

**Theorem 4.23.** Suppose that \( V \) and \( W \) are finite dimensional vector spaces. Suppose that \( \beta = \{v_1, \ldots, v_n\} \) is a basis of \( V \) and \( \beta' = \{w_1, \ldots, w_m\} \) is a basis of \( W \). Then the map \( \Lambda : \mathcal{L}_F(V,W) \to \mathbb{M}_{m,n}(F) \) defined by \( \Lambda(L) = M^\beta_{\beta'}(L) \) is an isomorphism.

Proof. \( \Lambda \) is a linear map by Lemma 4.22. It remains to verify that \( \Lambda \) is 1-1 and onto, from which it will follow that \( \Lambda \) is an isomorphism. Suppose that \( L_1, L_2 \in \mathcal{L}_F(V,W) \) and \( \Lambda(L_1) = \Lambda(L_2) \). Then \( M^\beta_{\beta'}(L_1) = M^\beta_{\beta'}(L_2) \) so that \( (L_1(v_i))_{\beta'} = (L_2(v_i))_{\beta'} \) for \( 1 \leq i \leq n \). Thus \( L_1(v_i) = L_2(v_i) \) for \( 1 \leq i \leq n \). Since \( \beta \) is a basis of \( V \), \( L_1 = L_2 \) by the uniqueness statement of Theorem 4.13. Thus \( \Lambda \) is 1-1. Suppose that \( A = (a_{ij}) \in \mathbb{M}_{m,n}(F) \). Write \( A = (A^1, \ldots, A^n) \) where \( A^i = (a_{1,i}, \ldots, a_{m,i}) \in F^m \) are the columns of \( A \). Let \( z_i = \sum_{j=1}^m a_{ji}w_j \in W \) for \( 1 \leq i \leq n \). By Theorem 4.13, there exists a linear map \( L : V \to W \) such that \( L(v_i) = z_i \) for \( 1 \leq i \leq n \). We have that \( (L(v_i))_{\beta'} = (z_i)_{\beta'} = A^i \) for \( 1 \leq i \leq n \). We have that \( \Lambda(L) = M^\beta_{\beta'}(L) = (A^1, \ldots, A^n) = A \), and thus \( \Lambda \) is onto. \( \square \)

**Corollary 4.24.** Suppose that \( V \) is a vector space of dimension \( n \) and \( W \) is a vector space of dimension \( m \). Then \( \mathcal{L}_F(V,W) \) is a vector space of dimension \( mn \).

**Lemma 4.25.** Suppose that \( U, V \) and \( W \) are finite dimensional vector spaces, with respective bases \( \beta_1, \beta_2, \beta_3 \). Suppose that \( L_1 \in \mathcal{L}_F(U,V) \) and \( L_2 \in \mathcal{L}_F(V,W) \). Then

\[
M^\beta_{\beta_3}(L_2L_1) = M^\beta_{\beta_3}(L_2)M^\beta_{\beta_2}(L_1).
\]

**Theorem 4.26.** Suppose that \( V \) is a finite dimensional vector space, and \( \beta \) is a basis of \( V \). Then the map \( \Lambda : \mathcal{L}_F(V,V) \to \mathbb{M}_{n,n}(F) \) defined by \( \Lambda(L) = M^\beta_{\beta}(L) \) is a ring isomorphism.
5. Bilinear Forms

Let $U, V, W$ be vector spaces over a field $F$. Let $\text{Bil}(U \times V, W)$ be the set of bilinear maps from $U \times V$ to $W$. $\text{Bil}(U \times V, W)$ is a vector space over $F$. We call an element of $\text{Bil}_F(U \times V, F)$ a bilinear form. A bilinear form is usually written as a pairing $<v, w> \in F$ for $v, w \in V$. A bilinear form $<,>$ is symmetric if $<v, w> = <w, v>$ for all $v, w \in V$. A symmetric bilinear form is nondegenerate if whenever $v \in V$ is such that $<v, w> = 0$ for all $w \in V$ then $v = 0$.

**Theorem 5.1.** Let $F$ be a field and $A \in M_{m,n}(F)$. Let $g_A : F^m \times F^n \rightarrow F$ be defined by $g_A(v, w) = v'^t A w$. Then $g_A \in B_1(F^m \times F^n, F)$. Further, the map $\Psi : M_{m,n}(F) \rightarrow B_1(F^m \times F^n, F)$ defined by $\Psi(A) = g_A$ is an isomorphism of $F$-vector spaces.

Also, in the case when $m = n$,

1. $\Psi$ gives a 1-1 correspondence between $n \times n$ matrices and bilinear forms on $F^n$.
2. $\Psi$ gives a 1-1 correspondence between $n \times n$ symmetric matrices and symmetric forms on $F^n$.
3. $\Psi$ gives a 1-1 correspondence between invertible $n \times n$ symmetric matrices and nondegenerate symmetric forms.

From now on in this section, suppose that $V$ is a vector space over a field $F$, and $<,>$ is a nondegenerate symmetric form on $V$. An important example is the dot product on $F^n$. We begin by stating the identity

$$<v, w> = \frac{1}{2} [<v + w, v + w> - <v, v> - <w, w>]$$

for $v, w \in V$.

Define $v, w \in V$ to be orthogonal (or perpendicular) if $<v, w> = 0$. Suppose that $S \subset V$ is a subset. Define

$$S^\perp = \{v \in V | <v, w> = 0 \text{ for all } w \in S\}.$$

**Lemma 5.2.** Suppose that $S$ is a subset of $V$. Then

1. $S^\perp$ is a subspace of $V$.
2. If $U$ is the subspace $\text{Span}(S)$ of $V$, then $U^\perp = S^\perp$.

**Lemma 5.3.** Suppose

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \in M_{m,n}(F).$$

Let $U = \text{Span}\{A_1^t, \ldots, A_m^t\} \subset F^n$, which is the column space of $A^t$. Then $U^\perp = \text{Kernel } L_A$.

A basis $\beta = \{v_1, \ldots, v_n\}$ of $V$ is called an orthogonal basis if $<v_i, v_j> = 0$ whenever $i \neq j$.

**Theorem 5.4.** Suppose that $V$ is a finite dimensional vector space with a symmetric nondegenerate bilinear form $<,>$. Then $V$ has an orthogonal basis.

**Proof.** We prove the theorem by induction on $n = \dim V$. If $n = 1$, the theorem is true since any basis is an orthogonal basis.

Assume that $\dim V = n > 1$, and that any subspace of $V$ of dimension less than $n$ has an orthogonal basis. There are two cases. The first case is when every $v \in V$ satisfies
< v, v >= 0. Then by (2), we have < v, w >= 0 for every v, w ∈ V. Thus any basis of V is an orthogonal basis.

Now assume that there exists v₁ ∈ V such that < v₁, v₁ > ≠ 0. Let U = Span{v₁}, a one dimensional subspace of V. Suppose that v ∈ V. Let
c = \frac{< v, v₁ >}{< v₁, v₁ >} ∈ F.
Then v - cv₁ ∈ U⊥, and thus v = cv₁ + (v - cv₁) ∈ U + U⊥. It follows that V = U + U⊥. We have that U ∩ U⊥ = {0}, since < v₁, v₁ > ≠ 0. Thus V = U ⊕ U⊥. Since U⊥ has dimension n - 1 by Lemma 3.2, we have by induction that U⊥ has an orthogonal basis {v₂, ..., vₙ}. Thus {v₁, v₂, ..., vₙ} is an orthogonal basis of V.

Theorem 5.5. Let V be a finite dimensional vector space over a field F, with a symmetric bilinear form < , >. Assume dim V > 0. Let V₀ be the subspace of V defined by

V₀ = V ⊥ = {v ∈ V | < v, w >/= 0 for all w ∈ V}.

Let {v₁, ..., vₙ} be an orthogonal basis of V. Then the number of integers i such that < vᵢ, vᵢ >/= 0 is equal to the dimension of V₀.

Proof. Order the {v₁, ..., vₙ} so that < vᵢ, vᵢ > ≥ 0 for 1 ≤ i ≤ r and < vᵢ, vᵢ > ≠ 0 for r < i ≤ n. Suppose that w ∈ V. Then w = c₁v₁ + ... + cₙvₙ for some c₁, ..., cₙ ∈ F.

< vᵢ, w >/= \sum_{j=1}^{n} cᵢ < vᵢ, vⱼ >/= cᵢ < vᵢ, vᵢ >/= 0

for 1 ≤ i ≤ r. Thus v₁, ..., vᵣ ∈ V₀.

Now suppose that v ∈ V₀. We have v = d₁v₁ + ... + dₙwₙ for some d₁, ..., dₙ ∈ F. We have

0 =< vᵢ, v >/= \sum_{j=1}^{n} dⱼ < vᵢ, vⱼ >/= dᵢ < vᵢ, vᵢ >

for 1 ≤ i ≤ r. Since < vᵢ, vᵢ > ≠ 0 for r < i ≤ n we have dᵢ = 0 for r < i ≤ n. Thus v = d₁v₁ + ... + dᵣvᵣ ∈ Span{v₁, ..., vᵣ}. Thus V₀ = Span{v₁, ..., vᵣ}. Since v₁, ..., vᵣ are linearly independent, they are a basis of V₀, and thus dim V₀ = r.

The dimension of V₀ in Theorem 5.5 is called the index of nullity of the form.

Theorem 5.6. (Sylvester’s Theorem) Let V be a finite dimensional vector space over R with a symmetric bilinear form. There exists an integer r with the following property. If {v₁, ..., vₙ} is an orthogonal basis of V, then there are precisely r integers i such that < vᵢ, vᵢ >/= 0.

Proof. Let {v₁, ..., vₙ} and {w₁, ..., wₙ} be orthogonal bases of V. We may suppose that their elements are arranged so that < vᵢ, vᵢ >/= 0 is 1 ≤ i ≤ r, < vᵢ, vᵢ >/= 0 for r + 1 ≤ i ≤ s and < wᵢ, wᵢ >/= 0 if s + 1 ≤ i ≤ n. Similarly, < wᵢ, wᵢ >/= 0 is 1 ≤ i ≤ r′, < wᵢ, wᵢ >/= 0 if r′ + 1 ≤ i ≤ s′. If s′ + 1 ≤ i ≤ n.

We first prove that v₁, ..., vᵣ, wᵣ₊₁, ..., wₙ are linearly independent. Suppose we have a relation

x₁v₁ + ... + xᵣvᵣ + yᵣ₊₁wᵣ₊₁ + ... + yₙwₙ = 0

for some x₁, ..., xᵣ, yᵣ₊₁, ..., yₙ ∈ F. Then

x₁v₁ + ... + xᵣvᵣ = -yᵣ₊₁wᵣ₊₁ + ... + yₙwₙ).
Since the left hand side is $\geq$ preceding equation with itself, we have $y$ independent.
Let $c$ of the two bases in the proof, we also obtain form.

$$\dim V = r$$

The integer $r$ in the proof of Sylvester’s theorem is call the index of positivity of the form.

6. Exercises

1. Let $V$ be an $n$-dimensional vector space over a field $F$. The dual space of $V$ is the vector space $V^* = L(V, F)$. Elements of $V^*$ are called functionals.

i) Suppose that $\beta = \{v_1, \ldots, v_n\}$ is a basis of $V$. For $1 \leq i \leq n$, define $v^*_i \in V^*$ by $v^*_i(v) = c_i$ if $v = c_1v_1 + \cdots + c_nv_n \in V$ (so that $v^*_i(v_j) = \delta_{ij}$). Prove that the $v^*_i$ are elements of $V^*$ and that they form a basis of $V^*$. (This basis is called the dual basis to $\beta$.)

ii) Suppose that $V$ has a nondegenerate symmetric bilinear form $<,>$. For $v \in V$ define $L_v : V \to F$ by $L_v(w) = <v, w>$ for $w \in V$. Show that $L_v \in V^*$, and that the map $\Phi : V \to V^*$ defined by $\Phi(v) = L_v$ is an isomorphism of vector spaces.

2. Suppose that $L : V \to W$ is a linear map of finite dimensional vector spaces. Define $\hat{L} : W^* \to V^*$ by $\hat{L}(\varphi) = \varphi L$ for $\varphi \in W^*$.

i) Show that $\hat{L}$ is a linear map.

ii) Suppose that $L$ is 1-1. Is $\hat{L}$ 1-1? Is $\hat{L}$ onto? Prove your answers.

iii) Suppose that $L$ is onto. Is $\hat{L}$ 1-1? Is $\hat{L}$ onto? Prove your answers.

iv) Suppose that $A \in M_{m,n}(F)$ where $F$ is a field, and $L = L_A : F^m \to F^m$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $F^n$, and $\{f_1, \ldots, f_m\}$ be the standard basis of $F^m$. Let $\beta = \{e_{\beta_1}, \ldots, e_{\beta_m}\}$ be the dual basis of $(F^n)^*$ and let $\beta' = \{f_{\beta_1}, \ldots, f_{\beta_n}\}$ be the dual basis of $(F^m)^*$. Compute $M_{\beta'}(\hat{L})$.

3. Let $V$ be a vector space of dimension $n$, and $W$ be a subspace of $V$. Define

$$\text{Perp}_{V^*}(W) = \{ \varphi \in V^* \mid \varphi(w) = 0 \text{ for all } w \in W \}.$$ 

i) Prove that $\dim W + \dim \text{Perp}_{V^*}(W) = n$.

ii) Suppose that $V$ has a nondegenerate symmetric bilinear form $<,>$. Prove that the map $v \mapsto L_v$ of Problem 1.ii) induces an isomorphism of

$$W^\perp = \{ v \in V \mid <v, w> = 0 \text{ for all } w \in W \}$$

with $\text{Perp}_{V^*}(W)$. Conclude that $\dim W + \dim W^\perp = n$.

iii) Use the conclusions of Problem 3.ii) to prove that the column rank of a matrix is equal to the row rank of a matrix (see Definition 7.4). This gives a different proof of part of Theorem 7.5.

iv) (rank-nullity theorem) Suppose that $A \in M_{m,n}(F)$ is a matrix. Show that the dimension of the solution space

$$\{ x \in F^n \mid Ax = \vec{0} \}$$
is equal to $n - \text{rank } A$. Thus with the notation of Lemma 5.3,
$$\dim U^\perp + \text{rank } A = n.$$  

v) Consider the complex vector space $\mathbb{C}^n$ with the dot product as nondegenerate symmetric bilinear form. Give an example of a subspace $W$ of $\mathbb{C}^n$ such that $\mathbb{C}^n \neq W \oplus W^\perp$. Verify for your example that $\dim W + \dim W^\perp = \dim \mathbb{C}^2 = 2$.

4. Suppose that $V$ is a finite dimensional vector space with a nondegenerate symmetric bilinear form $< , >$. An operator $\Phi$ on $V$ is a linear map $\Phi : V \to V$. If $\Phi$ is an operator on $V$, show that there exists a unique operator $\Psi : V \to V$ such that $< \Phi(v), w > = < v, \Psi(w) >$ for all $v, w \in V$. $\Psi$ is called the transpose of $\Phi$, and we write $\Psi = \Phi^t$.

Hint: Define a map $\Psi : V \to V$ as follows. For $w \in V$, let $L : V \to F$ be the map defined by $L(v) = < \Phi(v), w >$ for $v \in V$. Verify that $L$ is linear, so that $L \in V^*$. By Problem 1.ii), there exists a unique $w' \in V$ such that for all $v \in V$, we have $L(v) = < v, w' >$. Define $\Psi(w) = w'$. Prove that $\Psi$ is linear, and that $< \Phi(v), w > = < v, \Psi(w) >$ for all $v, w \in V$. Don’t forget to prove that $\Psi$ is unique!

5. Let $V = F^n$ with the dot product as nondegenerate symmetric bilinear form. Suppose that $L : V \to V$ is an operator. Show that $L^t = L_A$ if $A \in M_{n \times n}(F)$ is the $n \times n$ matrix such that $L = L_A$.

6. Give a “down to earth” (but “coordinate dependent”) proof of Problem 4, starting like this. Let $\beta = \{ v_1, \ldots, v_n \}$ be a basis of $V$. The linear map $\Phi_\beta : V \to F^n$ defined by $\Phi_\beta(v) = (v)\beta$ is an isomorphism, with inverse $\Psi : F^n \to V$ defined by $\Psi([x_1, \ldots, x_n]^t) = x_1v_1 + \cdots + x_nv_n$. Define a product $[ , ]$ on $F^n$ by $[x, y] = < \Psi(x), \Psi(y) >$. Verify that $[ , ]$ is a nondegenerate symmetric bilinear form on $F^n$.

By our classification of nondegenerate symmetric bilinear forms on $F^n$, we know that $[x, y] = x^tBy$ for some symmetric invertible matrix $B \in M_{n \times n}(F)$.

7. Determinants

**Theorem 7.1.** Suppose that $F$ is a field, and $n$ is a positive integer. Then there exists a unique function $\text{Det} : M_{n,n}(F) \to F$ which satisfies the following properties:

1. $\text{Det}$ is multilinear in the columns of matrices in $M_{n,n}(F)$.
2. $\text{Det}$ is alternating in the columns of matrices in $M_{n,n}(F)$; that is $\text{Det}(A) = 0$ whenever two columns of $A$ are equal.
3. $\text{Det}(I_n) = 1$.

Existence can be proven by defining a function on $A = (a_{ij}) \in M_{n,n}(F)$ by
$$\text{Det}(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma)a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}$$
where for a permutation $\sigma \in S_n$,
$$\text{sgn}(\sigma) = \begin{cases} 
1 & \text{if } \sigma \text{ is even} \\
-1 & \text{if } \sigma \text{ is odd}.
\end{cases}$$

It can be verified that $\text{Det}$ is multilinear, alternating and $\text{Det}(I_n) = 1$.

Uniqueness is proven by showing that if a function $\Phi : M_{n,n}(F) \to F$ is multilinear and alternating on columns and satisfies $\Phi(I_n) = 1$ then $\Phi = \text{Det}$.

**Lemma 7.2.** Suppose that $A, B \in M_{n,n}(F)$. Then
1. \( \text{Det}(AB) = \text{Det}(A)\text{Det}(B) \).
2. \( \text{Det}(A^i) = \text{Det}(A) \).

**Lemma 7.3.** Suppose that \( A \in M_{n,n}(F) \) Then the following are equivalent

1. \( \text{Det}(A) \neq 0 \)
2. The columns \( \{A^1, \ldots, A^n\} \) of \( A \) are linearly independent.
3. The solution space to \( Ax = 0 \) is \((0)\).

**Proof.** 2. is equivalent to 3. by the dimension formula, applied to the linear map \( L_A : F^n \to F^n \).

We will now establish that 1. is equivalent to 2. Suppose that \( \{A^1, \ldots, A^n\} \) are linearly dependent. Then, after possibly permuting the columns of \( A \) such that
\[
A_1: \cdots : A_{s-1}, A_s, \cdots, A_n^m
\]
\( A \) has rank \( s \). Now suppose that \( \{A^1, \ldots, A^n\} \) are independent. Then, after permuting the columns of \( A \).

\[
\text{Det}(A) = \text{Det}(A_1, \ldots, A_n^	ext{m-1}, A_m)
\]
\[
= \text{Det}(A_1, \ldots, A_n^	ext{m-1}, c_1A_1 + \cdots + c_m A_n - 1)
\]
\[
= c_1\text{Det}(A_1, \ldots, A_n^	ext{m-1}, A_1') + \cdots + c_m\text{Det}(A_1, \ldots, A_n^	ext{m-1}, A_n')
\]
\[
= 0.
\]

Now suppose that \( \{A^1, \ldots, A^n\} \) are linearly independent. Then \( \text{dim Image} L_A = n \). By the dimension theorem, \( \text{dim Kernel} L_A = 0 \). Thus \( L_A \) is an isomorphism, so \( A \) is invertible with an inverse \( B \). We have that \( 1 = \text{Det}(I_n) = \text{Det}(AB) = \text{Det}(A)\text{Det}(B) \). Thus \( \text{Det}(A) \neq 0 \). \( \square \)

**Definition 7.4.** Suppose that \( A \in M_{m,n}(F) \).

1. The column rank of \( A \) is the dimension of the column space \( \text{Span}\{A^1, \ldots, A^n\} \), which is a subspace of \( F^m \).
2. The row rank of \( A \) is the dimension of the row space \( \text{Span}\{A_1, \ldots, A_m\} \), which is a subspace of \( F_n \).
3. The rank of \( A \) is the largest integer \( r \) such that there exists an \( r \times r \) submatrix \( B \) of \( A \) such that \( \text{Det}(B) \neq 0 \).

**Theorem 7.5.** Suppose that \( A \in M_{m,n}(F) \). Then the column rank of \( A \), the row rank of \( A \) and the rank of \( A \) are all equal.

**Proof.** Let \( s \) be the column rank of \( A \), and let \( r \) be the rank of \( A \). We will show that \( s = r \).

There exists an \( r \times r \) submatrix \( B \) of \( A \) such that \( \text{Det}(B) \neq 0 \). There exist \( 1 \leq i_1 < i_2 < \cdots < i_r \leq m \) and \( 1 \leq j_1 < j_2 < \cdots < j_r \leq n \) such that \( B \) is obtained from \( A \) by deleting the rows \( A_{i_1}, \ldots, A_{i_r} \) and the columns \( A_{j_1}, \ldots, A_{j_r} \) such that \( j \notin \{j_1, \ldots, j_r\} \). We have that the columns of \( B \) are linearly independent, so the columns \( A_{i_1}^1, \ldots, A_{i_r}^n \) of \( A \) are linearly independent. Thus \( s \geq r \).

We now will prove that \( r \geq s \) by induction on \( n \), from which it follows that \( r = s \). The case \( n = 1 \) is easily verified; \( r \) and \( s \) are either 0 or 1, depending on if \( A \) is zero or not.

Assume that the induction statement is true for matrices of size less than \( n \). After permuting the columns of \( A \), we may assume that the first \( s \) columns of \( A \) are linearly independent. Since the column space is a subspace of \( F^m \), we have that \( s \leq m \). Let \( B \) be the submatrix consisting of the first \( s \) rows of \( A \). If \( s < n \), then by induction we have that the rank of \( B \) is \( s \), so that \( A \) has rank \( r \geq s \).

We have reduced to the case where \( s = n \leq m \). Let \( C \) be the \((n-1) \times m \) submatrix of \( A \) consisting of the first \( n-1 \) columns of \( A \). By induction, there exists an \((n-1) \times (n-1) \) submatrix \( E \) of \( C \) whose determinant is non zero. After possibly interchanging rows of \( C \), we may assume that \( E \) consists of the first \( n-1 \) rows of \( C \). For \( 1 \leq i \leq n \), let \( E_i \) be the
column vector consisting of the first \( n - 1 \) rows of the \( i \)-th column of \( A \), and for \( 1 \leq i \leq n \) and \( n \leq j \leq m \), let \( E_i^j \) be the column vector consisting of the first \( n - 1 \) rows of the \( i \)-th column of \( A \), followed by the \( j \)-th row of the \( i \)-th column of \( A \).

Since \( \text{Det}(E) \neq 0 \), \( \{E_1, \ldots, E_{n-1}\} \) are linearly independent, and are thus a basis of \( F^{n-1} \). Thus there are unique elements \( a_1, \ldots, a_{n-1} \in F \) such that

\[
(3) \quad a_1E_1 + \cdots + a_{n-1}E_{n-1} = E_n.
\]

Suppose that all determinants of \( n \times n \) submatrices of \( A \) are zero. Then \( \{E_1^1, \ldots, E_n^n\} \) are linearly dependent for \( n \leq j \leq m \). By uniqueness of the relation (3), we must then have that

\[
a_1E_1^1 + \cdots + a_{n-1}E_{n-1}^j - E_n = 0
\]

for \( n \leq j \leq m \). Thus

\[
a_1A^1 + \cdots + a_{n-1}A^{n-1} - A^n = 0,
\]

a contradiction to our assumption that \( A \) has column rank \( n \). Thus some \( n \times n \) submatrix of \( A \) has nonzero determinant, and thus \( r \geq s \).

Taking the transpose of \( A \), the above argument shows that the row rank of \( A \) is equal to the rank of \( A \).

\[\square\]

8. Rings

A set \( R \) is a ring if it has an addition operation under which \( A \) is an abelian group, and an associative multiplication operation such that \( a(b + c) = ab + ac \) and \( (b + c)a = ba + ca \) for all \( a, b, c \in A \). \( A \) further has a multiplicative identity \( 1_A \) (written 1 when there is no danger of confusion). A ring \( A \) is a commutative ring if \( ab = ba \) for all \( a, b \in A \).

Suppose that \( A \) and \( B \) are rings. A ring homomorphism \( \varphi : A \to B \) is a mapping such that \( \varphi \) is a homomorphism of abelian groups, \( \varphi(ab) = \varphi(a)\varphi(b) \) for \( a, b \in A \) and \( \varphi(1_A) = 1_B \).

Suppose that \( B \) is a commutative ring, \( A \) is a subring of \( B \) and \( S \) is a subset of \( B \). The subring of \( B \) generated by \( S \) and \( A \) is the intersection of all subrings \( T \) of \( B \) which contain \( A \) and \( S \). The subring of \( B \) generated by \( S \) and \( A \) is denoted by \( A[S] \). It should be verified that \( A[S] \) is in fact a subring of \( B \), and that

\[
A[S] = \left\{ \sum_{i_1, i_2, \ldots, i_r=0}^{n} a_{i_1, \ldots, i_r} s_1^{i_1} \cdots s_r^{i_r} \mid r \in \mathbb{N}, n \in \mathbb{N}, a_{i_1, \ldots, i_r} \in A, s_1, \ldots, s_r \in S \right\}.
\]

and let \( R_i = R \) for \( i \in \mathbb{N} \). Let

\[
T = \{\{a_i\} \in \times_{i \in \mathbb{N}} R_i \mid a_i = 0 \text{ for } i \gg 0\}.
\]

We can also write a sequence \( \{a_i\} \in T \) as \( (a_0, a_1, a_2, \ldots, a_r, 0, 0, \ldots) \) for some \( r \in \mathbb{N} \). \( T \) is a ring with addition \( \{a_i\} + \{b_i\} = \{a_i + b_i\} \) and \( \{a_i\}\{b_i\} = \{c_k\} \) where \( c_k = \sum_{i+j=k} a_i b_j \).

The zero element \( 0_T \) is the sequence all of whose terms are 0, and \( 1_T \) is the sequence all of whose terms are zero expect the zero’th term which is 1. That is, \( 0_T = (0, 0, \ldots) \) and \( 1_T = (1, 0, 0, \ldots) \). Let \( x \) be the sequence whose first term is 1 and all other terms are zero; that is, \( x = (0, 1, 0, 0, \ldots) \). Then \( x^t \) is the sequence whose \( i \)-th term is 1 and all other terms are zero. The natural map of \( R \) into \( T \) which maps \( a \in R \) to the sequence whose zero-th term is \( a \) and all other terms are zero identifies \( R \) with a subring of \( T \). Suppose
that \( \{a_i\} \in T \). Then there is some natural number \( r \) such that \( a_i = 0 \) for all \( i > r \). We have that
\[
\{a_i\} = (a_0, a_1, \ldots, a_r, 0, 0, \ldots) = \sum_{i=0}^{r} a_i x^i.
\]
Thus the subring \( R[x] \) of \( T \) generated by \( x \) and \( R \) is equal to \( T \). \( R[x] \) is called a polynomial ring.

**Lemma 8.1.** Suppose that \( f(x), g(x) \in R[x] \) have expansions \( f(x) = a_0 + a_1 x + \cdots + a_r x^r \in R[x] \) and \( g(x) = b_0 + b_1 x + \cdots + b_r x^r \in R[x] \) with \( a_0, \ldots, a_r \in R \) and \( b_0, \ldots, b_r \in R \). Suppose \( f(x) = g(x) \). Then \( a_i = b_i \) for \( 0 \leq i \leq r \).

**Proof.** We have \( f(x) = (a_0, a_1, \ldots, a_r, 0, \ldots) \) and \( g(x) = (b_0, b_1, \ldots, b_r, 0, \ldots) \) as elements of \( \times_{i \in \mathbb{N}} R_i \). Since two sequences in \( \times_{i \in \mathbb{N}} R_i \) are equal if and only if their coefficients are equal, we have that \( a_i = b_i \) for all \( i \). \( \square \)

**Theorem 8.2.** (Universal Property of Polynomial Rings) Suppose that \( A \) and \( B \) are commutative rings and \( \varphi : A \to B \) is a ring homomorphism. Suppose that \( b \in B \). Then there exists a unique ring homomorphism \( \overline{\varphi} : A[x] \to B \) such that \( \overline{\varphi}(a) = \varphi(a) \) for \( a \in A \) and \( \overline{\varphi}(x) = b \).

**Proof.** We first prove existence. Define \( \overline{\varphi} : A \to B \) by \( \overline{\varphi}(f(x)) = \varphi(a_0) + \varphi(a_1)b + \cdots + \varphi(a_r) b^r \) for
\[
f(x) = a_0 + a_1 x + \cdots + a_r x^r \in R[x]
\]
with \( a_0, \ldots, a_r \in R \). This map is well defined, since the expansion (4) is unique by Lemma 8.1. It should be verified that \( \overline{\varphi} \) is a ring homomorphism with the desired properties.

We will now prove uniqueness. Now suppose that \( \Psi : A[x] \to B \) is a ring homomorphism such that \( \Psi(a) = \varphi(a) \) for \( a \in A \) and \( \Psi(x) = b \). Suppose that \( f(x) \in A[x] \). Write \( f(x) = a_0 + a_1 x + \cdots + a_r x^r \) with \( a_0, \ldots, a_r \in A \). We have
\[
\Psi(f(x)) = \Psi(a_0 + a_1 x + \cdots + a_r x^r) = \Psi(a_0) + \Psi(a_1) \Psi(x) + \cdots + \Psi(a_r) \Psi(x)^r
\]
\[
= \varphi(a_0) + \varphi(a_1) b + \cdots + \varphi(a_r) b^r = \overline{\varphi}(a_0 + a_1 x + \cdots + a_r x^r)
\]
\[
= \overline{\varphi}(f(x)).
\]
Since this identity is true for all \( f(x) \in A[x] \), we have that \( \Psi = \overline{\varphi} \). \( \square \)

Let \( F \) be a field, and \( F[x] \) be a polynomial ring over \( F \). Suppose that \( f(x) = a_0 + a_1 x + \cdots + a_n x^n \in F[x] \) with \( a_0, \ldots, a_n \in F \) and \( a_n \neq 0 \). The degree of \( f(x) \) is defined to be \( \deg f(x) = n \). If \( f(x) = 0 \) define \( \deg f(x) = -\infty \). For \( f, g \in F[x] \) we have
\[
\deg fg = \deg f + \deg g
\]
and
\[
\deg f + g \leq \max \{ \deg f, \deg g \}.
\]
f \in F[x] is a unit (has a multiplicative inverse in \( F[x] \)) if and only if \( f \) is a nonzero element of \( F \), if and only if \( \deg f = 0 \) (this follows from the formulas on degree). \( F[x] \) is a domain; that is it has no zero divisors. A basic result is Euclidean division.

**Theorem 8.3.** Suppose that \( f, g \in F[x] \) with \( f \neq 0 \). Then there exist unique polynomials \( q, r \in F[x] \) such that \( g = qf + r \) with \( \deg r < \deg f \).

**Corollary 8.4.** Suppose that \( f(x) \in F[x] \) is nonzero. Then \( f(x) \) has at most \( \deg f \) distinct roots in \( F \).
Corollary 8.5. \( F[x] \) is a principal ideal domain; that is every ideal in \( F[x] \) is generated by a single element.

An element \( f(x) \in F[x] \) is defined to be irreducible if it has degree \( \geq 1 \), and \( f \) cannot be written as a product \( f = gh \) with \( f, g \notin F \). A polynomial \( f(x) \in F[x] \) of positive degree \( n \) is monic if \( f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n \) for some \( a_0, \ldots, a_{n-1} \in F \).

Theorem 8.6. (Unique Factorization) Suppose that \( f(x) \in F[x] \) has positive degree. Then there is a factorization

\[
f(x) = cf_1^{n_1} \cdots f_r^{n_r}
\]

where \( r \) is a positive integer, \( f_1, \ldots, f_r \in F[x] \) are monic irreducible, \( 0 \neq c \in F \) and \( n_1, \ldots, n_r \) are positive integers.

This factorization is unique, in the sense that any other factorization of \( f(x) \) as a product of monic irreducibles is obtained from (6) by permuting the \( f_i \).

If \( F \) is an algebraically closed field (such as the complex numbers \( \mathbb{C} \)) the monic irreducibles are exactly the linear polynomials \( x - \alpha \) for \( \alpha \in F \).

Theorem 8.7. Suppose that \( F \) is an algebraically closed field and \( f(x) \in F[x] \) has positive degree. Then there is a (unique) factorization

\[
f(x) = c(x - \alpha_1)^{n_1} \cdots (x - \alpha_r)^{n_r}
\]

where \( r \) is a positive integer, \( \alpha_1, \ldots, \alpha_r \in F \) are distinct, \( 0 \neq c \in F \) and \( n_1, \ldots, n_r \) are positive integers.

Suppose that \( f(x), g(x) \in F[x] \) are not both zero. A common divisor of \( f \) and \( g \) in \( F[x] \) is an element \( h \in F[x] \) such that \( h \) divides \( f \) and \( h \) divides \( g \) in \( F[x] \). A greatest common divisor of \( f \) and \( g \) in \( F[x] \) is an element \( d \in F[x] \) such that \( d \) is a common divisor of \( f \) and \( g \) in \( F[x] \) and \( d \) divides every other common divisor of \( f \) and \( g \) in \( F[x] \).

Theorem 8.8. Suppose that \( f, g \in F[x] \) are not both zero. Then

1. there exists a unique monic greatest common divisor \( d(x) = \gcd(f, g) \) of \( f \) and \( g \) in \( F[x] \).
2. There exist \( p, q \in F[x] \) such that \( d = pf + qg \).
3. Suppose that \( K \) is a field containing \( F \). Then \( d \) is a greatest common divisor of \( f \) and \( g \) in \( K[x] \).

9. Eigenvalues and Eigenvectors

Suppose that \( F \) is a field.

Definition 9.1. Suppose that \( V \) is a vector space over a field \( F \) and \( L : V \rightarrow V \) is a linear map. An element \( \lambda \in F \) is an eigenvalue of \( L \) if there exists a nonzero element \( v \in V \) such that \( L(v) = \lambda v \). A nonzero vector \( v \in V \) such that \( L(v) = \lambda v \) is called an eigenvector of \( L \) with eigenvalue \( \lambda \). If \( \lambda \in F \) is an eigenvalue of \( L \), then

\[
E(\lambda) = \{ v \in V \mid L(v) = \lambda v \}
\]

is the eigenspace of \( \lambda \) for \( L \).

Lemma 9.2. Suppose that \( V \) is a vector space over a field \( F \), \( L : V \rightarrow V \) is a linear map, and \( \lambda \in F \) is an eigenvalue of \( L \). Then \( E(\lambda) \) is a subspace of \( V \) of positive dimension.
Theorem 9.3. Suppose that $V$ is a vector space over a field $F$, and $L : V \to V$ is a linear map. Let $\lambda_1, \ldots, \lambda_r \in F$ be distinct eigenvalues of $L$. Then the subspace of $V$ spanned by $E(\lambda_1), \ldots, E(\lambda_r)$ is a direct sum; that is

$$E(\lambda_1) + E(\lambda_2) + \cdots + E(\lambda_r) = E(\lambda_1) \bigoplus E(\lambda_2) \bigoplus \cdots \bigoplus E(\lambda_r).$$

Proof. We prove the Lemma by induction on $r$. The case $r = 1$ follows from the definition of direct sum. Now assume that the Lemma is true for $r - 1$ eigenvalues. Suppose that $v_i \in E(\lambda_i)$ for $1 \leq i \leq r$, $w_i \in E(\lambda_i)$ for $1 \leq i \leq r$ and

$$v_1 + v_2 + \cdots + v_r = w_1 + w_2 + \cdots + w_r. \quad (7)$$

We must show that $v_i = w_i$ for $1 \leq i \leq r$. Multiplying equation (7) by $\lambda_i$, we obtain

$$\lambda_i v_1 + \lambda_i v_2 + \cdots + \lambda_i v_r = \lambda_i w_1 + \lambda_i w_2 + \cdots + \lambda_i w_r. \quad (8)$$

Applying $L$ to equation (7) we obtain

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_r w_r. \quad (9)$$

Subtracting equation (8) from (9), we obtain

$$(\lambda_1 - \lambda_r)v_1 + \cdots + (\lambda_{r-1} - \lambda_r)v_{r-1} = (\lambda_r - \lambda_r)w_1 + \cdots + (\lambda_{r-1} - \lambda_r)w_{r-1}. \quad (10)$$

Now by induction on $r$, and since $\lambda_i - \lambda_r \neq 0$ for $1 \leq i \leq r - 1$, we get that $v_i = w_i$ for $1 \leq i \leq r - 1$. Substituting into equation (7), we then obtain that $v_r = w_r$, and we see that the sum is direct. \qed

Corollary 9.4. Suppose that $V$ is a finite dimensional vector space over a field $F$, and $L : V \to V$ is linear. Then $L$ has at most $\dim(V)$ distinct eigenvalues.

10. Characteristic Polynomials

Suppose that $F$ is a field. We let $F[t]$ be a polynomial ring over $F$.

Definition 10.1. Suppose that $F$ is a field, and $A \in M_{n,n}(F)$. The characteristic polynomial $A$ is the polynomial $p_A(t) = \det(tI_n - A) \in F[t]$.

$p_A(t)$ is a monic polynomial of degree $n$.

Lemma 10.2. Suppose that $V$ is an $n$-dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Let $\beta, \beta'$ be two bases of $V$. Then

$$p_{M_\beta^\beta(L)}(t) = p_{M_\beta^{\beta'}(L)}(t).$$

Proof. We have that

$$M_\beta^{\beta'}(L) = M_\beta^\beta(I)M_\beta^\beta(L)M_\beta^{\beta'}(I).$$

Let $A = M_\beta^\beta(L)$, $B = M_\beta^{\beta'}(L)$ and $C = M_\beta^{\beta'}(I)$, so that $B = C^{-1}AC$. Then

$$p_{M_\beta^{\beta'}(L)}(t) = \det(tI_n - B) = \det(tI_n - C^{-1}AC) = \det(C^{-1}(tI_n - A)C) = \det(C)^{-1}\det(tI_n - A)\det(C) = \det(tI_n - A) = p_{M_\beta^\beta(L)}(t).$$

\qed
Definition 10.3. Suppose that $V$ is an $n$-dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Let $\beta$ be a basis of $V$. Define the characteristic polynomial $p_L(t)$ of $L$ to be

$$p_L(t) = p_{M_\beta(L)}(t).$$

Lemma 10.2 shows that the definition of characteristic polynomial of $L$ is well defined; it is independent of choice of basis of $V$.

Theorem 10.4. Suppose that $V$ is an $n$-dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Then the eigenvalues of $L$ are the roots of $p_L(t) = 0$ which are in $F$. In particular, $L$ has at most $n$ distinct eigenvalues.

Proof. Let $\beta$ be a basis of $V$, and suppose that $\lambda \in F$. Then the following are equivalent:

1. $\lambda$ is an eigenvalue of $L$.
2. There exists a nonzero vector $v \in V$ such that $Lv = \lambda v$.
3. $\text{the Kernel of } \lambda I - L : V \to V$ is nonzero.
4. The solution space to $M_\beta(\lambda I - L)y = 0$ is non zero.
5. The solution space to $(\lambda I_n - M_\beta(L))y = 0$ is non zero.
6. $\text{Det}(\lambda I_n - M_\beta(L)) = 0$
7. $p_L(\lambda) = 0$.

$\square$

Definition 10.5. Suppose that $V$ is an $n$-dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. $L$ is diagonalizable if there exists a basis $\beta$ of $V$ such that $M_\beta(L)$ is a diagonal matrix.

Theorem 10.6. Suppose that $V$ is an $n$-dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Then the following are equivalent:

1. $L$ is diagonalizable.
2. There exists a basis of $V$ consisting of eigenvectors of $L$.
3. Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of $L$. Then $\dim E(\lambda_1) + \cdots + \dim E(\lambda_r) = n$.

Lemma 10.7. Suppose that $A \in M_{n,n}(F)$ is a matrix. Then $L_A$ is diagonalizable if and only if $A$ is similar to a diagonal matrix (there exists an invertible matrix $B \in M_{n,n}(F)$ and a diagonal matrix $D \in M_{n,n}(F)$ such that $B^{-1}AB = D$).

11. Triangulation

Let $V$ be a finite dimensional vector space over a field $F$, of positive dimension $n \geq 1$. Suppose that $L : V \to V$ is a linear map. A subspace $W$ of $V$ is $L$-invariant if $L(W) \subseteq W$.

A fan $\{V_1, \ldots, V_n\}$ of $L$ in $V$ is a finite sequence of $L$-invariant subspaces

$$V_1 \subset V_2 \subset \cdots \subset V_n = V$$

such that $\dim V_i = i$ for $1 \leq i \leq n$.

A fan basis (of the fan $\{V_1, \ldots, V_n\}$) for $L$ is a basis $\{v_1, \ldots, v_n\}$ of $V$ such that $\{v_1, \ldots, v_i\}$ is a basis of $V_i$ for $1 \leq i \leq n$. Given a fan, a fan basis always exists.

Theorem 11.1. Let $\beta = \{v_1, \ldots, v_n\}$ be a fan basis for $L$. Then the matrix $M_\beta(L)$ of $L$ with respect to the basis $\beta$ is an upper triangular matrix.
Proof. Let \( \{ V_1, \ldots, V_n \} \) be the fan corresponding to the fan basis. \( L(V_i) \subset V_i \) for all \( i \) implies there exist \( a_{ij} \in F \) such that

\[
\begin{align*}
L(v_1) & = a_{11}v_1 \\
L(v_2) & = a_{21}v_1 + a_{22}v_2 \\
& \vdots \\
L(v_n) & = a_{n1}v_1 + \cdots + a_{nn}v_n.
\end{align*}
\]

Thus

\[
M_\beta^L(L) = (L(v_1)_\beta, L(v_2)_\beta, \ldots, L(v_n)_\beta) = (a_{ij})
\]

is upper triangular. \( \square \)

**Definition 11.2.** A linear map \( L : V \to V \) is triangular if there exists a basis \( \beta \) of \( V \) such that \( M_\beta^L(L) \) is upper triangular. A matrix \( A \in M_{nn}(F) \) is triangulizable if there exists an invertible matrix \( B \in M_{nn}(F) \) such that \( B^{-1}AB \) is upper triangular.

A matrix \( A \) is triangulizable if and only if the linear map \( L_A : F^n \to F^n \) is triangulizable; if \( \beta \) is a fan basis for \( L_A \), then

\[
M_\beta^L(L_A) = M_{st}^I(M_{st}^L(L_A))M_\beta^I = B^{-1}AB
\]

where \( I : V \to V \) is the identity map and \( st \) denotes the standard basis of \( F^n \), and \( B = M_{st}^I(I) \).

Suppose that \( V \) is a vector space, and \( W_1, W_2 \) are subspaces such that \( V = W_1 \bigoplus W_2 \). Then every element \( v \in V \) has a unique expression \( v = w_1 + w_2 \) with \( w_1 \in W_1 \) and \( w_2 \in W_2 \). We may thus define a projection \( \pi_1 : V \to W_1 \) by \( \pi_1(v) = w_1 \) if \( v = w_1 + w_2 \) with \( w_1 \in W_1 \) and \( w_2 \in W_2 \), and define a projection \( \pi_2 : V \to W_1 \) by \( \pi_1(v) = w_2 \) for \( v = w_1 + w_2 \in V \). These projections are linear maps, which depend on both \( W_1 \) and \( W_2 \). Composing with the inclusion \( W_1 \subset V \), we can view \( \pi_1 \) as a map from \( V \) to \( V \). Composing with the inclusion \( W_2 \subset V \), we can view \( \pi_2 \) as a map from \( V \) to \( V \). Then \( \pi_1 + \pi_2 = I_V \).

**Theorem 11.3.** Let \( V \) be a non zero finite dimensional vector space over a field \( F \), and let \( L : V \to V \) be a linear map. Suppose that the characteristic polynomial \( p_L(t) \) factors into linear factors in \( F[t] \) (this will always happen if \( F \) is algebraically closed). Then there exists a fan of \( L \) in \( V \), so that \( L \) is triangulizable.

**Proof.** We prove the theorem by induction on \( n = \dim V \). If \( n = 1 \), then any basis of \( V \) is a fan basis for \( V \). Assume that the theorem is true for linear maps of vector spaces \( W \) over \( F \) of dimension less than \( n = \dim V \). Since \( p_L(t) \) splits into linear factors in \( F[t] \), and \( p_L(t) \) has degree \( n > 0 \), \( p_L(t) \) has a root \( \lambda_1 \) in \( F \), which is thus an eigenvalue of \( L \). Let \( v_1 \in V \) be an eigenvector of \( L \) with the eigenvalue \( \lambda_1 \). Let \( V_1 \) be the one dimensional subspace of \( V \) generated by \( \{ v_1 \} \). Extend \( \{ v_1 \} \) to a basis \( \beta = \{ v_1, v_2, \ldots, v_n \} \) of \( V \). Let \( W = \text{Span}\{ v_2, \ldots, v_n \} \). \( \beta' = \{ v_2, \ldots, v_n \} \) is a basis of \( W \). \( V \) is the direct sum \( V = V_1 \bigoplus W \). Let \( \pi_1 : V \to V_1 \) and \( \pi_2 : V \to W \) be the projections. The composed linear map is \( \pi_2L : W \to W \). From

\[
M_\beta^L(L) = \begin{pmatrix}
\lambda_1 & * \\
0 & M_{\beta'}^L(\pi_2L)
\end{pmatrix},
\]

we calculate \( p_L(t) = (t - \lambda_1)p_{\pi_2L}(t) \). Since \( p_L(t) \) factors into linear factors in \( F[t] \), \( p_{\pi_2L}(t) \) also factors into linear factors in \( F[t] \).
By induction, there exists a fan of $\pi_2 L$ in $W$, say $\{W_1, \ldots, W_{n-1}\}$. Let $V_i = V_1 + W_{i-1}$ for $2 \leq i \leq n$. The subspaces $V_i$ form a chain

$$V_1 \subset V_2 \subset \cdots \subset V_n = V.$$  

Let $\{u_1, \ldots, u_{n-1}\}$ be a fan basis for $\pi_2 L$, so that $\{u_1, \ldots, u_j\}$ is a basis of $W_j$ for $1 \leq j \leq n-1$. Then $\{v_1, u_1, \ldots, u_{n-1}\}$ is a basis of $V_i$. We will show that $\{V_1, \ldots, V_n\}$ is a fan of $L$ in $V$, with fan basis $\{v_1, u_1, \ldots, u_{n-1}\}$.

We have that $L = IL = (\pi_1 + \pi_2) L = \pi_1 L + \pi_2 L$. Let $v \in V_i$. We have an expression $v = cv_1 + w_{i-1}$ with $c \in F$ and $w_{i-1} \in W_{i-1}$. $\pi_1 L(v) = \pi_1 (L(v)) \in V_1 \subset V_i$. $\pi_2 L(v) = \pi_2 L(cv_1) + \pi_2 L(w_{i-1})$. Now $\pi_2 L(cv_1) = c\pi_2 (v_1) = c\pi_2 (\lambda_1 v_1) = 0$ and $\pi_2 L(w_{i-1}) \in W_{i-1}$ since $\{w_1, \ldots, w_{n-1}\}$ is a fan basis for $\pi_2 L$. Thus $\pi_2 L(v) \in W_{i-1}$ and so $L(v) \in V_i$. Thus $V_i$ is $L$-invariant, and we have shown that $\{V_1, \ldots, V_n\}$ is a fan of $L$ in $V$. 

12. The minimal polynomial of a linear operator

Suppose that $L : V \rightarrow V$ is a linear map, where $V$ is a finite dimensional vector space over a field $F$. Let $\mathcal{L}_F(V, V)$ be the $F$-vector space of linear maps from $V$ to $V$. $\mathcal{L}_F(V, V)$ is a ring with multiplication given by composition of maps. Let $I : V \rightarrow V$ be the identity map, which is the multiplicative identity of the ring $\mathcal{L}_F(V, V)$. There is a natural inclusion of the field $F$ as a subring (and subvector space) of $\mathcal{L}_F(V, V)$, obtained by mapping $\lambda \in F$ to $\lambda I \in \mathcal{L}_F(V, V)$. Let $F[L]$ be the subring of $\mathcal{L}_F(V, V)$ generated by $F$ and $L$.

$$F[L] = \{a_0 I + a_1 L + \cdots + a_r L^r \mid r \in \mathbb{N} \text{ and } a_0, \ldots, a_r \in F\}.$$ 

$F[L]$ is a commutative ring. By the universal property of polynomial rings, there is a surjective ring homomorphism $\varphi$ from the polynomial ring $F[t]$ onto $F[L]$, obtained by mapping $f(t) \in F[t]$ to $f(L)$. $\varphi$ is also a vector space homomorphism. Since $F[L]$ is a subspace of $\mathcal{L}_F(V, V)$, which is an $F$-vector space of dimension $n^2$, $F[L]$ is a finite dimensional vector space. Thus the kernel of $\varphi$ is nonzero. Let $m_L(t)$ be the monic generator of the kernel of $\varphi$. We have an isomorphism

$$F[L] \cong F[t]/(m_L(t)).$$ 

$m_L(t)$ is the minimal polynomial of $L$.

We point out here that if $g(t), h(t) \in F[t]$ are polynomials, then since $F[L]$ is a commutative ring, we have equality of composition of operators

$$g(L) \circ h(L) = h(L) \circ g(L).$$

We generally write $g(L)h(L) = h(L)g(L)$ to denote the composition of operators. If $v \in V$, we will also write $Lv$ for $L(v)$ and $f(L)v$ for $f(L)(v)$.

The above theory also works for matrices. The $n \times n$ matrices $M_{n,n}(F)$ is a vector space over $F$, and is also a ring with matrix multiplication. $F$ is embedded as a subring by the map $\lambda \rightarrow \lambda I_n$, where $I_n$ denotes the $n \times n$ identity matrix. Given $A \in M_{n,n}(F)$, $F[A]$ is a commutative subring, and there is a natural surjective ring homomorphism $F[t] \rightarrow F[A]$. The kernel is generated by a monic polynomial $m_A(t)$ which is the minimal polynomial of $A$.

Now suppose that $V$ is a finite dimensional vector space over $F$, and $\beta$ is a basis of $V$. Then the map $\mathcal{L}_F(V, V) \rightarrow M_{n,n}(F)$ defined by $\varphi \rightarrow M^{\beta}_{\beta}(\varphi)$ for $\varphi \in \mathcal{L}_F(V, V)$ is a vector space isomorphism, and a ring isomorphism.

Suppose that $L : V \rightarrow V$ is a linear map, and let $A = M^{\beta}_{\beta}(L)$. Then we have an induced isomorphism of $F[L]$ with $F[A]$. For a polynomial $f(x) \in F[x]$, we have that the image of
\( f(L) \) in \( F[A] \) is 

\[
M_\beta^2(f(L)) = f(M_\beta^2(L)) = f(A).
\]

Thus we have that \( m_A(t) = m_L(t) \).

**Lemma 12.1.** Suppose that \( A \) is an \( n \times n \) matrix with coefficients in a field \( F \), and \( K \) is an extension field of \( F \). Let \( f(x) \) be the minimal polynomial of \( A \) over \( F \), and let \( g(x) \) be the minimal polynomial of \( A \) over \( K \). Then \( f(x) = g(x) \).

**Proof.** Let \( a_1, a_2, \ldots, a_r \in K \) be elements of \( K \) which are linearly independent over \( F \). We will use the following observation. Suppose that \( A_1, \ldots, A_r \in M_{nn}(F) \) and \( a_1A_1 + \cdots + a_rA_r = 0 \) in \( M_{nn}(K) \). Then since each coefficient of \( a_1A_1 + \cdots + a_rA_r \) is zero, we have that \( A_1 = A_2 = \cdots = A_r = 0 \).

Write \( g(x) = x^r + b_{r-1}x^{r-1} + \cdots + b_0 \) with \( b_i \in K \). Let \( \{a_1 = 1, \ldots, a_r\} \) be a basis of the subspace of \( K \) (recall that \( K \) is a vector space over \( F \)) spanned by \( \{1, b_{r-1}, \ldots, b_0\} \). Expand \( b_i = \sum_{j=1}^r c_{ij}a_j \) with \( c_{ij} \in F \). Let \( f_1(x) = x^r + c_{r-1,1}x^{r-1} + \cdots + c_{0,1} \) and \( f_j(x) = c_{r-1,j}x^{r-1} + \cdots + c_{0,j} \) for \( j \geq 2 \). \( f_i(x) \in F[x] \) for all \( i \), and \( g(x) = \sum_{j=1}^r a_j f_j(x) \).

We have that \( 0 = g(A) = \sum_{i=1}^r a_i f_i(A) \), which implies that \( f_i(A) = 0 \) for all \( i \), as observed above. Thus \( f(x) \) divides \( f_i(x) \) in \( F[x] \) for all \( i \), and then \( f(x) \) divides \( g(x) \) in \( K[x] \). Since \( f(A) = 0, g(x) \) divides \( f(x) \) in \( K[x] \) so \( f \) and \( g \) generate the same ideal, so \( f = g \) is the unique monic generator.

\[
\square
\]

13. THEOREM OF HAMILTON-CAYLEY

**Theorem 13.1.** Let \( V \) be a non zero finite dimensional vector space, over an algebraically closed field \( F \), and let \( L : V \rightarrow V \) be a linear map. Let \( p_L(t) \) be its characteristic polynomial. Then \( p_L(L) = 0 \).

Here “\( p_L(L) = 0 \)” means that \( p_L(L) \) is the zero map on \( V \).

**Proof.** There exists a fan \( \{V_1, \ldots, V_n\} \) of \( L \) in \( V \), with associated fan basis \( \beta = \{v_1, \ldots, v_n\} \).

The matrix \( M_\beta(L) \in M_{n,n}(F) \) is upper triangular. Write \( M_\beta(L) = (a_{ij}) \). Then

\[
p_L(t) = \text{Det}(tI_n - M_\beta(L)) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}).
\]

We shall prove by induction on \( i \) that

\[
(L - a_{11}I) \cdots (L - a_{ii}I)v = 0
\]

for all \( v \in V_i \).

We first prove this in the case \( i = 1 \). Then \( (L - a_{11}I)v = Lv - a_{11}v = 0 \) for \( v \in V_1 \), since \( V_1 \) is in the eigenspace for the eigenvalue \( a_{11} \).

Now assume that \( i > 1 \), and that

\[
(L - a_{11}I) \cdots (L - a_{i-1,i-1}I)v = 0
\]

for all \( v \in V_{i-1} \). Suppose that \( v \in V_i \). Then \( v = v' + cv_i \) for some \( v' \in V_{i-1} \) and \( c \in F \).

Let \( z = (L - a_{ii}I)v' \). \( z \in V_{i-1} \) because \( L(V_{i-1}) \subseteq V_{i-1} \) and \( a_{ii}v' \in V_{i-1} \). By induction,

\[
(L - a_{11}I) \cdots (L - a_{i-1,i-1}I)(L - a_{ii}I)v' = (L - a_{11}I) \cdots (L - a_{i-1,i-1}I)z = 0.
\]

We have \( (L - a_{ii}I)cv_i \in V_{i-1} \), since \( L(v_i) = a_{ii}v_i + a_{i2}v_2 + \cdots + a_{i1}v_1 \). Thus

\[
(L - a_{11}I) \cdots (L - a_{i-1,i-1}I)(L - a_{ii}I)cv_i = 0,
\]

and thus

\[
(L - a_{11}I) \cdots (L - a_{ii}I)v = 0.
\]

19
Corollary 13.2. Let $F$ be an algebraically closed field and suppose that $A \in M_{n,n}(F)$. Let $p_A(t)$ be its characteristic polynomial. Then $p_A(A) = 0$.

Here “$p_A(A) = 0$” means that $p_A(A)$ is the zero matrix.

Proof. Let $L = L_A : F^n \to F^n$. Then $0 = p_L(L)$. Let $\beta$ be the standard basis of $F^n$. Since $A = M^\beta_\beta(L)$, we have that $p_L(t) = p_A(t) \in F[t]$. Now
\[ p_A(A) = p_L(M^\beta_\beta(L)) = M^\beta_\beta(p_L(L)) = M^\beta_\beta(0) = 0 \]
where the right most zero in the above equation denotes the $n \times n$ zero matrix. \hfill $\square$

Corollary 13.3. Suppose that $F$ is a field, and $A \in M_{n,n}(F)$. Then $p_A(A) = 0$.

Proof. Let $K$ be an algebraically closed field containing $F$. From the natural inclusion $M_{n,n}(F) \subset M_{n,n}(K)$ we may regard $A$ as an element of $M_{n,n}(K)$. The computation $p_A(t) = \text{Det}(I_n - A)$ does not depend on the field $F$ or $K$. By Corollary 13.2, $p_A(A) = 0$. \hfill $\square$

Theorem 13.4. (Hamilton-Cayley) Suppose that $V$ is a non zero finite dimensional vector space over a field $F$ and $L : V \to V$ is a linear map. Then $p_L(L) = 0$.

Proof. Let $\beta$ be a basis of $V$, and let $A = M^\beta_\beta(L) \in M_{n,n}(F)$. We have $p_L(t) = p_A(t)$. By Corollary 13.3, $p_A(A) = 0$. Now
\[ M^\beta_\beta(p_L(L)) = p_L(M^\beta_\beta(L)) = p_A(A) = 0. \]
Thus $p_L(L) = 0$. \hfill $\square$

Corollary 13.5. Suppose that $V$ is a non zero finite dimensional vector space over a field $F$ and $L : V \to V$ is a linear map. Then the minimal polynomial $m_L(t)$ divides the characteristic polynomial $p_L(t)$ in $F[t]$.

Proof. By the Hamilton-Cayley Theorem, $p_L(t)$ is in the kernel of the homomorphism $F[t] \to F[L]$ which takes $t$ to $L$ and is the identity on $F$. Since $m_L(t)$ is a generator for this ideal, $m_L(t)$ divides $p_L(t)$. \hfill $\square$

14. Invariant Subspaces

Suppose that $V$ is a finite dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Suppose that $W_1, \ldots, W_r$ are $L$-invariant subspaces of $V$ such that $V = W_1 \oplus \cdots \oplus W_r$. Let $\beta_i = \{v_{i,1}, \ldots, v_{i,\sigma(i)}\}$ be bases of $W_i$, and let
\[ \beta = \{v_{1,1}, \ldots, v_{1,\sigma(1)}, v_{2,1}, \ldots, v_{r,\sigma(r)}\} \]
which is a basis of $V$. Since the $W_i$ are $L$-invariant, the matrix of $L$ with respect to $\beta$ is a block matrix
\[ M^\beta_\beta(L) = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_r \end{pmatrix} \]
where $A_i = M^\beta_{\beta_i}(L|W_i)$ are $\sigma(i) \times \sigma(i)$ matrices.
Lemma 14.1. Suppose that $V$ is a vector space over a field $F$ and $L : V \to V$ is a linear map. Suppose that $f(t) \in F[t]$ and let $W$ be the kernel of $f(L) : V \to V$. Then $W$ is an $L$-invariant subspace of $V$.

Proof. Let $x \in W$. Then $f(L)Ix = Lf(I)x = L0 = 0$. Thus $Lx \in W$. □

Theorem 14.2. Let $f(t) \in F[t]$. Suppose that $f = f_1f_2$ with $f_1, f_2 \in F[t]$ and $\deg f_1 \geq 1$, $\deg f_2 \geq 1$ and $\gcd(f_1, f_2) = 1$. Suppose that $V$ is a vector space over a field $F$, $L : V \to V$ is a linear map and $f(L) = 0$. Let $W_1 = \ker f_1(L)$ and $W_2 = \ker f_2(L)$. Then $V = W_1 \oplus W_2$.

Proof. Since $\gcd(f_1, f_2) = 1$, there exist $g_1, g_2 \in F[t]$ such that $1 = g_1(t)f_1(t) + g_2(t)f_2(t)$, so that
\begin{equation}
1 = g_1(L)f_1(L) + g_2(L)f_2(L) = I
\end{equation}
is the identity map of $V$. Let $x \in V$.
\[x = g_1(L)f_1(L)x + g_2(L)f_2(L)x.\]

Then $g_1(L)f_1(L)x \in W_2$ since
\[f_2(L)g_1(L)f_1(L)x = g_1(L)f_1(L)f_2(L)x = g_1(L)f(L)x = g_1(L)0 = 0.
\]
Similarly, $g_2(L)f_2(L)x \in W_1$. Thus $V = W_1 + W_2$. To show that the sum is direct, we must show that an expression $x = w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$ is uniquely determined by $x$. Apply $g_1(L)f_1(L)$ to this expression to get
\[g_1(L)f_1(L)x = g_1(L)f_1(L)w_1 + g_1(L)f_1(L)w_2 = 0 + g_1(L)f_1(L)w_2.
\]
Now apply (10) to $w_2$, to get
\[w_2 = g_1(L)f_1(L)w_2 \in W_2 = g_1(L)f_1(L)w_2 + 0,
\]
which implies that $w_2 = g_1(L)f_1(L)x$ is uniquely determined by $x$. Similarly, $w_1 = g_2(L)f_2(L)x$ is uniquely determined by $x$. □

Theorem 14.3. Let $V$ be a vector space over a field $F$, and let $L : V \to V$ be a linear map. Suppose $f(t) \in F[t]$ satisfies $f(L) = 0$. Suppose that $f(t) = f_1(t)f_2(t)\cdots f_r(t)$ where $f_i(t) \in F[t]$ and $\gcd(f_i, f_j) = 1$ if $i \neq j$. Let $W_i = \ker f_i(L)$ for $1 \leq i \leq r$. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

Proof. We prove the theorem by induction on $r$, the case $r = 1$ being trivial. We have that $\gcd((f_1, \cdots, f_r)) = 1$ since $\gcd((f_1, f_i)) = 1$ for $2 \leq i \leq r$. Theorem 14.2 thus implies $V = W_1 \oplus W$, where $W = \ker f_2(L)f_3(L)\cdots f_r(L)$. $f_j(L) : W \to W$ for $2 \leq j \leq r$. By induction on $r$, we have that $W = U_2 \oplus \cdots \oplus U_r$, where for $j \geq 2$, $U_j$ is the kernel of $f_j(L) : W \to W$. Thus $V = W_1 \oplus U_2 \oplus \cdots \oplus U_r$.

We will prove that $W_j = U_j$ for $j \geq 2$, which will establish the theorem. We have that $U_j \subset W_j$. Let $v \in W_j$, with $j \geq 2$. Then $v \in W_j \cap W = U_j$. □

Corollary 14.4. Suppose $V$ is a vector space over a field $F$. Let $L : V \to V$ be a linear map. Suppose that a monic polynomial $f(t) \in F[t]$ satisfies $f(L) = 0$, and $f(t)$ splits into linear factors in $F[t]$ as
\[f(t) = (t - \alpha_1)^{m_1}\cdots(t - \alpha_r)^{m_r}\]
where $\alpha_1, \ldots, \alpha_r \in F$ are distinct. Let $W_i = \ker (L - \alpha_i)^{m_i}$. Then
\[V = W_1 \oplus W_2 \oplus \cdots \oplus W_r.\]
As an application of this Corollary, we consider the following example
Let $A(C)$ be the set of analytic functions on $C$, and $D = \frac{d}{dz}$. Suppose that $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_{n-1} \in C$. Let
$$V = \{ f \in A(C) \mid D^n f + a_{n-1}D^{n-1} f + \cdots + a_0 f = 0 \}.$$ 
Then $V$ is a complex vector space. Let
$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0 \in \mathbb{C}[t].$$
Suppose that $p(t)$ factors as
$$p(t) = (t - \alpha_1)^{m_1} \cdots (t - \alpha_r)^{m_r}$$
with $\alpha_1, \ldots, \alpha_r \in C$ distinct. Then by Corollary 14.4, we have that
$$V = \bigoplus_{i=1}^{r} W_i$$
where
$$W_i = \{ f \in A(C) \mid (D - \alpha_i I)^{m_i} f = 0 \}.$$ 
Now from the following Lemma, we are able to construct a basis of $V$. In fact, we see that
$$\{ e^{\alpha_1 x}, xe^{\alpha_1 x}, \ldots, x^{m_1-1}e^{\alpha_1 x}, e^{\alpha_2 x}, xe^{\alpha_2 x}, \ldots, x^{m_r-1}e^{\alpha_r x} \}$$
is a basis of $V$.

**Lemma 14.5.** Let $\alpha \in C$, and $m \in \mathbb{N}$. Let
$$W = \{ f \in A(C) \mid (D - \alpha I)^m f = 0 \}.$$ 
Then
$$\{ e^{\alpha x}, xe^{\alpha x}, \ldots, x^{m-1}e^{\alpha x} \}$$
is a basis of $W$.

**Proof.** It can be verified by induction on $m$ that
$$(D - \alpha I)^m f = e^{\alpha x}D^m(e^{-\alpha x} f)$$
for $f \in A(C)$. Thus $f \in W$ if and only if $D^m(e^{-\alpha x} f) = 0$. The functions whose $m$-th derivatives are zero are the polynomials of degree $\leq m-1$. Hence the space of solutions to $(D - \alpha I^m) f = 0$ is the space generated by $\{ e^{\alpha x}, xe^{\alpha x}, \ldots, x^{m-1}e^{\alpha x} \}$. It remains to verify that these functions are linearly independent. Suppose that there exist $c_i \in \mathbb{C}$ such that
$$c_0e^{\alpha s} + c_1se^{\alpha s} + \cdots + c_{m-1}s^{m-1}e^{\alpha s} = 0$$
for all $s \in \mathbb{C}$. Let
$$q(t) = c_0 + c_1t + \cdots + c_{m-1}t^{m-1} \in \mathbb{C}[t].$$
We have $q(s)e^{\alpha s} = 0$ for all $s \in \mathbb{C}$. But $e^{\alpha s} \neq 0$ for all $s \in \mathbb{C}$, so $q(s) = 0$ for all $s \in \mathbb{C}$. Since $\mathbb{C}$ is infinite, and the equation $q(t) = 0$ has at most a finite number of roots if $q(t) \neq 0$, we must have $q(t) = 0$. Thus $c_i = 0$ for $0 \leq i \leq m - 1$. \hfill $\Box$
15. Cyclic Vectors

**Definition 15.1.** Let $V$ be a finite dimensional vector space over a field $F$, and let $L : V \to V$ be a linear map. $V$ is called $L$-cyclic if there exists $v \in V$ such that

$$V = F[L]v = \{ f(L)v \mid f(t) \in F[t] \}.$$  

We will say that the vector $v$ is $L$-cyclic.

**Lemma 15.2.** Let $V$ be a finite dimensional vector space over a field $F$, and let $L : V \to V$ be a linear map. Suppose that $V$ is $L$-cyclic and $v \in V$ is an $L$-cyclic vector. Let $d$ be the degree of the minimal polynomial $m_L(t) \in F[t]$ of $L$. Then $\{ v, Lv, \ldots, L^{d-1}v \}$ is a basis of $V$.

**Proof.** Suppose $w \in V$. Then $w = f(L)v$ for some $f(t) \in F[t]$. We have $f(t) = p(t)m_L(t) + r(t)$ with $p(t), r(t) \in F[t]$ and $\deg r(t) < d$. Thus $w = f(L)v = r(L)v \in \text{Span}\{v, Lv, \ldots, L^{d-1}v\}$. This shows that $V = \text{Span}\{v, L v, \ldots, L^{d-1}v\}$. Suppose that there is a dependence relation

$$c_0v + c_1Lv + \cdots + c_sL^s v = 0$$

where $0 \leq s \leq d - 1$, all $c_i \in F$ and $c_s \neq 0$. Let $f(t) = c_0 + c_1t + \cdots + c_st^s \in F[t]$. Let $w \in V$. Then there exist $g(t) \in F[t]$ such that $w = g(L)v$.

$$f(L)w = f(L)g(L)v = g(L)f(L)v = g(L)0 = 0.$$  

Thus the operator $f(L) = 0$, which implies that $m_L(t)$ divides $f(t)$, which is impossible since $f(t)$ is nonzero and has degree less than $d$. Thus $\{ v, L v, \ldots, L^{d-1}v \}$ are linearly independent, and a basis of $V$. □

With the hypotheses of Lemma 15.2, express $m_L(t) = a_0 + a_1t + \cdots + a_{d-1}t^{d-1} + t^d$ with the $a_i \in F$. Let $\beta = \{ v, L v, \ldots, L^{d-1}v \}$, a basis of $V$. Then

$$M_\beta^\beta(L) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & -a_1 \\ 0 & 1 & 0 & 0 & -a_2 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & -a_{d-2} \\ 0 & 0 & 0 & 1 & -a_{d-1} \end{pmatrix}.$$  

16. Nilpotent Operators

Suppose that $V$ is a vector space, and $L : V \to V$ is a linear map. We will say that $L$ is nilpotent if there exists a positive integer $r$ such that $L^r = 0$. If $r$ is the smallest positive integer such that $L^r = 0$, then $m_L(t) = t^r$.

**Theorem 16.1.** Let $V$ be a finite dimensional vector space over a field $F$, and let $L : V \to V$ be a nilpotent linear map. Then $V$ is a direct sum of $L$-invariant $L$-cyclic subspaces.

**Proof.** We prove the theorem by induction on $\dim V$. If $V$ has dimension 1, then $V$ is $L$-cyclic. Assume that the theorem is true for vector spaces of dimension less than $s = \dim V$.

Let $r$ be the smallest integer such that $L^r = 0$, so that the minimal polynomial of $L$ is $m_L(t) = t^r$. Since $L^{r-1} \neq 0$, there exists $w \in V$ such that $L^{r-1}w \neq 0$. Let $v = L^{r-1}w$. $Lv = 0$ so Kernel $L \neq 0$.

$$\dim L(V) + \dim \text{Kernel } L = \dim V$$
implies $\dim L(V) < \dim V$.

By induction on $s = \dim V$, $L(V)$ is a direct sum of $L$-invariant subspaces which are $L$-cyclic, say

$$L(V) = W_1 \bigoplus \cdots \bigoplus W_m.$$ 

Let $w_i \in W_i$ be an $L|W_i$-cyclic vector, so that if $r_i$ is the degree of the minimal polynomial $p_{L_i}(t)$, then $W_i$ has a basis $\{w_i, Lw_i, \ldots, L^{r_i-1}w_i\}$ by Lemma 15.2. Since $L|W_i$ is nilpotent, $m_{L_i|W_i}(t) = t^{r_i}$. Let $v_i \in V$ be such that $Lv_i = w_i$.

Let $V_i$ be the cyclic subspace $F[L]v_i$ of $V$. $L^{r_i+1}v_i = L^{r_i}w_i = 0$ implies that $\{v_i, Lv_i, \ldots, L^{r_i}v_i\}$ spans $V_i$. Suppose that

$$d_0v_i + d_1Lv_i + \cdots + d_{r_i}L^{r_i}v_i = 0$$

with $d_0, \ldots, d_{r_i} \in F$. Then

$$0 = L(0) = L(d_0v_i + d_1Lv_i + \cdots + d_{r_i}L^{r_i}v_i) = d_0Lv_i + \cdots + d_{r_i-1}L^{r_i-1}v_i+ d_{r_i}L^{r_i}v_i = d_0w_i + \cdots + d_{r_i-1}L^{r_i-1}w_i.$$ 

Thus $d_0 = \cdots = d_{r_i-1} = 0$, since $\{w_i, \ldots, L^{r_i-1}w_i\}$ is a basis of $W_i$. Now since $L^{r_i}v_i = L^{r_i-1}w_i \neq 0$, we have $d_{r_i} = 0$. Thus $\{v_i, Lv_i, \ldots, L^{r_i}v_i\}$ are linearly independent, and are thus a basis of $V_i$.

We will prove that the subspace $V' = V_1 + \cdots + V_m$ of $V$ is a direct sum. We have to prove that if

$$(11) \quad 0 = u_1 + \cdots + u_m$$

with $u_i \in V_i$, then $u_i = 0$ for all $i$. Since $u_i \in V_i$, we have an expression $u_i = f_i(L)v_i$ where $f_i(t) \in F[t]$ is a polynomial. Thus (11) becomes

$$(12) \quad f_1(L)v_1 + \cdots + f_m(L)v_m = 0.$$ 

Apply $L$ to (12), and recall that $Lf_i(L) = f_i(L)L$, to get

$$f_1(L)w_1 + \cdots + f_m(L)w_m = 0.$$ 

Now $W_1 + \cdots + W_m$ is a direct sum decomposition of $L(V)$ by $L$-invariant subspaces, so $f_i(L)w_i = 0$ (for all $i$ with $1 \leq i \leq m$) which implies that $f_i(L)|(W_i) = 0$, since $W_i$ is $w_i$-cyclic. Thus $m_{L_i|W_i}(t) = t^{r_i}$ divides $f_i(t)$. In particular, $t$ divides $f_i(t)$ in $F[t]$. Write $f_i(t) = g_i(t)t$ for some polynomial $g_i(t)$. Then $f_i(L) = g_i(L)L$. (12) implies

$$g_1(L)w_1 + \cdots + g_m(L)w_m = 0.$$ 

Thus $g_i(L)w_i = 0$ for all $i$, since $L(V)$ is the direct sum of the $W_i$. This implies that $t^{r_i}$ divides $g_i(t)$ in $F[t]$ so that $t^{r_i+1}$ divides $f_i(t)$ which implies that $u_i = f_i(L)v_i = 0$. Thus $V'$ is the direct sum $V' = V_1 \bigoplus \cdots \bigoplus V_m$.

An element of $L(V)$ is of the form

$$f_1(L)w_1 + \cdots + f_m(L)w_m = f_1(L)Lv_1 + \cdots + f_m(L)Lv_m = L(f_1(L)v_1 + \cdots + f_m(L)v_m)$$

for some polynomials $f_i(t) \in F[t]$. Thus $L(V') = L(V)$. Now let $v, v' \in V$. $Lv = Lv'$ for some $v' \in V'$. Then $L(v - v') = 0$. Thus $v = v' + (v - v') \in V' + \text{Kernel } L$. We conclude that $V = V' + \text{kernel } L$ (which may not be a direct sum).

Let $\beta' = \{L^iv_1 \mid 1 \leq i \leq m, 1 \leq j \leq r_i\}$ be the basis we have constructed of $V'$. We extend $\beta'$ to a basis of $V$ by using elements of kernel $L$, to get a basis $\beta = \{\beta', z_1, \ldots, z_e\}$ of $V$ where $z_1, \ldots, z_e \in \text{kernel } L$. Each $z_j$ satisfies $Lz_j = 0$, so $z_j$ is an eigenvector for $L$. 


and the one dimensional space generated by \( z_j \) is \( L \)-invariant and cyclic. Let this subspace be \( Z_j \). We have

\[
V = V' \bigoplus Z_1 \cdots \bigoplus Z_e = V_1 \bigoplus \cdots \bigoplus V_m \bigoplus Z_1 \cdots \bigoplus Z_e
\]

expressing \( V \) as a direct sum of \( L \)-cyclic subspaces.

\[
\square
\]

17. JORDAN FORM

**Definition 17.1.** Suppose that \( V \) is a finite dimensional vector space over a field \( F \), and \( L : V \to V \) is a linear map. A basis \( \beta \) of \( V \) is a Jordan basis for \( L \) if the matrix \( M^\beta_\beta(L) \) is a block matrix

\[
J = M^\beta_\beta(L) = \begin{pmatrix}
J_1 & 0 & \cdots & 0 \\
0 & J_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & J_m
\end{pmatrix},
\]

where each \( J_i \) is a Jordan block; that is a matrix of the form

\[
J_i = \begin{pmatrix}
\alpha_i & 1 & 0 & \cdots & 0 & 0 \\
0 & \alpha_i & 1 & \cdots & 0 & 0 \\
0 & 0 & \alpha_i & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \alpha_i & 1 \\
0 & 0 & 0 & \cdots & 0 & \alpha_i
\end{pmatrix}
\]

for some \( \alpha_i \in F \).

A matrix of the form \( J \) is called a Jordan form.

**Theorem 17.2.** Suppose that \( V \) is a finite dimensional vector space over a field \( F \), and \( L : V \to V \) is a linear map. Suppose that the minimal polynomial \( m_L(t) \) splits into linear factors in \( F[t] \). Then \( V \) has a Jordan basis for \( L \).

**Proof.** By assumption, we have a factorization \( m_L(t) = (t - \alpha_1)^{a_1} \cdots (t - \alpha_s)^{a_s} \) with \( \alpha_1, \ldots, \alpha_s \in F \) distinct. Let

\[
V_i = \{ v \in V \mid (L - \alpha_iI)^{a_i}v = 0 \}
\]

for \( 1 \leq i \leq s \). We proved in Lemma 14.1 and Corollary 14.4 that \( V_i \) are \( L \)-invariant subspaces of \( V \) and \( V = V_1 \bigoplus \cdots \bigoplus V_s \). It thus suffices to prove the theorem in the case that there exists \( \alpha \in F \) and \( r \in \mathbb{N} \) such that \( (L - \alpha I)^r = 0 \). In this case we apply Theorem 16.1 to the nilpotent operator \( T = L - \alpha I \), to show that \( V \) is a direct sum of \( T \)-cyclic subspaces. Since a \( T \)-invariant subspace is \( L \)-invariant, it suffices to prove the theorem in the case that \( T \) is nilpotent, and \( V \) is \( T \)-cyclic. Let \( v \) be a \( T \)-cyclic vector. There exists a positive integer \( r \) such that the minimal polynomial of \( T \) is \( m_T(t) = t^r \), and by Lemma 15.2, \( \{ T^{-1}v, \ldots, T^rv, v \} \) is a basis of \( V \). Recalling that \( T = L - \alpha I \), we have

\[
L(L - \alpha I)^{r-1}v = \alpha(L - \alpha I)^{r-1}v \\
L(L - \alpha I)^{r-2}v = (L - \alpha I)^{r-1}v + \alpha(L - \alpha I)^{r-2}v \\
\vdots \\
L^r v = (L - \alpha)^r v + \alpha v.
\]
Thus the matrix of $L$ with respect to the basis $\beta$ is the Jordan block

$$M_\beta^\beta(L) = \begin{pmatrix} \alpha & 1 & 0 & \cdots & 0 & 0 \\ 0 & \alpha & 1 & \cdots & 0 & 0 \\ 0 & 0 & \alpha & \cdots & 0 & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha & 1 \\ 0 & 0 & 0 & \cdots & 0 & \alpha \end{pmatrix}.$$  

□

The proof of the following theorem reduces the calculation of the Jordan form of an operator $L$ to a straightforward calculation, if we are able to determine the eigenvalues (roots of the characteristic equation) of $L$.

**Theorem 17.3.** Suppose that $V$ is a finite dimensional vector space over a field $F$, and $L : V \to V$ is a linear map. Suppose that the minimal polynomial $m_L(t)$ splits into linear factors in $F[t]$. For $\alpha \in F$ and $i \in \mathbb{N}$, define

$$s_i(\alpha) = \dim \text{Kernel} \left( L - \alpha I \right)^i.$$

Then the Jordan form of $L$ is uniquely determined, up to permutation of Jordan blocks, by the $s_i(\alpha)$.

In particular, the Jordan Form of $L$ is uniquely determined up to permutation of Jordan blocks.

**Proof.** Suppose that $\alpha \in F$. Let $s_i = s_i(\alpha)$ for $1 \leq i \leq n = \dim V$. Suppose that $\beta$ is a Jordan basis of $V$. Let

$$A = M_\beta^\beta(L) = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix}$$

where the

$$J_i = \begin{pmatrix} \alpha_i & 1 & 0 & \cdots & 0 \\ 0 & \alpha_i & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_i \end{pmatrix}$$

are Jordan blocks. Suppose that $J_i$ has size $n_i$ for $1 \leq i \leq s$. We have that

$$\dim \text{Kernel} \left( L - \alpha I \right)^i = \dim \text{Kernel} \left( L M_\beta^\beta(L - \alpha I)^i \right) = \dim \text{Kernel} \left( L (A - \alpha I_n)^i \right)$$

where $I_n$ is the $n \times n$ identity matrix and $L M_\beta^\beta(L - \alpha I)^i$ is the associated linear map from $F^n$ to $F^n$. We have that

$$(A - \alpha I_n)^i = \begin{pmatrix} (J_1 - \alpha I_{n_1})^i & 0 & \cdots & 0 \\ 0 & (J_2 - \alpha I_{n_2})^i & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & (J_s - \alpha I_{n_s})^i \end{pmatrix}.$$
Now if \( \alpha_j \neq \alpha \), then
\[
(J_j - \alpha I_{n_j})^i = \begin{pmatrix}
(\alpha_j - \alpha)^i & * & \cdots & * \\
0 & (\alpha_j - \alpha)^i & * & \cdots & * \\
\vdots & \vdots & \ddots & \vdots & \ddots \\
0 & 0 & 0 & (\alpha_j - \alpha)^i 
\end{pmatrix},
\]
and
\[
(J_j - \alpha_j I_{n_j})^i = \begin{pmatrix}
0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & 0 
\end{pmatrix}
\]
for \( 1 \leq i \leq n_j - 1 \), where the 1 in the first row occurs in the \( i + 1 \)-st column. We also have that \((J_j - \alpha_j I_{n_j})^n_j = 0\). We have
\[
\dim \text{ Kernel } L((J_j - \alpha I_{n_j})^i) = \begin{cases} 
0 & \text{if } \alpha_j \neq \alpha \\
\min\{i, n_j\} & \text{if } \alpha_j = \alpha.
\end{cases}
\]
Let \( r_j \) be the number of Jordan blocks \( J_j \) of size \( n_j = i \) and with \( \alpha_j = \alpha \). Let
\[
\begin{align*}
s_1 &= r_1 + \cdots + r_n \\
&\vdots \\
s_i &= r_1 + 2r_2 + 3r_3 + \cdots + i(r_i + \cdots + r_n) \\
&\vdots \\
s_n &= r_1 + 2r_2 + 3r_3 + \cdots + nr_n.
\end{align*}
\]
From the above system we obtain \( r_1 = 2s_1 - s_2, r_i = 2s_i - s_{i+1} - s_{i-1} \) for \( 2 \leq i \leq n - 1 \) and \( r_n = s_n - s_{n-1} \). Since the Jordan form of \( L \) is determined up to permutation of the Jordan blocks by the knowledge of the \( r_i = r_i(\alpha) \) for \( \alpha \in F \), and the \( s_i = s_i(\alpha) \) are completely determined by \( L \), the Jordan form of \( L \) is completely determined up to permutation of Jordan blocks.

18. Exercises

Suppose that \( V \) is a finite dimensional vector space over an algebraically closed field \( F \), and \( L : V \to V \) is a linear map.

1. If \( L \) is nilpotent and not the zero map, show that \( L \) is not diagonalizable.
2. Show that \( L \) can be written in the form \( L = D + N \) where \( D : V \to V \) is diagonalizable, \( N : V \to V \) is nilpotent, and \( DN = ND \).
3. Give a formula for the minimal polynomial of \( L \), in terms of its Jordan form.
4. Let \( p_L(t) \) be the characteristic polynomial of \( L \), and suppose that it has a factorization
\[
p_L(t) = \prod_{i=1}^{r} (t - \alpha_i)^{m_i}
\]
where \( \alpha_1, \ldots, \alpha_r \) are distinct. Let \( f(t) \) be a polynomial. Express the characteristic polynomial \( p_{f(t)}(t) \) as a product of factors of degree 1.
19. Inner Products

Let \( \overline{z} \) denote complex conjugation, for \( z \in \mathbb{C} \). \( z \in \mathbb{R} \) if and only if \( \overline{z} = z \).

Suppose that \( V \) is a vector space over \( F \), where \( F = \mathbb{R} \) or \( F = \mathbb{C} \). An inner product on \( V \) is a (respectively, real or complex) valued function on \( V \times V \) such that for \( v_1, v_2, w \in V \) and \( \alpha_1 \in F \)

1. \( < v_1 + v_2, w > = < v_1, w > + < v_2, w > \),
2. \( < \alpha_1 v_1, w > = \alpha_1 < v_1, w_1 > \),
3. \( < v, w > = \overline{< w, v >} \),
4. \( < w, w > \geq 0 \) and \( < w, w > = 0 \) if and only if \( w = 0 \).

An inner product space is a vector space with an inner product. A complex inner product is often called an Hermitian product.

A real inner product is a positive definite, symmetric bilinear form. However, an Hermitian inner product is not even a bilinear form. In spite of this difference, the theory for real and complex inner products is very similar.

From now on in this section, suppose that \( V \) is an inner product space, with an inner product \( < , > \).

For \( v \in V \), we define the norm of \( v \) by
\[
||v|| = \sqrt{< v, v >}.
\]

From the definition of an inner product, we obtain the polarization identities. If \( V \) is a real inner product space, then for \( v, w \in V \), we have
\[
< v, w > = \frac{1}{4} ||v + w||^2 - \frac{1}{4} ||v - w||^2.
\]
If \( V \) is an Hermitian inner product space, then for \( v, w \in V \), we have
\[
< v, w > = \frac{1}{4} ||v + w||^2 - \frac{1}{4} ||v - w||^2 + i \frac{1}{4} ||v + iw||^2 - i \frac{1}{4} ||v - iw||^2.
\]

A property of inner products that we will use repeatedly is the fact that if \( x, y \in V \) and \( < v, x > = < v, y > \) for all \( v \in V \), then \( x = y \).

We say that vectors \( v, w \in V \) are orthogonal or perpendicular if \( < v, w > = 0 \). Suppose that \( S \subset V \) is a subset. Define
\[
S^\perp = \{ v \in V | < v, w > = 0 \text{ for all } w \in S \}.
\]

**Lemma 19.1.** Suppose that \( S \) is a subset of \( V \). Then
1. \( S^\perp \) is a subspace of \( V \).
2. If \( U \) is the subspace \( \text{Span}(S) \) of \( V \), then \( U^\perp = S^\perp \).

A set of nonzero vectors \( \{v_1, \ldots, v_r\} \) in \( V \) are called orthogonal if \( < v_i, v_j > = 0 \) whenever \( i \neq j \).

**Lemma 19.2.** Suppose that \( \{v_1, \ldots, v_r\} \) are nonzero orthogonal vectors in \( V \). Then \( v_1, \ldots, v_r \) are linearly independent.

A set of vectors \( \{u_1, \ldots, u_r\} \) in \( V \) are called orthonormal if \( < u_i, u_j > = 0 \) if \( i \neq j \) and \( ||u_i|| = 1 \) for all \( i \).

**Lemma 19.3.** Suppose that \( \{u_1, \ldots, u_r\} \) are orthonormal vectors in \( V \). Then \( u_1, \ldots, u_r \) are linearly independent.

A basis of \( V \) consisting of orthonormal vectors is called an orthonormal basis.
Lemma 19.4. Let \( \{v_1, \ldots, v_s\} \) be a set of nonzero orthogonal vectors in \( V \), and \( v \in V \). Let
\[
c_i = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle}.
\]
1. We have that
\[
v - \sum_{i=1}^{s} c_i v_i \in \text{Span}\{v_1, \ldots, v_s\}^\perp.
\]
2. Let \( a_1, \ldots, a_s \in F \). Then
\[
\|v - \sum_{i=1}^{s} c_i v_i\| \leq \|v - \sum_{i=1}^{s} a_i v_i\|
\]
with equality if and only if \( a_i = c_i \) for all \( i \). Thus \( c_1 v_1 + \cdots + c_s v_s \) gives the best approximation to \( v \) as a linear combination of \( v_1, \ldots, v_s \).

Proof. We first prove 1. For \( 1 \leq j \leq s \), we have
\[
\langle v - \sum_{i=1}^{s} c_i v_i, v_j \rangle = \langle v, v_j \rangle - \sum_{i=1}^{s} c_i \langle v_i, v_j \rangle = \langle v, v_j \rangle - c_j \langle v_j, v_j \rangle = 0.
\]
We now prove 2. We have that
\[
\|v - \sum_{i=1}^{s} a_i v_i\|^2 = \|v - \sum_{i=1}^{s} c_i v_i + \sum_{i=1}^{s} (c_i - a_i) v_i\|^2 = \|v - \sum_{i=1}^{s} c_i v_i\|^2 + \|\sum_{i=1}^{s} (c_i - a_i) v_i\|^2
\]
since \( v - \sum_{i=1}^{s} c_i v_i \in \text{Span}\{v_1, \ldots, v_s\}^\perp \).

Theorem 19.5. (Gram Schmidt) Suppose that \( V \) is a finite dimensional inner product space. Suppose that \( \{u_1, \ldots, u_r\} \) is a set of orthonormal vectors in \( V \). Then \( \{u_1, \ldots, u_r\} \) can be extended to an orthonormal basis \( \{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\} \) of \( V \).

Proof. First extend \( u_1, \ldots, u_r \) to a basis \( \{u_1, \ldots, u_r, v_{r+1}, \ldots, v_n\} \) of \( V \). Inductively define
\[
w_i = v_i - \sum_{j=0}^{i-1} \langle v_i, u_j \rangle u_j
\]
and
\[
u_i = \frac{1}{\|w_i\|} w_i
\]
for \( r + 1 \leq i \leq n \).

Let \( V_r = \text{Span}\{u_1, \ldots, u_r\} \) and \( V_i = \text{Span}\{u_1, \ldots, u_r, v_{r+1}, \ldots, v_i\} \) for \( r + 1 \leq i \leq n \). Then \( V_i = \text{Span}\{u_1, \ldots, u_i\} \) for \( r + 1 \leq i \leq n \). By Lemma 19.4, we have that \( u_i \in (V^{i-1})^\perp \) for \( r + 1 \leq i \leq n \). Thus \( \{u_1, \ldots, u_n\} \) are an orthonormal set of vectors which form a basis of \( V \).

Corollary 19.6. Suppose that \( W \) is a subspace of \( V \). Then \( V = W \oplus W^\perp \).

Proof. First construct, using Gram Schmidt, an orthonormal basis \( \{w_1, \ldots, w_s\} \) of \( W \). Now apply Gram Schmidt again to extend this to an orthonormal basis \( \{w_1, \ldots, w_s, u_1, \ldots, u_r\} \) of \( V \). Let \( U = \text{Span}\{u_1, \ldots, u_r\} \). We have that \( V = W \oplus U \). It remains to show that \( U = W^\perp \). Let \( x \in U \). Then \( x = \sum_{i=1}^{r} c_i u_i \) for some \( c_i \in F \). For all \( j \),
\[
\langle x, w_j \rangle = \sum_{i=1}^{r} c_i \langle u_i, w_j \rangle = 0.
\]
Thus $x \in W^\perp$ and $U \subset W^\perp$.

Suppose $x \in W^\perp$. Expand $x = \sum_{i=1}^r c_i u_i + \sum_{j=1}^s d_j w_j$ for some $c_i, d_j \in F$. For all $k$,

$$0 = <x, w_k> = \sum_{i=1}^r c_i <u_i, w_k> + \sum_{j=1}^s d_j <w_j, w_k> = d_k.$$

Thus $x \in U$ and $W^\perp \subset U$. □

We now can give another proof of Theorem 7.5, when $F = \mathbb{R}$.

**Corollary 19.7.** Suppose that $A \in M_{m,n}(\mathbb{R})$. Then the row rank of $A$ is equal to the column rank of $A$.

Suppose that $V$, $W$ are inner product spaces (both over $\mathbb{R}$ or over $\mathbb{C}$), with inner products $<,>$ and $[,]$ respectively. We will say that a linear map $\varphi : V \to W$ is a map of inner product spaces if $<v, w> = [\varphi(v), \varphi(w)]$ for all $v, w \in V$.

**Corollary 19.8.** Suppose that $V$ is a real inner product space of dimension $n$. Consider $\mathbb{R}^n$ as an inner product space with the dot product. Then there is an isomorphism $\varphi : V \to \mathbb{R}^n$ of inner product spaces.

**Corollary 19.9.** Suppose that $V$ is an Hermitian inner product space of dimension $n$. Consider $\mathbb{C}^n$ as an inner product space with the standard Hermitian product, $[v, w] = v^\dagger w$ for $v, w \in V$. Then there is an isomorphism $\varphi : V \to \mathbb{C}^n$ of inner product spaces.

**Proof.** Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of $V$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$. We can thus define a linear map $\varphi : V \to \mathbb{C}^n$ by $\varphi(v) = a_1 e_1 + \cdots + a_n e_n$ if $v = a_1 v_1 + \cdots + a_n v_n$ with $a_1, \ldots, a_n \in \mathbb{C}$. Since $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{C}^n$, $\varphi$ is an isomorphism. We have $<v_i, v_j> = \delta_{ij}$ and $[\varphi(v_i), \varphi(v_j)] = [e_i, e_j] = \delta_{ij}$ for $1 \leq i, j \leq n$.

Suppose that $v = a_1 v_1 + \cdots + a_n v_n \in V$ and $w = b_1 v_1 + \cdots + b_n v_n \in V$. Then

$$<v, w> = \sum_{i,j=1}^n a_i b_j \delta_{ij} = \sum_{i=1}^n a_i b_i.$$

We also calculate

$$[\varphi(v), \varphi(w)] = \left[ \sum_{i=1}^n a_i e_i, \sum_{j=1}^n b_j e_j \right] = \sum_{i,j=1}^n a_i b_j \delta_{ij} = \sum_{i=1}^n a_i b_i.$$

Thus $<v, w> = [\varphi(v), \varphi(w)]$ for all $v, w \in V$. □

### 20. Symmetric, Hermitian, Orthogonal and Unitary Operators

Throughout this section, assume that $V$ is a finite dimensional inner product space. Let $<,>$ be the inner product. Suppose that $L : V \to V$ is a linear map.

Recall that the dual space of $V$ is the vector space $V^* = L_F(V, K)$, and $\dim V^* = \dim V$.

**Lemma 20.1.** Suppose that $z \in V$. Then the map $\varphi_z : V \to F$ defined by $\varphi_z(v) = <v, z>$ is a linear map. Suppose that $\psi \in V^*$. Then there exists a unique $z \in V$ such that $\psi = \varphi_z$.

**Proof.** Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of $V$ with dual basis $\{v_1^*, \ldots, v_n^*\}$. Suppose that $\psi \in V^*$. Then there exist (unique) $c_1, \ldots, c_n \in F$ such that $\psi = c_1 v_1^* + \cdots + c_n v_n^*$. Let $z = c_1 v_1 + \cdots + c_n v_n \in V$. We have that $\psi(v_i) = c_i$ for $1 \leq i \leq n$.

$$\varphi_z(v_i) = <v_i, z> = c_1 <v_1, v_1> + \cdots + c_n <v_n, v_n> = c_i = \psi(v_i)$$
for $1 \leq i \leq n$. Since $\{v_1, \ldots, v_n\}$ is a basis of $V$, we have that $\psi = \varphi$. Suppose that $w \in V$ and $\varphi_w = \psi$. Expand $v = d_1v_1 + \cdots + d_nv_n$. We have that
\[ c_i = \varphi_w(v_i) = \bar{d}_i \]
for $1 \leq i \leq n$. Thus $d_i = c_i$ for $1 \leq i \leq n$, and $w = z$. \hfill $\Box$

**Theorem 20.2.** There exists a unique linear map $L^* : V \rightarrow V$ such that $<Lv, w> = <v, L^*w>$ for all $v, w \in V$.

**Proof.** For fixed $w \in V$, the map $\psi_w : V \rightarrow F$ defined by $\psi_w(v) = <L(v), w>$ is a linear map. Hence $\psi_w \in V^*$. By Lemma 20.1, there exists a unique vector $z_w \in V$ such that $\psi_w = \varphi_{z_w}$ (where $\varphi_{z_w}(v) = <v, z_w>$ for $v \in V$). Thus we have a function $L^* : V \rightarrow V$ defined by $L^*(w) = z_w$ for $w \in V$.

We will now verify that $L^*$ is linear. Suppose $w_1, w_2 \in V$. For $v \in V$,
\[ <v, L^*(w_1 + w_2)\times\langle v, L^*(w_1) + L^*(w_2) \rangle. \]

Since this identity holds for all $v \in V$, we have that $L^*(w_1 + w_2) = L^*(w_1) + L^*(w_2)$.

Suppose $w \in V$ and $c \in F$.
\[ <v, L^*(cw)\times\langle v, cL^*(w) \rangle. \]

Since this identity holds for all $v \in V$, we have that $L^*(cw) = cL^*(w)$. Thus $L^*$ is linear.

Now we prove uniqueness. Suppose that $T : V \rightarrow V$ is a linear map such that $<L(v), w> = <v, T(w)>$ for all $v, w \in V$. Then for any $w \in W$, $<v, T(w) > = <v, L^*(w) >$ for all $v \in V$. Thus $T(w) = L^*(w)$, and $L^*$ is unique. \hfill $\Box$

$L^*$ is called the adjoint of $L$.

**Definition 20.3.** $L$ is called self adjoint if $L^* = L$.

If $V$ is a real inner product space, a self adjoint operator is called symmetric, and we sometimes write $L^* = L^t$. If $V$ is an Hermitian inner product space, a self adjoint operator is called Hermitian.

**Lemma 20.4.** Let $V = \mathbb{R}^n$ with the standard inner product. Suppose that $L = L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ for some $A \in M_{nn}(\mathbb{R})$. Then $L^* = L_{A^t}$.

**Lemma 20.5.** Let $V = \mathbb{C}^n$ with the standard Hermitian inner product. Suppose that $L = L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ for some $A \in M_{nn}(\mathbb{C})$. Then $L^* = L_{A^t}^\dagger$.

**Proof.** Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$. Let $B \in M_{n,n}(\mathbb{C})$ be the unique matrix such that $L^* = L_B$. Write $A = (a_{ij})$ and $B = (b_{ij})$. We have that $e_i^tBe_j = b_{ij}$. For $1 \leq i, j \leq n$, we have
\[ <Le_i, e_j> = <e_i, L^*e_j> = <e_i, Be_j> = e_i^t\overline{Be_j} = e_i\overline{B}e_j = \overline{b}_{ij}. \]

We also calculate
\[ <Le_i, e_j> = <Ae_i, e_j> = (Ae_i)^t\tau_j = e_i^tA^t\tau_j = \overline{a}_{ji}. \]

Thus $b_{ij} = \overline{a}_{ji}$, and $B = A^\dagger$. \hfill $\Box$

**Lemma 20.6.** The following are equivalent:
1. $LL^* = I$.
2. $||Lv|| = ||v||$ for all $v \in V$.
3. $<Lv, Lw> = <v, w>$ for all $v, w \in V$
Proof. Suppose that 1. holds, so that $LL^* = I$. Then $L^*$ is a right inverse of $L$ so that $L$ is invertible with inverse $L^*$. Thus $L^*L = I$. Suppose that $v \in V$. Then
\[ ||Lv||^2 = <Lv, Lv> = <v, L^*Lv> = <v, v> = ||v||^2, \]
and thus 2. holds.

Now suppose that 2. holds. Then 3. holds by the appropriate polarization identity (13) or (14).

Finally suppose that 3. holds. For fixed $w \in V$, we have $<v, w> = <Lv, Lw> = <v, L^*Lw>$ for all $v \in V$. Thus $w = L^*Lw$. Since this holds for all $w \in W$, $L^*L = I$, so that $L$ is an isomorphism with $LL^* = I$, and 1. holds. $\square$

The above Lemma motivates the following definition.

**Definition 20.7.** $L : V \to V$ is an isometry if $LL^* = I$, where $I$ is the identity map of $V$.

If $V$ is a real inner product space, then an isometry is called an orthogonal transformation (or real unitary). If $V$ is an Hermitian inner product space, then an isometry is called a unitary transformation.

**Lemma 20.8.** Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of $V$. Then $L$ is an isometry if and only if $\{Lv_1, \ldots, Lv_n\}$ is an orthonormal basis of $V$.

**Theorem 20.9.** If $W$ is an $L$-invariant subspace of $V$ and $L = L^*$, then $W^\perp$ is an $L$-invariant subspace of $V$.

Proof. Let $v \in W^\perp$. For all $w \in W$ we have $<Lv, w> = <v, L^*w> = <v, Lw> = 0$, since $Lw \in W$. Hence $Lv \in W^\perp$. $\square$

**Lemma 20.10.** Suppose that $V$ is an Hermitian inner product space and $L$ is Hermitian. Suppose that $v \in V$. Then $<Lv, v> \in \mathbb{R}$.

Proof. We have $<Lv, v> = <v, L^*v> = <v, Lv> = \overline{<Lv, v>}$. Thus $<Lv, v> \in \mathbb{R}$. $\square$

## 21. The Spectral Theorem

**Lemma 21.1.** Let $A \in M_{nn}(\mathbb{R})$ be symmetric. Then $A$ has a real eigenvalue.

Proof. Regarding $A$ as a matrix in $M_{nn}(\mathbb{C})$, we have that $A$ is Hermitian. Now $A$ has a complex eigenvalue $\lambda$, with an eigenvector $v \in \mathbb{C}^n$. Let $<,>$ be the standard Hermitian product on $\mathbb{C}^n$. We have that $A = A^* = \overline{A}$, so $L_A : \mathbb{C}^n \to \mathbb{C}^n$ is Hermitian, by Lemma 20.5. By Lemma 20.10, we then have that $<Av, v> = \lambda <v, v>$ is real. Since $<v, v>$ is real we have that $\lambda$ is real. $\square$

**Theorem 21.2.** Suppose that $V$ is a finite dimensional real vector space with an inner product. Suppose that $L : V \to V$ is a symmetric linear map. Then $L$ has an eigenvector.

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be an orthonormal basis of $V$. We have an isomorphism $\Phi : V \to \mathbb{R}^n$ defined by $\Phi(v) = (v)_{\beta}$ for $v \in V$. Let $[\cdot, \cdot]$ be the dot product on $\mathbb{R}^n$. Let $A = M^\beta(L) \in M_{nn}(\mathbb{R})$. We have
\[ <v, w> = [(v)_{\beta}, (w)_{\beta}] \]
for $v, w \in V$. For $v, w \in V$, we calculate
\[ <L(v), w> = [(L(v))_{\beta}, (w)_{\beta}] = [A(v)_{\beta}, (w)_{\beta}] = [(v)_{\beta}, A^t(w)_{\beta}] \]
and
\[ < L(v), w > = < v, L(w) > = [(v)_\beta, A(w)_\beta]. \]

Since \([(v)_\beta, A^t(w)_\beta] = [(v)_\beta, A(w)_\beta] \) for all \(v, w \in V\), we have that \(A = A^t\). Thus \(A\) has a real eigenvalue \(\lambda\) which necessarily has a real eigenvector \(v\) by Lemma 21.1. Let \(y = \Phi^{-1}(x) \in V\).

\[(L(y))_\beta = A(y)_\beta = Ax = \lambda x = (\lambda y)_\beta.\]

Thus \(L(y) = \lambda y\). Since \(y\) is nonzero, as \(x\) is nonzero, we have that \(y\) is an eigenvector of \(L\). \(\square\)

**Theorem 21.3.** Suppose that \(V\) is a finite dimensional real vector space with an inner product. Suppose that \(L : V \rightarrow V\) is a symmetric linear map. Then \(V\) has an orthonormal basis consisting of eigenvectors of \(L\).

**Proof.** We prove the theorem by induction on \(n = \dim V\). By Theorem 21.2 there exists an eigenvector \(v \in V\) for \(L\). Let \(W\) be the one dimensional subspace of \(V\) generated by \(v\). \(W\) is \(L\)-invariant. By Theorem 20.9, \(W^\perp\) is also \(L\)-invariant. We have that \(V \cong W \bigoplus W^\perp\), so \(\dim W^\perp = \dim V - 1\). The restriction of \(L\) to \(W^\perp\) is a symmetric linear map of \(W^\perp\) to itself. By induction on \(n\), there exists an orthonormal basis \(\{u_2, \ldots, u_n\}\) of \(W^\perp\) consisting of eigenvectors of \(L\). Thus \(\{u_1 = \frac{1}{\|v\|}v, u_2, \ldots, u_n\}\) is an orthonormal basis of \(V\) consisting of eigenvectors of \(L\). \(\square\)

**Corollary 21.4.** Let \(A \in M_{nn}(\mathbb{R})\) be symmetric. Then there exists an orthogonal matrix \(P \in M_{n,n}(\mathbb{R})\) such that \(P^tAP = P^{-1}AP\) is a diagonal matrix.

**Proof.** \(\mathbb{R}^n\) has an orthonormal basis \(\beta = \{u_1, \ldots, u_n\}\) of eigenvectors of \(L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n\) by Lemma 20.4 and Theorem 21.3. Let \(\lambda_1, \ldots, \lambda_n\) be the respective eigenvalues. Then \(M^\beta(L_A)\) is the diagonal matrix \(D\) with \(\lambda_1, \lambda_2, \ldots, \lambda_n\) as diagonal elements. Let \(\beta^t = \{e_1, \ldots, e_n\}\) be the standard basis of \(\mathbb{R}^n\). Let \(P = M^\beta(I) = (u_1, \ldots, u_n)\). \(P^{-1} = M^{\beta^t}(I)\). \(L_P\) is an isometry by Lemma 20.8, since \(Pe_i = u_i\) for \(1 \leq i \leq n\). Thus \(P\) is orthogonal and \(P^{-1} = P^t\) by Lemmas 20.4 and 20.6.

\[ D = M^\beta(L_A) = M^{\beta^t}(I)M^\beta(I)L_AM^\beta(I) = P^{-1}AP = P^tAP. \]

\(\square\)

The proof of Theorem 21.3 also proves the following theorem.

**Theorem 21.5.** Suppose that \(V\) is a finite dimensional complex vector space with an Hermitian inner product. Suppose that \(L : V \rightarrow V\) is a Hermitian linear map. Then \(V\) has an orthonormal basis consisting of eigenvectors of \(L\). All eigenvalues of \(L\) are real.

**Corollary 21.6.** Let \(A \in M_{nn}(\mathbb{C})\) be Hermitian \((A = A^t)\). Then there exists a unitary matrix \(U \in M_{n,n}(\mathbb{C})\) such that \(U^tAU = U^{-1}AU\) is a real diagonal matrix.