

A

Linear ALgebra.

Alec Zabel-Mena

June 29, 2021

Chapter 1

Vector Spaces and Modules.

1.1 Definitions and Examples

Definition. We call a nonempty set V a **vector space** over a field F , if given a binary operation $+: V \times V \rightarrow V$ called **vector addition** and an operation $\cdot: F \times V \rightarrow V$ called **scalar multiplication**, we have that $(V, +)$ forms an abelian group, and for all $v, w \in V$ and $\alpha, \beta \in F$:

[label=(0)]

1. $\alpha(v + w) = \alpha v + \alpha w$.
2. $(\alpha + \beta)v = \alpha v + \beta v$.
3. $\alpha(\beta v) = (\alpha\beta)v$.
4. $1v = v$, where 1 is the identity element of F under its multiplication.

Lemma 1.1.1. *Let V be a vector space over a field F . Then the operation $\cdot: F \times V \rightarrow V$ of scalar multiplication is a group homomorphism of V into V .*

Proof. Taking $\cdot: F \times V \rightarrow V$ by $(\alpha, v) \rightarrow \alpha v$, restrict \cdot to V , i.e. consider $\cdot|_V: V \rightarrow V$ by $v \rightarrow \alpha v$ for $\alpha \in F$. By (1) of the scalar multiplication rules, we get that $\cdot|_V$ is a homomorphism; which makes \cdot a homomorphism. ■

Example 1.1. [label=(0)]

1. Let F be a field and $F \subseteq K$ a field extension of F . Then K as a vector space over F with $+$ the usual addition of K and \cdot the multiplication of K restricted to F by the first part, i.e. the product $\cdot: v \rightarrow \alpha v$ with $\alpha \in F$.
2. Let F be a field and consider F^n the set of ordered n -tuples of elements of F , for some $n \in \mathbb{Z}^+$. Take $+: (v, w) \rightarrow v + w$ by $(v_1, \dots, v_n) + (w_1, \dots, w_n) = (v_1 + w_1, \dots, v_n + w_n)$, where $v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in F^n$, and $\cdot: (\alpha, v) \rightarrow \alpha v$ by $\alpha(v_1, \dots, v_n) = (\alpha v_1, \dots, \alpha v_n)$. Then F^n is a vector space over F .

3. Let F be any field and let $F[x]$ be the polynomial field over F . Take $+$ to be polynomial addition, and \cdot the multiplication of a constant in F by a polynomial in $F[x]$. Then $F[x]$ as a vector space over F .
4. Let $F[x]$ be the polynomial field over a field F and consider the set $P_n = \{f \in F[x] : \deg f < n\}$. Then P_n as a subset of $F[x]$ forms a vector space over F under the same operations $+$ and \cdot (this last example motivates the following definition).

Definition. Let V be a vector space over a field F . We say a subset $W \subseteq V$ is a **subspace** of V if W is also a vector space over F .

Lemma 1.1.2. Let V be a vector space over a field F , and let $W \subseteq V$ be a subspace of V . Then for all $w_1, w_2 \in W$ and $\alpha, \beta \in F$, $\alpha w_1 + \beta w_2 \in W$.

Proof. Since W is a vector space we have that $\alpha w_1, \beta w_2 \in W$; then by closure of vector addition, $\alpha w_1 + \beta w_2 \in W$. ■

Definition. Let U and V be vector spaces over a field F . We call a mapping $T : U \rightarrow V$ a **homomorphism** of U into V if:

$$T(u_1 + u_2) = T(u_1) + T(u_2). \quad T(\alpha u_1) = \alpha T(u_1).$$

for all $u_1, u_2 \in U$ and $\alpha \in F$. If T is 1-1 from U onto V , then we call T an **isomorphism** and we say U is **isomorphic** to V and write $U \simeq V$. We define the **kernal** of T to be $\ker T = \{u \in U : T(u) = 0\}$. We call the set of all homomorphism of U into V $\text{hom}(U, V)$.

Example 1.2. Let F be a field and consider the vector spaces F^n and P_n defined in examples (2) and (4). Then $P_n \simeq F^n$. Take the map $a_0 + a_1x + \cdots + a_nx^{n-1} \rightarrow (a_0, \dots, a_{n-1})$, which defines an isomorphism.

Lemma 1.1.3. If V is a vector space over a field F , then for all $\alpha \in F$ and $v \in V$:

$$\alpha 0 = 0. \quad 0v = 0. \quad (-\alpha)v = -(\alpha v). \quad \alpha v = 0 \text{ and } v \neq 0 \text{ implies } \alpha = 0.$$

Proof. [label=(0)]

$$\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0, \text{ hence } \alpha 0 = 0.$$

$$0v = (0 + 0)v = 0v + 0v, \text{ hence } 0v = 0.$$

3. We have $0 = 0v$, that is $0 = (\alpha + (-\alpha))v = \alpha v + (-\alpha)v$. Adding both sides by $-(\alpha v)$ we get the desired result.

4. If $\alpha \neq 0$ and $v \neq 0$, then $0 = \alpha^{-1}0 = \alpha^{-1}(\alpha v) = 1v = v$ which makes $v = 0$, which cannot happen. So $\alpha = 0$. ■

Lemma 1.1.4. Let V be a vector space over a field F and let $W \subseteq V$ be a subspace of V . Then V/W is a vector space over F where for $v_1 + W, v_2 + W \in V/W$ and $\alpha \in F$ we have:

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2 + W). \quad (\alpha v_1 + W) = \alpha(v_1 + W).$$

2. *Proof.* Since V as an abelian group, and W a subgroup of V under $+$, we get that V/W as the quotient group of V over W ; which as abelian since W as abelian.

Suppose now that for $v, v' \in V$ that $v + W = v' + W$, then for $\alpha \in F$ we have $\alpha(v + W) = \alpha(v' + W)$, and by hypotheses, we have $v - v' \in W$. Now since W as a subspace, $\alpha(v - v') \in W$ as well, so $\alpha v + W = \alpha v' + W$, so the product as well defined.

Now consider $v, v' \in W$ and $\alpha, \beta \in F$. By our product we have that $\alpha(v + w + W) = \alpha(v + w) + W = (\alpha v + \alpha w) + W = (\alpha v + W) + (\alpha w + W)$, $(\alpha + \beta)(v + W) = (\alpha + \beta)v + W = (\alpha v + \beta v) + W = \alpha(v + W) + \beta(v + W)$, $\alpha(\beta v + W) = \alpha\beta v + W = (\alpha\beta)v + W$, and finally, $1(v + w) = 1v + W = v + W$. Therefore V/W as a vector space over F . ■

Definition. Let V be a vector space over F and let $W \subseteq V$ be a subspace of V . We call the vector space formed by taking the quotient group of V over W , V/W the **quotient space** of V over W .

Theorem 1.1.5 (The First Homomorphism Theorem for Vector Spaces). *If $T : U \rightarrow V$ as a homomorphism of U onto V , and $W = \ker T$, then $V \simeq U/W$. If U as a vector space and $W \subseteq U$ as a subspace of U , then there as a homomorphism of U onto U/W .*

Proof. By the fundamental theorem of homomorphisms, we have that, as groups, $V \simeq U/W$. That there as a homomorphism from U onto U/W follows immediately. ■

Definition. Let V be a vector space over a field F and let $\{U_i\}_{i=1}^n$ be a collection of subspaces of V . We call V the **internal direct sum** of $\{U_i\}$ if every element of V can be written uniquely as a vector sum of elements of each U_i for $1 \leq i \leq n$; That as for $v \in V$, $v = u_1 + \dots + u_n$ as unique where $u_i \in U_i$.

Lemma 1.1.6. *Let $\{V_i\}_{i=1}^n$ be a collection of vector spaces over a field F and let $V = \prod_{i=1}^n V_i$ and define $+: V \times V \rightarrow V$ by $(v_1, \dots, v_n) + (v'_1, \dots, v'_n) = (v_1 + v'_1, \dots, v_n + v'_n)$ and define $\cdot: F \times V \rightarrow V$ by $\alpha(v_1, \dots, v_n) = (\alpha v_1, \dots, \alpha v_n)$. Then V as a vector space over F .*

Proof. Since V_i as a vector space for all $1 \leq i \leq n$, they are all abelian groups, hence V as closed under $+$, and inherits associativity, as well as commutativity. Now letting $0 = (0_1, \dots, 0_n)$, where 0_i as the identity of V_i , we get for any $v \in V$ that $v + 0 = 0 + v = v$, so 0 as the identity. Likewise for any $v \in V$, $-v = (-v_1, \dots, -v_n)$ serves as the inverse for v . So $(V, +)$ forms an abelian group.

Now by the axioms of scalar multiplication on each of the V_i , let $v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in V$ and $\alpha, \beta \in F$. We get $\alpha(v + w) = \alpha(v_1 + w_1, \dots, v_n + w_n) = (\alpha(v_1 + w_1), \dots, \alpha(v_n + w_n)) = (\alpha v_1 + \alpha w_1, \dots, \alpha v_n + \alpha w_n) = (\alpha v_1, \dots, \alpha v_n) + (\alpha w_1, \dots, \alpha w_n) = \alpha v + \alpha w$. We also get $(\alpha + \beta)v = ((\alpha + \beta)v_1, \dots, (\alpha + \beta)v_n) = (\alpha v_1 + \beta v_1, \dots, \alpha v_n + \beta v_n) = (\alpha v_1, \dots, \alpha v_n) + (\beta v_1, \dots, \beta v_n) = \alpha v + \beta v$. Through similar calculation, we get that $\alpha(\beta v) = (\alpha\beta)v$ and $1v = v$; which makes V into a vector space. ■

Definition. Let $\{V_i\}_{i=1}^n$ be a collection of vector spaces over a field F and let $V = \prod_{i=1}^n V_i$ and define $+: V \times V \rightarrow V$ by $(v_1, \dots, v_n) + (v'_1, \dots, v'_n) = (v_1 + v'_1, \dots, v_n + v'_n)$ and define $\cdot: F \times V \rightarrow V$ by $\alpha(v_1, \dots, v_n) = (\alpha v_1, \dots, \alpha v_n)$. We call V , as a vector space over F the **external direct sum** of $\{V_i\}$ and write $V = V_1 \oplus \dots \oplus V_n$, or $V = \bigoplus_{i=1}^n V_i$.

Theorem 1.1.7. *Let V be a vector space and let $\{U_i\}_{i=1}^n$ be a collection of subspaces of V . If V is the internal direct sum of $\{U_i\}$ then V is isomorphic to the external direct sum of $\{U_i\}$; that is: $V \simeq \bigoplus_{i=1}^n U_i$.*

Proof. Let $v \in V$. By hypothesis $v = u_1 + \cdots + u_n$ with $u_i \in U_i$ for $1 \leq i \leq n$, and it is a unique representation of v . Define then, the map $T : V \rightarrow \bigoplus_{i=1}^n U_i$ by the map $v = u_1 + \cdots + u_n \rightarrow (u_1, \dots, u_n)$. Since v has a unique representation by definition, T is well defined; moreover it is 1-1, as $(u_1, \dots, u_n) = (w_1, \dots, w_n)$ implies $u_i = w_i$ for all $1 \leq i \leq n$, hence $u_1 + \cdots + u_n = w_1 + \cdots + w_n$, and since this sum is unique, they both represent a vector $v \in V$. That T is onto follows directly from definition.

Finally, let $v, w \in V$, then $v = u_1 + \cdots + u_n$ and $w = w_1 + \cdots + w_n$. Hence $T(v + w) = T(u_1 + w_1 + \cdots + u_n + w_n) = (u_1 + w_1, \dots, u_n + w_n) = (u_1, \dots, u_n) + (w_1, \dots, w_n) = T(v) + T(w)$. Similarly, $T(\alpha v) = (\alpha v)$. ■

Remark. That V is the internal direct sum of $\{U_i\}$ and that $V \simeq U_1 \oplus \cdots \oplus U_n$ by the above theorem permits us to write $V = U_1 \oplus \cdots \oplus U_n$, or $V = \bigoplus_{i=1}^n U_i$.

1.2 Linear Independence and Bases.

Definition. If V is a vector space over a field F and give $v_1, \dots, v_n \in V$, then we call any element $v \in V$ of the form $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ for $\alpha_1, \dots, \alpha_n \in F$ a **linear combination** of v_1, \dots, v_n over F .

Definition. Let V be a vector space. We call the set of all linear combinations of finite sets of elements of a nonempty subset $S \subseteq V$ the **linear span** of S ; and we write $\text{span } S$.

Lemma 1.2.1. *If V is a vector space, and $S \subseteq V$ is nonempty, then $\text{span } S$ is a subspace of V .*

Proof. Since $\text{span } S$ is the set of all linear combinations of finite sets of elements of S , it is clear that $\text{span } S \subseteq V$. Now let $v, w \in \text{span } S$, then $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ and $w = \mu_1 w_1 + \cdots + \mu_m w_m$; where $\lambda_i, \mu_j \in F$ and $v_i, w_j \in S$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. Now consider $\alpha, \beta \in F$, then $\alpha v + \beta w = \alpha(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \beta(\mu_1 w_1 + \cdots + \mu_m w_m) = (\alpha \lambda_1) v_1 + \cdots + (\alpha \lambda_n) v_n + (\beta \mu_1) w_1 + \cdots + (\beta \mu_m) w_m$ which is a linear combination of the finite set $\{v_1, \dots, v_n, w_1, \dots, w_m\}$ of elements of S . Therefore $\alpha v + \beta w \in \text{span } S$. ■

Lemma 1.2.2. *If $S, T \subseteq V$, then:*

[label=(0)] $S \subseteq T$ implies $\text{span } S \subseteq \text{span } T$. $\text{span } (S \cup T) = \text{span } S + \text{span } T$. $\text{span } (\text{span } S) = \text{span } S$.

2. Proof. [label=(0)]

Let $v \in \text{span } S$, then $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, with $v_1, \dots, v_n \in S$. Since $S \subseteq T$, $v_1, \dots, v_n \in T$, hence $v \in \text{span } T$.

2. Let $v \in \text{span}(S \cup T)$, then $v = \lambda_1 v_1 + \cdots + \lambda_n v_n + \mu_1 w_1 + \cdots + \mu_m w_m = (\lambda_1 v_1 + \cdots + \lambda_n v_n) + (\mu_1 w_1 + \cdots + \mu_m w_m)$, where $v_i \in S$ and $w_j \in T$. Then $v \in \text{span } S + \text{span } T$.

Now for $v \in \text{span } S + \text{span } T$, $v = u + w$ with $u \in \text{span } S$ and $w \in \text{span } T$, hence v is a linear combination of the finite set $\{u_1, \dots, u_n, w_1, \dots, w_n\}$ of elements of $S \cup T$, hence $v \in \text{span}(S \cup T)$.

3. Clearly $\text{span } S \in \text{span}(\text{span } S)$. Suppose then that $v \in \text{span}(\text{span } S)$. Then $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ where $v_i = \beta_{i1} v_{i1} + \cdots + \beta_{im} v_{im}$ where $v_{ij} \in S$. Hence $v = ((\alpha_1 \beta_{11}) v_{11} + \cdots + (\alpha_1 \beta_{1m}) v_{1m}) + \cdots + (\alpha_n \beta_{n1}) v_{n1} + \cdots + (\alpha_n \beta_{nm}) v_{nm}$. Therefore $\text{span}(\text{span } S) \subseteq \text{span } S$. ■

Definition. We call a vector space V over a field F **finite dimensional** over F if there is a finite subset $S \subseteq V$ whose linear span is V ; that is $\text{span } S = V$.

Example 1.3. F^n is finite dimensional. Let $S = \{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}$. Then $\text{span } S = F^n$.

Definition. Let V be a vector space over a field F . We say that a set of $\{v_1, \dots, v_n\}$ of elements of V **linearly dependent** over F if there exist $\lambda_1, \dots, \lambda_n \in F$, not all 0 such that $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. We call $\{v_1, \dots, v_n\}$ **linearly independent** over F if it is not linearly dependent over F ; that is $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ implies $\lambda_1 = \cdots = \lambda_n = 0$.

Example 1.4. [label=(0)]

1. In F^3 , the vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ are linearly independent, where as $(1, 1, 0)$, $(3, 1, 3)$, $(5, 3, 3)$ are linearly dependent.
2. Consider the set \mathbb{C} of complex numbets as a vector space over \mathbb{R} . The vectors $1, i$ are linearly independent over \mathbb{R} since $i \notin \mathbb{R}$. However, $1, i$ is not linearly independent over \mathbb{C} , as $i^2 + 1 = 0$ by definition; where $\lambda_1 = i$ and $\lambda_2 = 1$.

Lemma 1.2.3. *If $v_1, \dots, v_n \in V$ are linearly independent, then every element in $\text{span } \{v_1, \dots, v_n\}$ can be represented unquely as a linear combination of v_1, \dots, v_n .*

Proof. Let $v \in \text{span } \{v_1, \dots, v_n\}$ such that $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ and $v = \mu_1 v_1 + \cdots + \mu_n v_n$. Then $\lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n$, then $(\lambda_1 - \mu_1) v_1 + \cdots + (\lambda_n - \mu_n) v_n = 0$. By linear independence, this implies that $\lambda_i - \mu_i = 0$, for all $1 \leq i \leq n$. Therefore v is uniquely represented. ■

Theorem 1.2.4. *If $v_1, \dots, v_n \in V$, then they are linearly independent, or v_k is a linear combination of v_1, \dots, v_{k-1} for $1 \leq k \leq n$.*

Proof. If v_1, \dots, v_n are linearly independent, then we are done. Now suppose that they are linearly dependent, then $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ for $\lambda_1, \dots, \lambda_n$ not all 0. Let k be the largest such integer for which $\lambda_k \neq 0$, and $\lambda_i = 0$ for all $k < i$. Then $\lambda_1 v_1 + \cdots + \lambda_n v_n = \lambda_1 v_1 + \cdots + \lambda_k v_k$ where $\lambda_1, \dots, \lambda_k$ are not all 0 for $1 \leq i \leq k$. Then we have that $v_k = (\lambda_k^{-1} \lambda_1) v_1 + \cdots + (\lambda_k^{-1} \lambda_{k-1}) v_{k-1}$ which is a linear combination of v_1, \dots, v_{k-1} . ■

Corollary. If $v_1, \dots, v_n \in V$ have W as a linear span, and if v_1, \dots, v_k are linearly independent, then there is a linearly independent subset of $\{v_1, \dots, v_n\}$ of the form $\{v_1, \dots, v_k, v_{i_1}, \dots, v_{i_r}\}$ which span W .

Proof. If v_1, \dots, v_n are linearly independent, then we are done. If not, let j be the smallest such integer for which v_j is a linear combination of its predecessors. Since v_1, \dots, v_k are linearly independent, we get $k < j$. then consider the set $S = \{v_1, \dots, v_n\} \setminus v_j = \{v_1, \dots, v_k, \dots, v_{j-1}, v_{j+1}, \dots, v_n\}$ which has $n - 1$ elements. Clearly, $\text{span } S \subseteq W$.

Now let $w \in W$, then $w = \lambda_1 v_1 + \dots + \lambda_n v_n$. Since v_j is a linear combination of v_1, \dots, v_{j-1} , we get that $w = \lambda'_1 v_1 + \dots + \lambda'_k v_k + \dots + \lambda'_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \dots + \lambda_n v_n$ which makes $W \subseteq \text{span } S$.

Now if we proceed by removing all vectors which are linear combinations of their predecessors, we get a set $\{v_1, \dots, v_k, v_{i_1}, \dots, v_{i_r}\}$ with span S ; by the preceding argument, we get again that $W \subseteq \text{span } S$. ■

Corollary. If V is a finite dimensional vector space, then there is a finite set of linearly independent vectors $\{v_1, \dots, v_n\}$ such that $\text{span } \{v_1, \dots, v_n\} = V$.

Proof. By definition, since V is finite dimensional, there is a finite set of vectors $\{u_1, \dots, u_m\}$ with linear span V . Then by the previous corollary, there is a subset $\{v_1, \dots, v_n\}$ of linearly independent vectors whose span is also V . ■

Definition. We call a subset S of a vector space V a **basis** if S consists of linearly independent vectors, and $\text{span } S = V$.

What the above corollary states, is that if V is a finite dimensional vector space, and u_1, \dots, u_m (not necessarily independent), $\text{span } V$, then u_1, \dots, u_m contain a basis of V .

Example 1.5. A basis need not be finite. Consider the polynomial field $F[x]$, the set $\{1, x, x^2, \dots, x_n, \dots\}$ forms a basis of $F[x]$. However, the set $\{1, x, x^2, \dots, x^n\}$ span the subspace P_n of $F[x]$.

Lemma 1.2.5. If V is a finite dimensional vector space, then $V \simeq F^n$ for some $n \in \mathbb{Z}^+$.

Proof. By lemma 1.2.3 and the above corollary, any $v \in V$ is the unique combination of basis elements v_1, \dots, v_n ; that is $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Now take the map $v \rightarrow (\lambda_1, \dots, \lambda_n)$ is well defined, 1-1 by linear independence and onto. Hence $V \simeq F^n$. ■

Remark. In fact if $\{v_1, \dots, v_n\}$ is a basis for V , then $|\{v_1, \dots, v_n\}| = n$.

Lemma 1.2.6. If $v_1, \dots, v_n \in V$ forms a basis, and $w_1, \dots, w_m \in V$ are linearly independent, then $m \leq n$. Moreover, the set $\{v_1, \dots, v_n\}$ is maximally linearly independent.

Proof. For any arbitrary vector $v \in V$, v is a linear combination of v_1, \dots, v_n by lemma 1.2.3, hence $\{v_1, \dots, v, v\}$ is linearly dependent. This makes $\{v_1, \dots, v_n\}$ maximally independent.

Now $w_m \in V$ is a linear combination of v_1, \dots, v_n ; moreover they span V by theorem 1.2.4, therefore, by the previous corollary there is a subset $\{w_m, v_{i_1}, \dots, v_{i_k}\}$ with $k \leq n - 1$ which is a basis of V .

Repeating by taking $w_{m-1}, w_m, \dots, v_{i_k}$; we get, eventually, a basis $\{w_{m-1}, w_m, \dots, v_{j_1}, \dots, v_{j_s}\}$, with $s \leq n-1$. Repeating then of the vectors w_2, \dots, w_{m-2} , we get a basis $\{w_2, \dots, w_{m-1}, \dots, v_\alpha\}$. Since w_1, \dots, w_m are linearly independent, w_1 is not a linear combination of the others, hence the basis contains some v . Now the basis above has $m-1$ w_i 's, at the cost of one $v \in V$, hence $m-1 \leq n-1$; thus $m \leq n$. ■

Corollary. *Any two bases have the same number of elements.*

Proof. Let $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_m\}$ be bases with n and m elements respectively. Since they are both linearly independent, by above we get $m \leq n$ and $n \leq m$. Therefore $m = n$. ■

Corollary. *$F^n \simeq F^m$ if and only if $n = m$.*

Proof. F^n has the basis $\{(1, 0, \dots, 0)_n, \dots, (0, 0, \dots, 1)_n\}$ and F^m has basis $\{(1, 0, \dots, 0)_m, \dots, (0, 0, \dots, 1)_m\}$, and any isomorphism must map a basis to a basis. ■

Corollary. *If V is finite dimensional over F , with $V \simeq F^n$ for some unique n , then any basis in V has exactly n elements.*

Definition. If V is a finite dimensional vector space over a field F , with a basis $\{v_1, \dots, v_n\}$ of n elements, we call the n **dimension** of V over F and write $\dim_F V = n$ or $\dim V = n$.

Example 1.6. [label=(0)]

1. $\dim F^n = n$.
2. $\dim_F P_n = n$, and $\dim F[x] = \infty$ (since $F[x]$ is infinite dimensional).
3. $\dim_{\mathbb{R}} \mathbb{C} = 2$.

Corollary. *If V and U are finite dimensional vector spaces over a field F , with $\dim_F V = \dim_F U$, then $V \simeq U$.*

Proof. $V \simeq F^n$ and $F^n \simeq U$. By transitivity, we get $V \simeq U$. ■

Lemma 1.2.7. *If V is a finite dimensional vector space over F and of $u_1, \dots, u_m \in V$ are linearly independent, then there exist $u_{m+1}, \dots, u_{m+r} \in V$ such that $\{u_1, \dots, u_m, u_{m+1}, u_{m+r}\}$ is a basis of V .*

Proof. By finite dimensionality, there is a basis v_1, \dots, v_n of V , which span V . Hence $\text{span}\{u_1, \dots, u_m, v_1, \dots, v_n\} = V$, therefore by theorem 1.2.4 there is a subset $\{u_1, \dots, u_m, v_{i_1}, \dots, v_{i_r}\}$ which is a basis of V . Now just map $v_{i_j} \rightarrow u_{m+j}$ for each $1 \leq j \leq r$. ■

Remark. This gives us a method for constructing bases of vector spaces.

Lemma 1.2.8. *If V is finite dimensional, and if W is a subspace of V , then W is also finite dimensional. Moreover $\dim W \leq \dim V$ and $\dim V/W = \dim V - \dim W$.*

Proof. If $\dim V = n$, then any set of $n + 1$ vectors in V is linearly dependent, by maximality, hence so is any set of $n + 1$ vectors in W . Then there exists a maximal set of linearly independent elements in W , w_1, \dots, w_m , with $m \leq n$. If $w \in W$, then w_1, \dots, w_m, w are linearly dependent with $\lambda_1 w_1 + \dots + \lambda_m w_m + \lambda w = 0$. Now $\lambda \neq 0$, for that would imply w_1, \dots, w_m, w linearly independent. Hence $w = \mu_1 w_1 + \dots + \mu_m w_m$ where $\mu_i = \lambda^{-1} \lambda_i$. Thus we get $w \in \text{span}\{w_1, \dots, w_m\}$, i.e. $W = \text{span}\{w_1, \dots, w_m\}$, thus w_1, \dots, w_m form a basis of W . Therefore $m = \dim W \leq \dim V = n$.

Now take $V \rightarrow V/W$ by $v_1, \dots, v_r \rightarrow v'_1, \dots, v'_r$. By lemma 1.2.7, if $\{w_1, \dots, w_m\}$ form a basis of W , then there exist v_{m+1}, \dots, v_{m+r} such that $\{w_1, \dots, w_m, v_{m+1}, v_{m+r}\}$ form a basis for V . That is, for any $v \in V$, $v = \lambda_1 w_1 + \dots + \lambda_m w_m + \mu_1 v_1 + \dots + \mu_r v_r$. Then we get that $v' = \mu_1 v'_1 + \dots + \mu_r v'_r$, hence $\text{span}\{v'_1, \dots, v'_r\} = V/W$. Now if $\gamma_1 v'_1 + \dots + \gamma_r v'_r = 0$, then $\gamma_1 v'_1 + \dots + \gamma_r v_r \in W$, making $\gamma_1 v'_1 + \dots + \gamma_r v_r = \lambda_1 w_1 + \dots + \lambda_m w_m$. By linear independence, $\gamma_i, \lambda_j = 0$ for all $1 \leq i \leq r$ and $1 \leq j \leq m$. This V/M has a basis of $r = \dim V - \dim W$ elements. Therefore $\dim V/W = \dim v - \dim W$. ■

Corollary. *If U and W are finite dimensional subspaces of a vector space V , then $U + W$ is finite dimensional, and $\dim(U + W) = \dim U + \dim W - \dim U \cap W$.*

Proof. We have $U + W/W \simeq U/U \cap W$. Hence we get that $\dim U + W/W = \dim(U + W) - \dim W = \dim U/U \cap W = \dim U - \dim U \cap W$. Then $\dim(U + W) = \dim W + \dim U - \dim U \cap W$. ■

1.3 Dual Spaces.

Lemma 1.3.1. *Let V and W be vector spaces over a field F . Then $\text{hom}(V, W)$ is a vector space over F .*

Proof. First, let $T, L \in \text{hom}(V, W)$, and $\alpha, \beta \in F$. Then $T + L(\alpha v + \beta u) = \alpha T(v) + \beta T(u) + \alpha L(v) + \beta L(u) = \alpha(T + L)(v) + \beta(T + L)(u)$, so $T + L \in \text{hom}(V, W)$. Since $+$ is just function addition, it is associative. Likewise, the zero map $0 : V \rightarrow W$ by $v \rightarrow 0$ and the map $-T : V \rightarrow W$ by $v \rightarrow -T(v)$ define the identity of $\text{hom}(V, W)$ and the inverse of T respectively. This makes $(\text{hom}(V, W), +)$ into a group. Now by the properties of homomorphisms, we also see that $\alpha(T + L) = \alpha T + \alpha L$, $(\alpha + \beta)T = \alpha T + \beta T$, $\alpha(\beta T) = (\alpha\beta)T$ and $T(1v) = 1T(v)$. This makes $\text{hom}(V, W)$ a vector space. ■

Lemma 1.3.2. *If $S, T \in \text{hom}(V, W)$ such that $S(v_i) = T(v_i)$ for all v_i in a basis $\{v_1, \dots, v_n\}$ of V , then $S = T$.*

Proof. Since $\{v_1, \dots, v_n\}$ is a basis of V , we have for every $v \in V$, $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ for unique $\lambda_1, \dots, \lambda_n \in F$. Then we get $S(v) = \lambda_1 S(v_1) + \dots + \lambda_n S(v_n) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = T(v)$. Thus $S(v) = T(v)$ for all $v \in V$. ■

Theorem 1.3.3. *If V and W are vector spaces with $\dim V = m$ and $\dim W = n$, then $\dim \text{hom}(V, W) = mn$.*

Proof. Let $\{v_1, \dots, v_m\}$ and $\{w_1, \dots, w_n\}$ be bases for V and W , respectively. Then for any $v \in V$, $v = \lambda_1 v_1 + \dots + \lambda_m v_m$ for unique $\lambda_1, \dots, \lambda_m \in F$. Now let $T_{ij} \in \text{hom}(V, W)$ be

defined such that $T_{ij}(v_i) = 0$ for $i \neq j$ and $T_{ij}(v_j) = \lambda_i w_j$; for $1 \leq i \leq m$ and $1 \leq j \leq n$. We see there are mn possible such T_{ij} . Now let $S \in \text{hom}(V, W)$, then $S(v_i) \in W$, hence $S = \mu_{11}w_1 + \cdots + \mu_{1n}w_n$ for unique $\mu_{11}, \dots, \mu_{1n} \in F$. Then $S(v_i) = \mu_{i1}w_1 + \cdots + \mu_{in}w_n$ for unique $\mu_{i1}, \dots, \mu_{in} \in F$. Now let $S_0 = \mu_{11}T_{11} + \cdots + \mu_{1n}T_{1n} + \cdots + \mu_{m1}T_{m1} + \cdots + \mu_{mn}T_{mn}$. Then $S_0(v_k) = (\mu_{11}T_{11} + \cdots + \mu_{1n}T_{1n} + \cdots + \mu_{m1}T_{m1} + \cdots + \mu_{mn}T_{mn})(v_k) = \mu_{11}T_{11}(v_k) + \cdots + \mu_{1n}T_{1n}(v_k) + \cdots + \mu_{m1}T_{m1}(v_k) + \cdots + \mu_{mn}T_{mn}(v_k)$. Since $T_{ij}(v_k) = 0$ for $i \neq k$ we get $S_0(v_k) = \alpha_{k1}w_1 + \cdots + \alpha_{kn}w_n$. So $S_0(v_k) = S(v_k)$ for the basis $\{v_1, \dots, v_m\}$ of V ; this makes $S_0 = S$.

Now since $S = S_0$ is arbitrary, and subsequently a linear combination of the T_{ij} , we get that $\text{span}\{T_{11}, \dots, T_{1n}, \dots, T_{m1}, \dots, T_{mn}\} = \text{hom}(V, W)$. Now suppose for $\beta_{11}, \dots, \beta_{1n}, \dots, \beta_{m1}, \dots, \beta_{mn} \in F$ that $\beta_{11}T_{11} + \cdots + \beta_{1n}T_{1n} + \cdots + \beta_{m1}T_{m1} + \cdots + \beta_{mn}T_{mn} = 0$. Then we get that $(\beta_{11}T_{11} + \cdots + \beta_{1n}T_{1n} + \cdots + \beta_{m1}T_{m1} + \cdots + \beta_{mn}T_{mn})(v_k) = \beta_{k1}w_1 + \cdots + \beta_{kn}w_n = 0$. Since $\{w_1, \dots, w_n\}$ is a basis of W , this makes $\beta_{kj} = 0$ for all $1 \leq n$. Thus $\{T_{11}, \dots, T_{1n}, \dots, T_{m1}, \dots, T_{mn}\}$ linearly independent, and hence a basis of $\text{hom}(V, W)$. Therefore, $\dim \text{hom}(V, W) = mn$. ■

Corollary. $\dim \text{hom}(V, V) = m^2$.

Corollary. $\dim \text{hom}(V, F) = m$.

Definition. Let V be a vector space over a field F . We call the vector space $\text{hom}(V, F)$ the **dual space** of V and denote it $\text{dual } V$. We call elements of $\text{dual } V$ **linear functionals** on V into F .

If V is an infinite dimensional vector space, the $\text{dual } V$ is very big and of no interest. In these cases, we use properties of other possible structures of $\text{dual } V$ to find a restricted subspace. If V is finite dimensional, then $\text{dual } V$ is finite and always defined.

Lemma 1.3.4. *If V is a finite dimensional vector space, and $v \neq 0 \in V$, then there is a linear functional $\hat{v} \in \text{dual } V$ such that $\hat{v}(v) \neq 0$.*

Proof. Let $\{v_1, \dots, v_n\}$ be a bases of V and let $\hat{v}_i \in \text{dual } V$ be defined by $\hat{v}_i(v_j) = 0$ whenever $i \neq j$ and $\hat{v}_i(v_j) = 1$ otherwise. Then if $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, $\hat{v}_i(v) = \lambda_i$. Then $\{\hat{v}_1, \dots, \hat{v}_n\}$ forms a basis of $\text{dual } V$. Now if $v \neq 0 \in V$, by lemma 1.2.7, we get a basis $v_1 = v, v_2, \dots, v_n$. Thence there is a linear functional $\hat{v}_1(v_1) = \hat{v}_1(v) = 1$. ■

Definition. Let V be a finite dimensional vector space with basis $\{v_1, \dots, v_n\}$. We define the **dual basis** of $\{v_1, \dots, v_n\}$ to be a basis of linear functionals $\{\hat{v}_1, \dots, \hat{v}_n\}$ of $\text{dual } V$ such that $\hat{v}_i(v_j) = 0$ wheberver $i \neq j$ and $\hat{v}_i(v_i) = 1$ otherwise.

Lemma 1.3.5. *If V is a finite dimensional vector space, and $T \in \text{dual } V$ such that $T(v)$ is fixed, then the map $\psi : v \rightarrow T_v$, where $T_v(T) = T(v)$ defines an isomorphism of V onto $\text{dual}(\text{dual } V)$.*

Proof. Let $v_0 \in V$. Let $T \in \text{dual } V$ be a linear functional such that $T(v_0)$ is fixed. Then $T(v_0)$ defines a linear functional of $\text{dual } V$ into F . Let $T_{v_0} : \text{dual } V \rightarrow F$ be defined by $T_{v_0}(T) = T(v_0)$, for any $T \in \text{dual } V$. Notice that for $T, L \in \text{dual } V$ and $\alpha, \beta \in F$, we have $T_{v_0}(\alpha T + \beta L) = \alpha T(v_0) + \beta L(v_0) = \alpha T_{v_0}(T) + \beta T_{v_0}(L)$, which makes $T_{v_0} \in \text{dual}(\text{dual } V)$.

Now given any $v \in V$, we can associate it with a $T_v \in \text{dual}(\text{dual } V)$. Now define $\psi : V \rightarrow \text{dual}(\text{dual } V)$ by $\psi : v \rightarrow T_v$. Then for $v, w \in V$ and $\alpha, \beta \in F$ we have $T_{\alpha v + \beta w}(T) = \alpha T(v) + \beta T(w) = \alpha T_v(T) + \beta T_w(T)$, so ψ is a homomorphism of V onto $\text{dual}(\text{dual } V)$; ψ is onto by definition.

Now let $v \in \ker \psi$. So $\psi(v) = 0$; that means $t_v(T) = T(v) = 0$ for all $T \in \text{dual } V$. However, by lemma 1.3.3, there must be a $T \in \text{dual } V$ for which $T(v) \neq 0$ when $v \neq 0$. Therefore, if $v \in \ker T$, it must be that $v = 0$, that is $\ker T = (0)$. Thus ψ is 1-1, which makes it an isomorphism. ■

Definition. Let W be a subspace of a vector space V . We denote the **annihilator** of W to be $A(W) = \{T \in \text{dual } V : T(v) = 0\}$.

Example 1.7. [label=(0)]

1. Let $W_1, W_2 \subseteq V$ be subspaces of a finite dimensional vector space. Let $T \in A(W_1 + W_2)$. Then $T(w) = 0$ for $w \in W_1 + W_2$, hence $w = w_1 + w_2$ where $w_i \in W_i$ for $1 \leq i \leq 2$. So we get $T(w_1) + T(w_2) = 0$ which makes either both $T(w_1), T(w_2)$ 0 or inverses of each other. IN either case, $T(w_1) + T(w_2) \in A(W_1) + A(W_2)$ or $T(w_1) + T(w_2) \in A(W_1) \cap A(W_2) \subseteq A(W_1) + A(W_2)$. So $A(W_1 + W_2) \subseteq A(W_1) + A(W_2)$. On the other hand we have $A(W_1), A(W_2) \subseteq A(W_1 + W_2)$, hence $A(W_1) + A(W_2) \subseteq A(W_1 + W_2)$. Hence we have $A(W_1 + W_2) = A(W_1) + A(W_2)$.
2. Similarly, if $T \in A(W_1 \cap W_2)$, then $T(w) = 0$ for $w \in W_1 \cap W_2$, making $T(w) \in A(W_1) \cap A(W_2)$. By similar reasoning to before, we also get that $A(W_1) \cap A(W_2) \subseteq A(W_1 \cap W_2)$. So we get $A(W_1 \cap W_2) = A(W_1) \cap A(W_2)$

Let $\tilde{T} \in \text{dual } W$ such that $\tilde{T}(w) = T(w)$ for any $w \in W$; where $T \in \text{dual } V$. Now define the map $\psi : \text{dual } V \rightarrow \text{dual } W$ by $\psi : T \rightarrow \tilde{T}$. Then we see that $A(W) = \ker \psi$, which makes it a subspace.

Lemma 1.3.6. *If $S \subseteq V$ is a subset of a finite dimensional vector space, then $A(S) \subseteq A(\text{span } S)$.*

Proof. Since $S \subseteq \text{span } S$, it is clear that $A(S) \subseteq A(\text{span } S)$. Now let $v \in \text{span } S$. Then $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ where $v_1, \dots, v_n \in S$. Let $\psi : \text{dual } V \rightarrow \text{dual } S$ by $T \rightarrow \tilde{T}$ where $\tilde{T}(s) = T(s)$ for all $s \in S$. Then $\psi(v) = \lambda_1 \psi(v_1) + \dots + \lambda_n \psi(v_n) = \psi(v) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n)$. Since $\ker \psi = A(S)$, and $T(v_i) = 0$ for all $v_i \in S$ for $1 \leq i \leq n$, we get $\psi(v) = 0$ hence $v \in A(S)$; which puts $A(\text{span } S) \subseteq A(S)$. ■

Theorem 1.3.7 (The Second Homomorphism Theorem for Vector Spaces). *If V is a finite dimensional vector space, and $W \subseteq V$ is a subspace of V , then $\text{dual } W \simeq \text{dual } V / A(W)$, and $\dim A(W) = \dim V - \dim W$.*

Proof. Consider again the map $\psi : \text{dual } V \rightarrow \text{dual } W$ by $T \rightarrow \tilde{T}$, where $\tilde{T}(w) = T(w)$ for all $w \in W$; and recalling above that $A(W) = \ker T$.

Let $h \in \text{dual } W$. By lemma 1.2.7, if $\{w_1, \dots, w_m\}$ is a basis of W , then there is a basis $\{w_1, \dots, w_m, v_1, \dots, v_r\}$; hence $\dim V = r + m$. Let W_1 be a subspace of V such that $\text{Span}\{v_1, \dots, v_r\} = W_1$. Then $V = W \oplus W_1$. Now if $h \in \text{dual } W$, let $f \in \text{dual } V$ be defined

by $f(v) = w$ where $v = w + w_1 \in W \oplus W_1$. By definition, we have that $f \in \text{dual } V$ and $f = h$. So $\psi(f) = h$ making ψ onto. Since $A(W) = \ker \psi$, by the first homomorphism theorem for vector spaces, we get $\text{dual } W \simeq \text{dual } V/A(W)$.

Moreover, we get $\dim \text{dual } W = \dim \text{dual } V/A(W) = \dim \text{dual } V - \dim A(W)$, and since $\dim \text{dual } V = \dim V$ and $\dim \text{dual } W = \dim W$; we get $\dim A(W) = \dim V - \dim W$. ■

Corollary. $A(A(W)) = W$.

Proof. Notice that $A(A(W)) \subseteq \text{dual}(\text{dual } V)$. Clearly, $W \subseteq A(A(W))$, for if $\psi(w) = T_w$ by $T_w(f) = f(w)$ and $T_w = 0$ for all $f \in A(W)$. Now by above we get $\dim A(A(W)) = \dim \text{dual } V - \dim A(W) = \dim V - (\dim V - \dim W) = \dim W$. This makes $W \simeq A(A(W))$; and since $W \subseteq A(A(W))$, we get $W = A(A(W))$. ■

Theorem 1.3.8. *The system of homogeneous linear equations:*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{1.1}$$

where $a_{ij} \in F$ is of rank r , then there are $n - r$ linearly independent solutions in F^n .

Proof. Consider the system described by equation 1.1, with $a_{ij} \in F$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. Let U be a subspace of m vectors generated by $\{(a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn})\}$. Consider the basis $\{(1, 0, \dots, 0), \dots, (0, 0, \dots, 1)\}$ of F^n and let $\{\hat{v}_1, \dots, \hat{v}_n\}$ be its dual basis. Then $T \in \text{dual } F^n$ has the form $T = x_1\hat{v}_1 + \cdots + x_n\hat{v}_n$, with $x_i \in F$ for $1 \leq i \leq n$.

Now for $(a_{11}, \dots, a_{1n}) \in U$, $T(a_{11}, \dots, a_{1n}) = (x_1\hat{v}_1 + \cdots + x_n\hat{v}_n)(a_{11}, \dots, a_{1n}) = a_{11}x_1 + \cdots + a_{1n}x_n$, since $\hat{v}_i(v_j) = 0$ for $i \neq j$. Conversely, every solution (x_1, \dots, x_n) gives an element of the form $x_1\hat{v}_1 + \cdots + x_n\hat{v}_n$ in $A(U)$. Therefore, the number of linearly independent solutions of equation 1.1 is $\dim A(U) = \dim V - \dim U = n - r$. ■

Corollary. *If $n > m$, then there is a solution (x_1, \dots, x_n) where not all x_i is 0.*

1.4 Inner Product Spaces.

Definition. We define a vector space V over \mathbb{C} to be an **inner product space** if there exists a binary operation $\langle, \rangle : V \times V \rightarrow \mathbb{C}$ such that for all $v, u, w \in V$ and $\alpha, \beta \in \mathbb{C}$:

- (1) $\langle u, v \rangle = \overline{\langle v, u \rangle}$.
- (2) $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ if and only if $u = 0$.
- (3) $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$.

Example 1.8. (1) In \mathbb{C}^n , let $u = (\alpha_1, \dots, \alpha_n)$ and $v = (\beta_1, \dots, \beta_n)$ and define $\langle u, v \rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i}$. Notice that $\sum \alpha_i \overline{\beta_i} = \sum_{i=1}^n \overline{\beta_i} \alpha_i = \overline{\sum \beta_i \overline{\alpha_i}}$; so $\langle u, v \rangle = \overline{\langle v, u \rangle}$. We also have that $\langle u, u \rangle \geq 0$ and is 0 only when $u = 0$. Moreover, if $w = (\gamma_1, \dots, \gamma_n)$ and $\alpha, \beta \in \mathbb{C}$, then $\langle \alpha u + \beta v, w \rangle = \sum (\alpha \alpha_i + \beta \beta_i) \overline{\gamma_i} = \alpha \sum \alpha_i \overline{\gamma_i} + \beta \sum \beta_i \overline{\gamma_i} = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$. So \langle, \rangle defines an inner product over \mathbb{C}^n .

- (2) Let $\mathbb{C}^{[0,1]}$ be the set of all complex valued functions continuous on the domain $[0, 1]$. If $f, g \in \mathbb{C}^{[0,1]}$, define $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)} dt$. Then \langle, \rangle defines an inner product over $\mathbb{C}^{[0,1]}$. Let $f, g, h \in \mathbb{C}^{[0,1]}$ and $\alpha, \beta \in \mathbb{C}$. We have then that $\langle f, g \rangle = \int f\overline{g} = \int \overline{\overline{f}g} = \overline{\int \overline{f}g} = \overline{\langle g, f \rangle}$. Moreover, $\int_0^1 f\overline{f} dt \geq 0$; now $\langle f, f \rangle = 0$ if $f = 0$. Now if $\int f\overline{f} dt = 0$, letting $f(t) = x(t) + iy(t)$, by the product of conjugates, and the sum rule, $x(t) = y(t) = 0$, i.e. $f = 0$. Again, by the rules of complex integrals, $\langle \alpha f + \beta g, h \rangle = \int (\alpha f + \beta g)\overline{h} = \alpha \int f\overline{h} + \beta \int g\overline{h}$.

Definition. Let V be an inner product space over \mathbb{C} . The **norm** of $v \in V$ is the map $\| \cdot \| : V \rightarrow \mathbb{R}$ by $\|v\| = \sqrt{\langle v, v \rangle}$.

Lemma 1.4.1. If V is an inner product space, with $u, v \in V$ and $\alpha, \beta \in \mathbb{C}$, then $\langle \alpha u + \beta v, \alpha u + \beta v \rangle = \alpha\overline{\alpha}\langle u, u \rangle + \alpha\overline{\beta}\langle u, v \rangle + \overline{\alpha}\beta\langle v, u \rangle + \beta\overline{\beta}\langle v, v \rangle$.

Proof. Take (3) on the inner product $\langle \alpha u + \beta v, \alpha u + \beta v \rangle$ to get: $\langle \alpha u + \beta v, \alpha u + \beta v \rangle = \alpha\langle u, \alpha u + \beta v \rangle + \beta\langle v, \alpha u + \beta v \rangle = \alpha\langle \alpha u + \beta v, u \rangle + \beta\langle \alpha u + \beta v, v \rangle = \alpha\overline{\alpha}\langle u, u \rangle + \alpha\overline{\beta}\langle u, v \rangle + \overline{\alpha}\beta\langle v, u \rangle + \beta\overline{\beta}\langle v, v \rangle$. ■

Corollary. $\|\alpha u\| = |\alpha|\|u\|$.

Proof. We have $\|\alpha u\|^2 = \langle \alpha u, \alpha u \rangle = \alpha\overline{\alpha}\langle u, u \rangle$. Since $\alpha\overline{\alpha} = |\alpha|^2$ we have $\|\alpha u\| = |\alpha|^2\|u\|^2$ which gives us the result. ■

Lemma 1.4.2. If $a, b \in \mathbb{R}$ such that $a > 0$ and $a\lambda^2 + 2b\lambda + c \geq 0$ for all $\lambda \in \mathbb{R}$, then $b^2 \leq ac$.

Proof. We complete the squares. $a\lambda^2 + 2b\lambda + c = \frac{1}{a}(a\lambda + b)^2 + (c - \frac{b^2}{a}) \geq 0$. Choosing $\lambda = -\frac{b}{a}$, we get $c - \frac{b^2}{a} \geq 0$. ■

Theorem 1.4.3 (The Cauchy-Schwarz Inequality). If V is an inner product space over \mathbb{C} with $u, v \in V$, then $|\langle u, v \rangle| \leq \|u\|\|v\|$.

Proof. If $\langle u, v \rangle \in V = \mathbb{R}$, and $u \neq 0$, then for any $\lambda \in \mathbb{R}$, $\langle u\lambda + v, u\lambda + v \rangle = \lambda^2\langle u, u \rangle + 2\lambda\langle u, v \rangle + \langle v, v \rangle \geq 0$. Letting $a = \langle u, u \rangle$, $b = \langle u, v \rangle$ and $c = \langle v, v \rangle$ we get $a\lambda^2 + 2b\lambda + c \geq 0$. By the above lemma, then $b^2 \leq ac$; i.e. $|\langle u, v \rangle|^2 \leq \|u\|^2\|v\|^2$.

Now take $\alpha = \langle u, u \rangle \in V \neq \mathbb{R}$. Then $\alpha \neq 0$. Now we observe that $\langle \frac{u}{\alpha}, v \rangle = \frac{1}{\alpha}\langle u, v \rangle = \frac{1}{\langle u, v \rangle}\langle u, v \rangle = 1$; so $\langle \frac{u}{\alpha}, v \rangle \in \mathbb{R}$. Then by above, we have $1 = |\langle \frac{u}{\alpha}, v \rangle| \leq \|\frac{u}{\alpha}\|\|v\| = \frac{1}{|\alpha|}\|u\|\|v\|$, that is $1 \leq \frac{\|u\|\|v\|}{|\alpha|}$; giving the desired result. ■

Example 1.9. (1) Let $V = \mathbb{C}^n$ and $\langle u, v \rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i}$ with $u = (\alpha_1, \dots, \alpha_n)$ and $v = (\beta_1, \dots, \beta_n)$. Then we have $|\sum \alpha_i \overline{\beta_i}| \leq \sum |\alpha_i|^2 \sum |\beta_i|^2$.

(2) If $V = \mathbb{C}^{[0,1]}$ with $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)} dt$, then we have $|\int_0^1 f\overline{g}| \leq \int_0^1 |f|^2 \int_0^1 |g|^2$.

Definition. If V is an inner product space, we say that $u, v \in V$ are **orthogonal** (or that u is **orthogonal** to v) if $\langle u, v \rangle = 0$.

Example 1.10. If u is orthogonal to v , then $\langle u, v \rangle = \overline{\langle v, u \rangle} = \overline{0} = 0$, making v orthogonal to u .

Definition. If V is an inner product space, and $W \subseteq V$ is a subspace of V we call the **orthogonal complement** of W the space $W^\perp = \{x \in V : \langle x, w \rangle = 0, \text{ for all } w \in W\}$.

Lemma 1.4.4. W^\perp is a subspace of V .

Proof. Clearly $W^\perp \subseteq V$. Moreover, let $a, b \in W^\perp$ and $\alpha, \beta \in \mathbb{C}$, then $\langle \alpha a + \beta b, w \rangle = \alpha \langle a, w \rangle + \beta \langle b, w \rangle = 0$, so $\alpha a + \beta b \in W^\perp$. ■

Example 1.11. Note that $W \cap W^\perp = \{x \in V : \langle x, w \rangle = 0\}$. If $w \in W^\perp$, then $\langle w, w \rangle = 0$ making $w = 0$, hence $W \cap W^\perp = \{0\}$.

Definition. We call a set of vectors $\{v_i\}_{i \in \mathbb{Z}^+}$ of an inner product space V **orthonormal** if:

- (1) $\langle v_i, v_i \rangle = 1$.
- (2) $\langle v_i, v_j \rangle = 0$ whenever $i \neq j$.

Lemma 1.4.5. If $\{v_i\}$ are a set of orthonormal vectors of V , then $\{v_i\}$ is also linearly independent. Moreover, if $\{v_i\}$ is finite and $w = \alpha_1 v_1 + \cdots + \alpha_n v_n$, then $\alpha_i = \langle w, v_i \rangle$ for each $1 \leq i \leq n$.

Proof. Suppose that $\alpha_1 v_1 + \cdots + \alpha_n v_n + \cdots = 0$, then $\langle \alpha_1 v_1 + \cdots + \alpha_n v_n + \cdots, v_i \rangle = \alpha_1 \langle v_1, v_i \rangle + \cdots + \alpha_n \langle v_n, v_i \rangle + \cdots = 0$. Since $\{v_i\}$ is orthonormal, we get that $\alpha_i = 0$ for each i , implying linear independence. Now if $\{v_i\}_{i=1}^n$ is finite, letting $w = \alpha_1 v_1 + \cdots + \alpha_n v_n$; by above we get that $\langle w, v_i \rangle = \alpha_i$ by orthonormality. ■

Lemma 1.4.6. If $\{v_1, \dots, v_n\}$ are orthonormal in V , and $w \in V$, then $u = w - \langle w, v_1 \rangle v_1 - \cdots - \langle w, v_n \rangle v_n$ is orthogonal to each v_i for $1 \leq i \leq n$.

Proof. $\langle u, v_i \rangle = \langle w - \langle w, v_1 \rangle v_1 - \cdots - \langle w, v_n \rangle v_n, v_i \rangle = \langle w, v_i \rangle - \langle w, v_i \rangle \langle v_1, v_i \rangle + \cdots + \langle w, v_n \rangle \langle v_n, v_i \rangle = 0$, making u orthogonal to v_i . ■

Theorem 1.4.7 (The Gram-Schmidt Orthogonalization Theorem). Let V be a finite dimensional inner product space. Then V has an orthonormal set as a basis.

Proof. Let $\dim V = n$ and let $\{v_1, \dots, v_n\}$ be a basis of V . Take w_1, \dots, w_n as follows: $v_1 | w_1$, $w_2 \in \text{span}\{w_1, v_2\}$ and $w_3 \in \text{span}\{w_1, w_2, v_3\}$; in general take $w_i \in \text{span}\{w_1, \dots, w_{i-1}, v_i\}$. Let $v_1 = \|v_1\| w_1$, then $\langle w_1, w_1 \rangle = \langle \frac{v_1}{\|v_1\|}, \frac{v_1}{\|v_1\|} \rangle = \frac{1}{\|v_1\|^2} \langle v_1, v_1 \rangle = 1$; hence $\|w_1\| = 1$. Now consider $\langle \alpha w_1 + v_2, w_1 \rangle = 0$. Then $\alpha \langle w_1, w_1 \rangle + \langle v_2, w_1 \rangle = 0$; since $\|w_1\| = 1$, then $\alpha = -\langle v_2, w_1 \rangle$. Now let $u_2 = -\langle v_2, w_1 \rangle w_1 + v_2$. u_2 is orthogonal to w_1 by lemma 1.4.6 and since v_1 and v_2 are linearly independent, so must w_1 and v_2 . So $u_2 \neq 0$. Now let $\|u_2\| w_2 = u_2$. We have then by above that, $\{w_1, w_2\}$ is orthonormal. Continuing along, suppose then that $\{w_1, \dots, w_i\}$ are orthonormal, where $\|u_i\| w_i = u_i$, and where $u_i = -\langle v_i, w_1 \rangle - \cdots - \langle v_i, w_{i-1} \rangle w_{i-1} + v_i$. Take $u_{i+1} = -\langle v_{i+1}, w_1 \rangle - \cdots - \langle v_{i+1}, w_i \rangle w_i + v_{i+1}$. By the above and lemma 1.4.5, w_1, \dots, w_i, v_{i+1} are linearly independent, so $u_{i+1} \neq 0$. Putting $\|u_{i+1}\| w_{i+1} = u_{i+1}$, clearly $\langle w_{i+1}, w_{i+1} \rangle = 1$. We also have, by the construction, that $\langle u_{i+1}, w_1 \rangle = \cdots = \langle u_{i+1}, w_i \rangle = 0$. So w_1, \dots, w_n are orthonormal.

Constructing $\{w_1, \dots, w_n\}$ from the basis $\{v_1, \dots, v_n\}$ this way gives an orthonormal set of n linearly independent vectors; i.e. a basis. ■

Corollary (Bessel's Inequality). *For all $v \in V$:*

$$\sum_{i=1}^m |\langle w_i, v \rangle|^2 \leq \|v\|^2. \quad (1.2)$$

Example 1.12. Let $V = \mathbb{R}_3[x]$ be the real field of all polynomials of $\deg < 3$. Define for $p(x), q(x) \in \mathbb{R}_3[x]$

$$\langle p, q \rangle = \int_{-1}^1 p(x)q(x) dx.$$

Now consider the basis $\{1, x, x^2\}$ of $\mathbb{R}_3[x]$. Take $w_1 = \frac{1}{\|1\|} = \frac{1}{\sqrt{\int_{-1}^1 dx}} = \frac{1}{\sqrt{2}}$. Take $u_2 = -\langle x, w_1 \rangle w_1 + x = -\frac{\langle x, w_1 \rangle}{\sqrt{2}} + x = x \neq 0$. Now take $w_2 = \frac{u_2}{\|u_2\|} = \frac{x}{\sqrt{\int_{-1}^1 x^2 dx}} = \frac{\sqrt{3}}{2}x$. Taking $u_3 = -\langle x^2, w_1 \rangle w_1 - \langle x^2, w_2 \rangle w_2 + x^2 = -\frac{1}{3} + x^2 \neq 0$; so taking $\frac{u_2}{\|u_3\|} = \frac{-\frac{1}{3} + x^2}{\sqrt{\int_{-1}^1 (-\frac{1}{3} + x^2) dx}} = \frac{\sqrt{10}}{4}(-1 + 3x^2)$, we get the orthonormal basis $\{x, -\frac{1}{3} + x^2, \frac{\sqrt{10}}{4}(-1 + 3x^2)\}$.

Theorem 1.4.8. *If V is a finite dimensional inner product space, and if $W \subseteq V$ is a subspace of V , then $V = W \oplus W^\perp$.*

Proof. Since $W \subseteq V$ is a subspace of V , W inherits the inner product of V (restrict \langle, \rangle to $W \times W$); similarly, W^\perp also inherits the inner product. By the Gram-Schmidt orthogonalization theorem, there is an orthonormal set of vectors $\{w_1, \dots, w_r\}$ which is a basis of W . Now if $v \in V$, by lemma 1.4.6 take $v_0 = v - \langle v, w_1 \rangle w_1 - \dots - \langle v, w_r \rangle w_r$ and $\langle v_0, w_i \rangle = 0$ for each $1 \leq i \leq r$. Then $v = v_0 + \langle v, w_1 \rangle w_1 + \dots + \langle v, w_r \rangle w_r \in W + W^\perp$. Since $W \cap W^\perp = 0$, we get $V = W \oplus W^\perp$. ■

Corollary. $(W^\perp)^\perp = W$.

Proof. If $w \in W$, then for any $u \in W$, $\langle u, w \rangle = 0$, hence $W \subseteq (W^\perp)^\perp$. Now $V = W^\perp \oplus (W^\perp)^\perp$ and we have $\dim W = \dim (W^\perp)^\perp$, which gives us $W = (W^\perp)^\perp$. ■

1.5 Modules.

Definition. Let R be a ring. We say a nonempty set M is a **left module** over R (or a **left R -module**) if there are operations $+: M \times M \rightarrow M$ and $\cdot: R \times M \rightarrow M$ such that $(M, +)$ is an abelian group, and for any $r, s \in R$ and $a, b \in M$:

- (1) $r(a + b) = ra + rb$.
- (2) $r(sa) = (rs)a$.
- (3) $(r + s)a = ra + sa$.

Similarly, we call M a **right module** (or **right R -module**) over R if $(a + b)r = ar + br$, $(as)r = a(sr)$, and $a(r + s) = ar + as$. We call M **unital** if R has a unit element, and $1m = m$ for all $m \in M$.

We focus on left modules.

Example 1.13. (1) All vector spaces are unital left modules over any field F .

- (2) Let G be a group together with an arbitrary operation $+$ and define an action $\cdot : \mathbb{Z} \times G \rightarrow G$ by $(n, a) \rightarrow na \in G$. Then the properties of exponents in groups gives $r(a + b) = ra + rb$, $r(sa) = (rs)a$, and $(r + s)a = ra + sa$. This makes every group a left \mathbb{Z} -module.
- (3) Let R be a ring, and let M be a left ideal of R . Take $r, m \rightarrow rm$. Since M is an ideal, $rm \in M$, and by the multiplicative associative, and distributive laws, M is a left R -module.
- (4) Any ring R is a left (and right) module over itself.
- (5) Let R be a ring, and (λ) a left ideal of R . Consider the quotient ring $R/(\lambda)$. define $+$ by $(a + \lambda) + (b + \lambda) = (a + b) + \lambda$ and $r(a + \lambda) = ra + \lambda$. Clearly these operations are well defined, and $(R/(\lambda), +)$ forms a group; moreover, $(a + \lambda) + (b + \lambda) = (a + b) + \lambda = (b + \lambda) + (a + \lambda)$, so $R/(\lambda)$ is abelian under $+$. Now notice that $r(a + b + \lambda) = r(a + b) + \lambda = ra + rb + \lambda = (ra + \lambda) + (rb + \lambda) = r(a + \lambda) + r(b + \lambda)$, $r(sa + \lambda) = rsa + \lambda = rs(a + \lambda)$, and $(r + s)(a + \lambda) = (r + s)a + \lambda = ra + rs + \lambda = r(a + \lambda) + s(a + \lambda)$. This makes $R/(\lambda)$ a left R -module. We call this module the **left quotient module** of R by (λ) .

Definition. Let M be an R -module (left or right) and $A \subseteq M$, we call A a **submodule** of M is $A \leq M$ and whenever $r \in R$ and $a \in A$, $ra \in A$, or $ar \in A$.

Definition. If M is an R -module with a collection of submodules $\{M_i\}_{i=1}^s$. We call M the **direct sum** of $\{M_i\}$ if for every $m \in M$, there are uniquely determined $m_i \in M_i$ for $1 \leq i \leq s$, such that $m = m_1 + \cdots + m_s$. We write $M = M_1 \oplus \cdots \oplus M_s$, or $M = \bigoplus_{i=1}^s M_i$.

Definition. An R -module is **cyclic** if there exists $m_0 \in M$ such that $m = rm_0$ (or $m = m_0r$) for all $m \in M$ and some $r \in R$.

Definition. We say an R -module is **finitely generated** if there exists $a_1, \dots, a_n \in M$ such that for every $m \in M$, $m = r_1a_1 + \cdots + r_na_n$ (or $m = a_1r_1 + \cdots + a_nr_n$) for $r_1, \dots, r_n \in R$. We call $\{a_i\}_{i=1}^n$ the **generating set**; and we call it a **minimal generating set** if $\{a_i\} \setminus a_j$ does not generate M , for $1 \leq i, j \leq n$. We call the size of a minimal generating set the **rank** of M and denote it $\text{rank } M$.

Most of the definitions are stated for both left and right R -modules. However, we consider the following theorems only for left R -modules.

Theorem 1.5.1 (The Fundamental Theorem on Finite Modules). *Let R be a Euclidean domain; then any finitely generated module M is the direct sum of a finite number of cyclic submodules.*

Proof. By definition, if M is finitely generated, then there are $a_1, \dots, a_n \in M$ for which every element of M is of the form $r_1a_1 + \cdots + r_na_n$, for $r_1, \dots, r_n \in R$. If M is indeed a direct sum of a finite collection of cyclic submodules, then each r_ia_i is uniquely determined.

By induction in the rank of M ; if $\text{rank } M = 1$, then M is generated by a single element m_0 . That is, for some $r \in R$, every element of M has the form rm_0 ; this makes M cyclic by definition, and hence the direct sum of itself.

Now suppose for $\text{rank } M = q$, that $M = \bigoplus_{i=1}^q M_i$, where M_i is a cyclic submodule. Suppose now that $\text{rank } M = q + 1$ and let $\{a_i\}_{i=1}^{q+1}$ be a minimal generating set for M . Then there are $r_1, \dots, r_{q+1} \in R$ for which $r_1 a_1 + \dots + r_{q+1} a_{q+1} = 0$ (the identity of $(M, +)$). If $r_1 a_1 = \dots = r_{q+1} a_{q+1} = 0$, then $M = \bigoplus_{i=1}^{q+1} M_i$ and we are done.

Now suppose that not all the $r_i a_i$ are 0. Since R is a Euclidean domain, with a degree function \deg , there is an element s_1 of minimum degree occurring as a coefficient in a relation of $\{a_i\}_{i=1}^{q+1}$. Then $s_1 a_1 + \dots + s_{q+1} a_{q+1} = 0$, where $\deg s_1 \leq \deg s_i$ for all $1 < i \leq q + 1$. Now if $r_1 a_1 + \dots + r_{q+1} a_{q+1} = 0$, then $s_1 | r_1$, for if $r_1 = m s_1 + t$ with $t = 0$ or $\deg t < \deg s_1$, then taking $(m s_1) a_1 + \dots + (m s_{q+1}) a_{q+1} = 0$ and subtracting $r_1 a_1 + \dots + r_{q+1} a_{q+1}$, we get $t a_1 + (r_2 - m s_2) a_2 + \dots + (r_{q+1} - m s_{q+1}) a_{q+1} = 0$, since $\deg t < \deg s_1$, and s_1 has minimum such degree, this makes $t = 0$.

We also have $s_1 | s_i$ for all $1 \leq i \leq q + 1$ (obviously $s_1 | s_1$). For, suppose that $s_1 \nmid s_i$ for all $1 < i \leq q + 1$, then $s_2 = m_2 s_1 + t$ with $\deg t < \deg s_1$. Now $a'_1 = a_1 + m_2 a_2 + \dots + m_{q+1} a_{q+1}$, $m_2 a_2, \dots, m_{q+1} a_{q+1}$ also generate M ; however, $s_1 a'_1 + t a_2 + s_3 a_3 + \dots + s_{q+1} a_{q+1} = 0$, so t is a coefficient occurring in some relation of $\{a_i\}$. But $\deg t < \deg s_1$, which contradicts that s_1 has minimum such degree, so $t = 0$ and hence $s_1 | s_2$. Similarly we get $s_1 | s_i$.

Now consider $a_1^* = a_1 + m_2 a_2 + \dots + m_{q+1} a_{q+1}$, a_2, \dots, a_{q+1} . They generate M ; moreover $s_1 a_1^* = s_1 a_1 + (s_1 m_2) a_2 + \dots + (s_1 m_{q+1}) a_{q+1} = s_1 a_1 + \dots + s_{q+1} a_{q+1} = 0$. If $r a_1^* = r a_1 + (r m_2) a_2 + \dots + (r m_{q+1}) a_{q+1} = 0$, then there is some relation on $\{a_i\}$ for which a_1 has coefficient r , i.e. $s_1 | r$, so $r_1 a_1^* = 0$. Letting M_1 the cyclic submodule generated by a_1^* , and M_2 the submodule finitely generated by $\{a_i\}_{i=2}^{q+1}$, we have $M_1 \cap M_2 = 0$ and $M = M_1 + M_2$; hence $M = M_1 \oplus M_2$. Now by hypothesis, we get $M_2 = M'_2 \oplus M_3 \oplus \dots \oplus M_{q+1}$, each of which is a cyclic submodule of M ; which completes the proof. ■

Corollary. Any finite abelian group is the direct product of cyclic groups.

Proof. Consider the finite abelian group G as a \mathbb{Z} -module. ■

Theorem 1.5.2. The number of non-isomorphic finite abelian groups of order p^n is $p(n)$; where $p(n)$ is the number of partitions of n .

Proof. Let G be a finite abelian group of order $\text{ord } G = p^n$, for $n, p \in \mathbb{Z}^+$ and p prime. By the corollary to the fundamental theorem, $G = G_1 \times \dots \times G_k$, where G_i is a cyclic group of order $\text{ord } G_i = p^{n_i}$, where $n_k \leq \dots \leq n_1 \leq n$. Then

$$G_1 \times G_2 = \frac{G_1 \text{ ord } G_2}{(G_1 \cap G_2)}.$$

Since $G_1 \times G_2$ is a direct product, $\text{ord } (G_1 \cap G_2) = (e)$, so $\text{ord } G_1 \times G_2 = \text{ord } G_1 \text{ ord } G_2 = p^{n_1} p^{n_2} = p^{n_1+n_2}$. Continuing this way we get $p^n = \text{ord } G = \text{ord } (G_1 \times \dots \times G_k) = p^{n_1+\dots+n_k}$, hence $n = n_1 + \dots + n_k$ making $\{n_i\}_{i=1}^k$ a partition of n .

On the other hand, if $\{n_i\}_{i=1}^k$ is a partition of n , then we construct G of $\text{ord } = p^n$ as follows: for $1 \leq i \leq k$, let G_i be a cyclic group of order $\text{ord } G_i = p^{n_i}$ and let G be the external direct product of $\{G_i\}_{i=1}^k$. G is an abelian group of order p^n . Hence for each partition of

n , there is a abelian group of order p^n , if we take p^{n_i} for $1 \leq i \leq k$, characterizing G up to isomorphism, we get a 1 – 1 correspondence of non-isomorphic finite abelian groups of order p^n and partitions of n . ■

Corollary. *The number of non-isomorphic finite abelian groups of order $p_1^{n_1} \dots p_k^{n_k}$ for p_i distinct primes is $p(n_1) \dots p(n_k)$.*