Chapter 6. Vector Spaces and Modules

6.1 Definition: Let R be a ring. A **module** over R (or an R-**module**) is a set U together with an element $0 \in U$ and two operations $+ : U \times U \to U$ and $* : R \times U \to U$, where we write $+(x, y)$ as $x + y$ and $*(t, x)$ as $t * x$, $t \cdot x$ or as tx, such that

(1) $+$ is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in U$,
(2) $+$ is: $x + y = y + x$ for all $x, y \in U$,
(3) 0 is an additive identity: $x + 0 = x$ for all $x \in U$,
(4) every element has an additive inverse: for all $x \in U$ there exists $y \in U$ with $x + y = 0$,
(5) $*$ is associative: $(st)x = s(tx)$ for all $s, t \in R$ and all $x \in U$,
(6) 1 is a multiplicative identity: $1 \cdot x = x$ for all $x \in U$,
(7) $*$ is distributive over $+$ in R: $(s + t)x = sx + tx$ for all $s, t \in R$ and all $x \in U$, and
(8) $*$ is distributive over $+$ in U: $t(x + y) = tx + ty$ for all $t \in R$ and all $x, y \in U$.

When F is a field, a module over F is also called a **vector space** over F.

6.2 Note: In an R-module U, the zero element is unique, the additive inverse of a given element $x \in U$ and we denote it by $-x$, and we have additive cancellation.

6.3 Definition: Let R be a ring and let W be an R-module. A **submodule** of W over R is a subset $U \subseteq W$ which is also an R-module using the (restrictions of) the same operations used in W. Note that for a subset $U \subseteq W$, the operations on W restrict to well-defined operations on U if and only if

(1) U is closed under $+$: for all $x, y \in U$ we have $x + y \in U$, and
(2) U is closed under $*$: for all $t \in R$ and all $x \in U$ we have $tx \in U$.

When the operations do restrict as above, U is a submodule of W if and only if

(3) U contains the zero element: $0 \in U$, and
(4) U is closed under negation: for all $x \in U$ we have $-x \in U$.

When F is a field and W is a vector space over F, a submodule of W is also called a **subspace** of W over F.

6.4 Example: Let R be a ring. Then R^n, R^ω and R^∞ are all R-modules, where

$$\begin{aligned}
R^n &= \big\{f : \{1, 2, \cdots, n\} \to R\big\} = \big\{(a_1, a_2, \cdots, a_n)\big|\text{each } a_k \in R\big\},\\
R^\omega &= \big\{f : \mathbf{Z}^+ \to R\big\} = \big\{(a_1, a_2, a_3, \cdots)\big|\text{ each } a_k \in R\big\} \text{ and}\\
R^\infty &= \big\{f : \mathbf{Z}^+ \to R\big|f(k) = 0 \text{ for all but finitely many } k \in \mathbf{Z}^+\big\}\\
&= \big\{(a_1, a_2, a_3, \cdots)\big|\text{ each } a_k \in R \text{ with } a_k = 0 \text{ for all but finitely many } k \in \mathbf{Z}^+\big\}
\end{aligned}$$

6.5 Example: When R is a ring, $M_{m\times n}(R)$ is an R-modulue.

6.6 Example: For two sets A and B, we denote the set of all functions $f : A \to B$ by B^A or by $\text{Func}(A, B)$. When A is a set and R is a ring, we define operations on $\text{Func}(A, R)$ by $(tf)(x) = t\, f(x)$, $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for all $x \in X$. The set $\text{Func}(A, R)$ is a ring under addition and multiplication and also an R-module under addition and multiplication by $t \in R$.

6.7 Example: Let R be a ring. Recall that a **polynomial** ,with coefficients in R, is an expression of the form $f(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ where $n \in \mathbf{N}$ and each $c_k \in R$. We denote the set of all such polynomials by $R[x]$ or by $P(R)$. We consider two polynomials to be equal when their coefficients are all equal, so for $f(x) = \sum_{k=0}^{n} a_k x^k$ and $g(x) = \sum_{k=0}^{n} b_k x^k$ we have $f = g$ when $a_k = b_k$ for all k. In particular, we have $f = 0$ when $a_k = 0$ for all indices k. When $0 \neq f \in R[x]$, the **degree** of f, denoted by $\deg(f)$, is the largest $n \in \mathbf{N}$ for which $c_n = 0$. The degree of the zero polynomial is $-\infty$. For $n \in \mathbf{N}$, we denote the set of all polynomials of degree at most n by $P_n(R)$. A (formal) **power series**, with coefficients in R, is an expression of the form $f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 + c_2 x^2 + \cdots$. We denote the set of all such power series by $R[[x]]$. Thus we have

$$P_n(R) = \Big\{ f(x) = \sum_{k=0}^{n} c_k x^k \Big| \text{ each } c_k \in R \Big\},$$

$$R[x] = P(R) = \bigcup_{n \in \mathbf{N}} P_n(R) = \Big\{ f(x) = \sum_{k=0}^{n} c_k x^k \Big| \, n \in \mathbf{N} \text{ and each } c_k \in R \Big\} \text{ and}$$

$$R[[x]] = \Big\{ \sum_{k=0}^{\infty} c_k x^k \Big| \text{ each } c_k \in R \Big\}.$$

We define operations on these sets as follows. Given $f(x) = \sum_{ikge0} a_k x^k$ and $g(x) = \sum_{k \geq 0} b_k x^k$ (where the sums may be finite or infinite) we define tf, $f + g$ and fg by

$$(tf)(x) = \sum_{k \geq o} t\, a_k x^k \ , \ (f+g)(x) = \sum_{ikge0} (a_k + b_k) x^k \ , \text{ and}$$

$$(fg)(x) = \sum_{i,j \geq 0} (a_i b_j) x^{i+j} = \sum_{k \geq 0} c_k x^k \ \text{ with } c_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

The sets $P_n(R)$, $R[x] = P(R)$ and $R[[x]]$ are all rings under addition and multiplication, and they are all R-modules under addition and multiplication by elements $t \in R$.

6.8 Definition: Let R be a ring. An **algebra** over R (or an **R-algebra**) is a set U with an element $0 \in U$ together with three operations $+ : U \times U \to U$, $* : U \times U \to U$ and $* : R \times U \to U$ such that U is an abelian group under $+$, such that the two multiplication operations satisfy $(xy)z = x(yz)$, $(x+y)z = xz + yz$, $x(y+z) = xy + xz$ and $t(xy) = (tx)y$ for all $t \in R$ and all $x, y, z \in U$.

6.9 Example: When R is a ring and A is a set, R^n, R^∞, R^ω, $M_n(R)$, $\mathrm{Func}(A, R)$, $P_n(R)$, $R[x]$, $R[[x]]$ and R^A are all R-algebras using their usual operations.

6.10 Theorem: *Let R be a ring, and let W be an R-module. Let A be a set, and for each $\alpha \in A$ let U_α be a submodule of W over R. Then $\bigcap_{\alpha \in A} U_\alpha$ is an R-module.*

Proof: I may include a proof later.

6.11 Definition: Let R be a ring, let W be an R-module, and let $S \subseteq U$. A **linear combination** of the set S (or of the elements in S) (over R) is an element $w \in W$ of the form $w = \sum_{i=1}^{n} t_i u_i$ for some $n \in \mathbf{N}$, $t_i \in R$ and $u_i \in U$ (we allow the case $n = 0$ to include the empty sum, which we take to be equal to 0). The **span** of S (over R), denoted by $\mathrm{Span}\,(S)$ or by $\mathrm{Span}_R(S)$, is the set of all such linear combinations, that is

$$\mathrm{Span}\,(S) = \Big\{ \sum_{i=1}^{n} t_i u_i \Big| n \in \mathbf{N}, t_i \in R, u_i \in S \Big\}.$$

When $U = \mathrm{Span}\,(S)$, we also say that S **spans** U or that U is **spanned by** S.

The submodule of W **generated by** S, denoted by $\langle S \rangle$, is the smallest submodule of W containing S, that is

$$\langle S \rangle = \bigcap \big\{ U \subseteq W \big| U \text{ is a submodule of } W \text{ with } S \subseteq U \big\}.$$

6.12 Theorem: *Let R be a ring, let W be an R-module, and let $S \subseteq W$. Then*

$$\langle S \rangle = \mathrm{Span}\,(S).$$

Proof: I may include a proof later.

6.13 Definition: Let R be a ring, let U be an R-module, and let $S \subseteq U$. We say that S is **linearly independent** (over R) when for all $n \in \mathbf{Z}^+$, for all $t_i \in R$, and for all distinct elements $u_i \in S$, if $\sum_{i=1}^{n} t_i u_i = 0$ then $t_i = 0$ for all indices i. Otherwise we say that S is **linearly dependent**. We say that S is a **basis** for U when S spans U and S is linearly independent. We say that U is a **free** R-module when it has a basis.

6.14 Note: A set S is a basis for an R-module U if and only if every element $x \in U$ can be written uniquely (up to the order of the terms in the sum) in the form $x = \sum_{i=1}^{n} t_i u_i$ where $n \in \mathbf{N}$, $0 \neq t_i \in F$ and $u_1, u_2, \cdots, u_n$ are distinct elements in S.

6.15 Example: For $e_k \in R^n$ given by $(e_k)_i = \delta_{k,i}$, the set $\mathcal{S} = \{e_1, e_2, \cdots, e_n\}$ is a basis for R^n, and we call it the **standard basis** for R^n. For $e_k \in R^\infty$ given by $(e_k)_i = \delta_{k,i}$, the set $\mathcal{S} = \{e_1, e_2, e_3, \cdots\}$ is a basis for R^∞, which we call the **standard basis** for R^∞. It is not immediately obvious whether the R-module R^ω has a basis.

6.16 Example: For each $k, l \in \{1, 2, \cdots, n\}$, let $E_{k,l} \in M_n(R)$ denote the matrix with $(E_{k,l})_{i,j} = \delta_{k,i}\delta_{l,j}$ (so the (k, l) entry is equal to 1 and all other entries are equal to 0). Then the set $\mathcal{S} = \big\{ E_{k,l} \big| k, l \in \{1, 2, \cdots, n\} \big\}$ is a basis for $M_{m \times n}(R)$, which we call the **standard basis** for $M_{m \times n}(R)$.

6.17 Example: For a ring R, the set $\mathcal{S} = \mathcal{S}_n = \big\{1, x, x^2, \cdots, x^n\big\}$ is the **standard basis** for $P_n(R)$, and the set $\mathcal{S} = \big\{1, x, x^2, x^3, \cdots\big\}$ is the **standard basis** for $P_n(R) = R[x]$. It is not immediately obvious whether the R-module $R[[x]]$ has a basis.

Ordered Bases and the Coordinate Map

6.18 Definition: Let R be a ring and let W be an R-module. Let $\mathcal{A} = (u_1, u_2, \cdots, u_n)$ be an ordered n-tuple of elements in W. A **linear combination** of $\mathcal{A}$ (over R) is an element $x \in W$ of the form $x = \sum_{i=1}^{n} t_i u_i$ with each $t_i \in R$. The **span** of U (over R) is the set

$$\text{Span}\,(\mathcal{A}) = \Big\{ \sum_{i=1}^{n} t_i u_i \Big| t \in R^n \Big\}.$$

When $U = \text{Span}\,(\mathcal{A})$ we say that $\mathcal{A}$ **spans** U or that U is **spanned by** $\mathcal{A}$. We say that $\mathcal{A}$ is **linearly independent** (over R) when for all $t \in R^n$, if $\sum_{i=1}^{n} t_i u_i = 0$ then $t = 0$. We say that $\mathcal{A}$ is an **ordered basis** for U when $\mathcal{A}$ is linearly independent and spans U.

6.19 Note: Let R be a ring, let W be an R-module, and let $u_1, u_2, \cdots, u_n \in W$. Then the span of the n-tuple $(u_1, u_2, \cdots, u_n)$ is equal to the span of the set $\{u_1, u_2, \cdots, u_n\}$, and the n-tuple $(u_1, u_2, \cdots, u_n)$ is linearly independent if and only of the elements $u_1, u_2, \cdots, u_n$ are distinct and the set $\{u_1, u_2, \cdots, u_n\}$ is linearly independent. For a submodule $U \subseteq W$, the n-tuple $(u_1, \cdots, u_n)$ is a basis for U if and only if the elements u_i are distinct and the set $\{u_1, \cdots, u_n\}$ is a basis for U.

6.20 Definition: Let R be a ring, let U be a free R-module, and let $\mathcal{A} = (u_1, u_2, \cdots, u_n)$ be an ordered basis for U. Given an element $x \in U$, since $\mathcal{A}$ spans U we can write x as a linear combination $x = \sum_{i=1}^{n} t_i u_i$ with $t \in R^n$, and since $\mathcal{A}$ is linearly independent the element $t \in R^n$ is unique. We denote the unique element $t \in R^n$ such that $x = \sum_{i=1}^{n} t_i u_i$ by $[x]_{\mathcal{A}}$. Thus for $x \in U$ and $t \in R^n$ we have

$$t = [x]_{\mathcal{A}} \iff x = \sum_{i=1}^{n} t_i u_i \,.$$

The elements $t_1, t_2, \cdots, t_n \in R$ are called the **coordinates** of x with respect to the ordered basis $\mathcal{A}$. In the case that R is a field, the vector $t = [x]_{\mathcal{A}}$ is called the **coordinate vector** of x with respect to $\mathcal{A}$. The bijective map $\phi_{\mathcal{A}} : U \to R^n$ given by

$$\phi_{\mathcal{A}}(x) = [x]_{\mathcal{A}}$$

is called the **coordinate map** from U to R^n induced by the ordered basis $\mathcal{A}$.

6.21 Theorem: *Let R be a ring, let U be a free R-module, and let $\mathcal{A}$ be a finite ordered basis for U. Then coordinate map $\phi_{\mathcal{A}} : U \to R^n$ is bijective and linear. That is*

$$\phi_{\mathcal{A}}(x + y) = \phi_{\mathcal{A}}(x) + \phi_{\mathcal{A}}(y) \quad \textit{and} \quad \phi_{\mathcal{A}}(tx) = r\,\phi_{\mathcal{A}}(x)$$

for all $x, y \in U$ and all $r \in R$.

Proof: I may include a proof later.

The Dimension of Finite Dimensional Vector Spaces

6.22 Note: Let R be a ring and let U be a free R-module. If $\mathcal{A}$ and $\mathcal{B}$ are bases for U with $\mathcal{A} \subseteq \mathcal{B}$, then we must have $\mathcal{A} = \mathcal{B}$ because if we had $\mathcal{A} \subsetneqq \mathcal{B}$ then we could choose $v \in \mathcal{B} \setminus \mathcal{A}$ then write v as a linear combination $v = \sum_{i=1}^{n} t_i u_i$ with each $u_i \in \mathcal{A}$, but then we would would have $0 = 1 \cdot v - \sum_{i=1}^{n} t_i u_i$ which is a linear combination of elements in $\mathcal{B}$ with coefficients not all equal to zero.

6.23 Theorem: *Let R be a ring and let U be a free R-module. Let $\mathcal{A}$ and $\mathcal{B}$ be two bases for U over R. Then $\mathcal{A}$ and $\mathcal{B}$ are either both finite or both infinite.*

Proof: If $U = \{0\}$ then $\mathcal{A} = \mathcal{B} = \emptyset$. Suppose that $U \neq \{0\}$. Suppose that one of the two bases is finite, say $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$. For each index i, write $u_i = \sum_{j=1}^{m_i} t_{i,j} v_{i,j}$ with $t_{i,j} \in R$ and $v_{i,j} \in \mathcal{B}$. Let $\mathcal{C} = \{v_{i,j} \big| 1 \leq i \leq n\,,\, 1 \leq j \leq m_i\}$. Note that $\mathcal{C}$ is finite, and $\mathcal{C}$ is linearly independent because $\mathcal{C} \subseteq \mathcal{B}$ and $\mathcal{B}$ is linearly independent, and $\mathcal{C}$ spans U because given $x \in U$ we can write $x = \sum_{i=1}^{n} t_i u_i$ and then we have $x = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (t_i s_{i,j}) v_{i,j} \in \text{Span}\,(\mathcal{C})$. Since $\mathcal{B}$ and $\mathcal{C}$ are both bases for U and $\mathcal{C} \subseteq \mathcal{B}$ we have $\mathcal{C} = \mathcal{B}$, so $\mathcal{B}$ is finite.

6.24 Theorem: *Let F be a field and let U be a vector space over F. Let $\mathcal{A}$ and $\mathcal{B}$ be finite bases for U over F. Then $|\mathcal{A}| = |\mathcal{B}|$.*

Proof: If $U = \{0\}$ then $\mathcal{A} = \mathcal{B} = \emptyset$. Suppose that $U \neq \{0\}$. Let $n = |\mathcal{A}|$ and say $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$. Replace the set $\mathcal{A}$ by the ordered n-tuple $\mathcal{A} = (u_1, u_2, \cdots, u_n)$. Let $\phi = \phi_{\mathcal{A}} : U \to F^n$ and consider the set $\phi(\mathcal{B}) = \{\phi(v) \big| v \in \mathcal{B}\} \subseteq F^n$. Note that $\phi(\mathcal{B})$ spans F^n because given $t \in F^n$ we can let $x = \sum_{i=1}^{n} t_i u_i$ so that $[x]_{\mathcal{A}} = t$, then we can write $x = \sum_{i=1}^{m} s_i v_i$ with $s_i \in R$ and $v_i \in \mathcal{B}$, and then we have $t = \phi(x) = \phi\Big(\sum_{i=1}^{m} s_i v_i\Big) = \sum_{i=1}^{m} s_i \phi(v_i) \in \text{Span}\,\big(\phi(\mathcal{B})\big)$. Also, we claim that $\phi(\mathcal{B})$ is linearly independent in F^n. Suppose that $\sum_{i=1}^{m} s_i y_i = 0$ where the y_i are distinct elements in F^n and each $s_i \in F$. Choose elements $x_i \in \mathcal{B}$ so that $\phi(x_i) = y_i$, and note that the elements x_i will be distinct because ϕ is bijective. Then we have $0 = \sum_{i=1}^{m} s_i y_i = \sum_{i=1}^{m} s_i \phi(x_i) = \phi\Big(\sum_{i=1}^{m} s_i x_i\Big)$. Since ϕ is injective it follows that $\sum_{i=1}^{m} s_i x_i = 0$. Since the elements x_i are distinct elements in $\mathcal{B}$ and $\mathcal{B}$ is linearly independent, it follows that every $s_i = 0$. Thus $\phi(\mathcal{B})$ is linearly independent, as claimed. Since $\phi(\mathcal{B})$ spans F^n it follows that $\big|\phi(\mathcal{B})\big| \geq n$, and since $\phi(\mathcal{B})$ is linearly independent it follows that $\big|\phi(\mathcal{B})\big) \leq n$, and so we have $\big|\phi(\mathcal{B})\big| = n = |\mathcal{A}|$. Since ϕ is bijective we have $|\mathcal{B}| = \big|\phi(\mathcal{B})\big| = |\mathcal{A}|$.

6.25 Definition: Let U be a vector space over a field F. When U has a finite basis, we say that U is **finite dimensional** (over F) and we define the **dimension** of U to be $\dim(U) = |\mathcal{A}|$ where $\mathcal{A}$ is any basis for U.

The Existence of a Basis for a Vector Space

6.26 Definition: Let S be a nonempty set of sets. A **chain** in S is a nonempty subset $T \subseteq S$ with the property that for all $A, B \in T$, either $A \subseteq B$ or $B \subseteq A$. For a subset $T \subseteq S$, an **upper bound** for T in S is an element $B \in S$ such that $A \subseteq B$ for all $A \in T$. A **maximal** element in S is an element $B \in S$ such that there is no $A \in S$ with $B \subseteq A$.

6.27 Theorem: *(Zorn's Lemma) Let S be a nonempty set. Suppose that every chain in S has an upper bound in S. Then S has a maximal element.*

Proof: We take Zorn's Lemma to be an axiom, which means that we accept it as true without proof.

6.28 Note: Let F be a field and let U be a vector space over F. Let $\mathcal{A}$ be a linearly independent subset of U and let $v \in U$. Then $\mathcal{A} \cup \{v\}$ is linearly independent if and only if $v \notin \text{Span}(\mathcal{A})$ We leave the proof of this result as an exercise. Note that the analogous result does not hold when U is a module over a ring R.

6.29 Theorem: *Every vector space has a basis. Indeed, every linearly independent set in a vector space is contained in a basis for the vector space.*

Proof: Let F be a field and let U be a vector space over F. Let $\mathcal{A}$ be a subset of U which is linearly independent over F. Let S be the collection of all linearly independent sets $\mathcal{B} \subseteq U$ with $\mathcal{A} \subseteq \mathcal{B}$. Note that $S \neq \emptyset$ because $\mathcal{A} \in S$. Let T be a chain in S. We claim that $\bigcup T \in S$. Since $T \neq \emptyset$ we can choose an element $\mathcal{B}_\prime \in T$ and then we have $\mathcal{A} \subseteq \mathcal{B}_\prime \subseteq \bigcup T$. Since for every $\mathcal{B} \in T$ we have $\mathcal{B} \subseteq U$ it follows that $\bigcup T \subseteq U$. It remains to show that $\bigcup T$ is linearly independent. Suppose that $\sum_{i=1}^{n} t_i u_i = 0$ where the u_i are distinct elements in $\bigcup T$ and $t_i \in F$. For each index i, since $u_i \in \bigcup T$ we can choose $\mathcal{B}_i \in T$ with $u_i \in \mathcal{B}_i$. Since T is a chain, for all indices i and j, either $\mathcal{B}_i \subseteq \mathcal{B}_j$ or $\mathcal{B}_| \subseteq \mathcal{B}_i$. It follows that we can choose an index k so that $\mathcal{B}_i \subseteq B_k$ for all indices i. Then we have $u_i \in \mathcal{B}_k$ for all i. Since the u_i are distinct elements in $\mathcal{B}_k$ with $\sum_{i=1}^{n} t_i u_i = 0$ and since $\mathcal{B}_\|$ is linearly independent it follows that $t_i = 0$ for every i. This shows that $\bigcup T$ is linearly independent, and so $\bigcup T \in S$, as claimed. Since $\bigcup T \in S$ it follows that T has an upper bound in S since for every $\mathcal{B} \in T$ we have $\mathcal{B} \subseteq \bigcup T$. By Zorn's Lemma, it follows that S has a maximal element. Let $\mathcal{B}$ be a maximal element in S. We claim that $\mathcal{B}$ is a basis for U. Since $\mathcal{B} \in S$ we know that $\mathcal{A} \subseteq \mathcal{B} \subseteq U$ and that $\mathcal{B}$ is linearly independent. Note also that $\mathcal{B}$ spans U because if we had $\text{Span}(\mathcal{B}) \subsetneqq U$ then we could choose $w \in U$ with $w \notin \text{Span}(\mathcal{B})$ and then $\mathcal{B} \cup \{w\}$ would be linearly independent by the above Note, but then $\mathcal{B}$ would not be maximal in S.

6.30 Example: When F is a field and A is any set, the vector spaces F^ω, $F[[x]]$ and $\text{Func}(A, F)$ all have bases. It is not easy to construct an explicit basis for any of these vector spaces.

6.31 Example: There exists a basis for $\mathbf{R}$ as a vector space over $\mathbf{Q}$, but it is not easy to construct an explicit basis.

Some Cardinal Arithmetic

6.32 Definition: Let S be a set of nonempty sets. A **choice function** on S is a function $f : S \to \bigcup S$ such that $f(A) \in A$ for every $A \in S$.

6.33 Theorem: *(The Axiom of Choice) Every set of nonempty sets has a choice function.*

Proof: A proof can be found in Ehsaan Hossain's Tutorial Lecture Notes.

6.34 Corollary: *Let A be a set. For each $\alpha \in A$, let X_α be a nonempty set. Then there exists a function $f : A \to \bigcup\limits_{\alpha \in A} X_\alpha$ with $f(\alpha) \in X_\alpha$ for all $\alpha \in A$.*

Proof: Let $S = \{X_\alpha | \alpha \in A\}$. Note that $\bigcup S = \bigcup\limits_{\alpha \in A} X_\alpha$. Let $g : S \to \bigcup S$ be a choice function for S, so we have $g(X_\alpha) \in X_\alpha$ for all $\alpha \in A$. Define the map $f : A \to \bigcup\limits_{\alpha \in A} X_\alpha$ by $f(\alpha) = g(X_\alpha)$ to obtain $f(\alpha) \in X_\alpha$ for all $\alpha \in A$, as required..

6.35 Theorem: *Let A and B be nonempty sets and let $f : A \to B$. Then*

(1) f is injective if and only if f has a left inverse, and
(2) f is surjective if and only if f has a right inverse.

Proof: The proofs can be found in last term's MATH 147 Lecture Notes. We remark that the proof of Part (1) does not require the Axiom of Choice but the proof of Part (2) does.

6.36 Definition: *For sets A and B, we say that A and B are* **equipotent** *(or that A and B have the same cardinality), and we write $|A| = |B|$ when there exists a bijection $f : A \to B$. We say that the cardinality of A is* **less than or equal to** *the cardinality of B, and we write $|A| \leq |B|$, when there exists an injective map $f : A \to B$. Note that by the above theorem, we have $|A| \leq |B|$ if and only if there exists a surjective map $fgB \to A$.*

6.37 Note: *It follows immediately from the above definitions that for all sets A, B and C we have*

(1) $|A| = |A|$,
(2) if $|A| = |B|$ then $|B| = |A|$, and
(3) if $|A| = |B|$ and $|B| = |C|$ then $|A| = |C|$, and also
(4) $|A| \leq |A|$, and
(5) if $|A| \leq |B|$ and $|B| \leq |C|$ then $|A| \leq |C|$.

Properties 1, 2 and 3 imply that equipotence is an equivalence relation on the class of all sets. Properties 4 and 5 are two of the 4 properties which appear in the definition a total order. The other 2 properties also hold, but they require proof. The third property is known as the Cantor-Schroeder-Bernstein Theorem, and we state it below. After that, we state and prove the fourth property.

6.38 Theorem: *(The Cantor-Schroeder-Bernstein Theorem) Let A and B be sets. If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.*

Proof: A proof can be found in last term's MATH 147 Lecture Notes (we remark that it does not require the Axiom of Choice).

6.39 Theorem: *Let A and B be sets. Then either $|A| \leq |B|$ or $|B| \leq |A|$.*

Proof: If $A = \emptyset$ we have $|A| \leq |B|$. If $B = \emptyset$ we have $|B| \leq |A|$. Suppose that $|A| \neq \emptyset$ and $B \neq \emptyset$. Let S be the set of all (graphs of) injective functions $f : X \to B$ with $X \subseteq A$. Note that $S \neq \emptyset$ since we can choose $a \in A$ and $b \in B$ and define $f : \{a\} \to B$ by $f(a) = b$. Let T be a chain in S. Note that $\bigcup T \in S$, as in the proof of the Axiom of Choice found in Ehsaan Hossain's notes, and so T has an upper bound in S. By Zorn's Lemma, it follows that S has a maximal element. Let (the graph of) $g : X \to B$ be a maximal element in S. Note that either $X = A$ or $g(X) = B$ since if we had $X \subsetneqq A$ and $g(X) \subsetneqq B$ then we could choose an element $a \in A \setminus X$ and an element $b \in B \setminus f(A)$ and then extend g to the injective map $h : X \cup \{a\} \to B$ defined by $h(x) = g(x)$ for $x \in X$ and $h(a) = b$, contradicting the maximality of g. In the case that $X = A$, since the map $g : A \to B$ is injective we have $|A| \leq |B$. In the case that $f(X) = B$, the map $g : X \to B$ is surjective so we have $|B| \leq |X| \leq |A|$.

6.40 Note: Recall that for sets X and Y, the set of all functions $f : Y \to X$ is denoted by X^Y. As an exercise, verify that given sets A_1, A_2, B_1 and B_2 with $|A_1| = |A_2|$ and $|B_1| = |B_2|$, we have

(1) $|A_1 \times B_1| = |A_2 \times B_2|$,
(2) $\left|{A_1}^{B_1}\right| = \left|{A_2}^{B_2}\right|$, and
(3) if $A_1 \cap B_1 = \emptyset$ and $A_2 \cap B_2 = \emptyset$ then $|A_1 \cup B_1| = |A_2 \cup B_2|$.

6.41 Definition: (Cardinal Arithmetic) For sets A, B and X, we write $|X| = |A|\,|B|$ when $|X| = |A \times B|$, $|X| = |A|^{|B|}$ when $|X| = |A^B|$, and $|X| = |A| + |B|$ when $|X| = |A' \cup B'|$ for disjoint sets A' and B' with $|A'| = |A|$ and $|B'| = |B|$ (for example the sets $A' = A \times \{1\}$ and $B' = B \times \{2\}$).

6.42 Note: Let B be an infinite set and let $n \in \mathbf{Z}^+$, and let $S_n = \{1, 2, \cdots, n\}$. As an exercise, show that there are injective maps $f : B \to B \times S_n$ and $g : B \times S_n \to B \times B$ so that $|B| \leq |B \times S_n| \leq |B \times B|$, then use the Cantor-Schroeder-Bernstein Theorem to show that if $|B| = |B \times B|$ then we have $|B \times S_n| = |B|$.

6.43 Theorem: *Let A be an infinite set. Then $|A \times A| = |A|$.*

Proof: Let S be the set of all (graphs of) bijective functions $f : X \times X \to X$ where X is an infinite subset of A. Note that $S \neq \emptyset$ because, as proven in last term's MATH 147 Lecture notes, we can choose a countable subset $X \in A$ and a bijection $f : X \times X \to X$. Let T be a chain in S. Note that $\bigcup T \in S$ as in the proof of the Axiom of Choice found in Ehsaan Hossain's notes, and so T has an upper bound in S. By Zorn's Lemma, S has a maximal element. Let (the graph of) $g : B \times B \to B$ be a maximal element in S. Note that since $g : B \times B \to B$ is bijective we have $|B \times B| = |B|$. By the previous theorem, either $|B| \leq |A \setminus B|$ or $|A \setminus B| \leq |B|$. We claim that $|A \setminus B| \leq |B|$.

Suppose, for a contradiction, that $|B| \leq |A \setminus B|$. Choose $C \subseteq A \setminus B$ with $|C| = |B|$. Note that the set $(B \cup C) \times (B \cup C)$ is the disjoint union

$$(B \cup C) \times (B \cup C) = (B \times B) \cup (B \times C) \cup (C \times B) \cup (C \times C)$$

and we have

$$\begin{aligned} \big|(B \times C) \cup (C \times B) \cup (C \times C)\big| &= |B \times C| + |C \times B| + |C \times C| \\ &= |B \times B| + |B \times B| + |B \times B| = |B| + |B| + |B| = |B \times \{1, 2, 3\}| = |B| = |C| \end{aligned}$$

and so the maximal bijective map $g : B \times B \to B$ can be extended to a bijective map $h : (B \cup C) \times (B \cup C) \to (B \cup C)$ contradicting the maximality of g. Thus the case in which $|B| \leq |A \setminus B|$ cannot arise, and so we must have $|A \setminus B| \leq |B|$, as claimed.

Since $|A \setminus B| \leq |B|$ we have

$$|A| = \big|(A \setminus B) \cup B\big| = |A \setminus B| + |B| \leq |B| + |B| = |B \times \{1, 2\}| = |B|$$

and hence

$$|A \times A| = |B \times B| = |B| = |A|.$$

6.44 Corollary: *Let A and B be sets.*

(1) If A is nonempty and B is infinite and $|A| \leq |B|$, then $|A|\,|B| = |B|$.
(2) If B is infinite and $|A| \leq |B|$ then $|A| + |B| = |B|$.

Proof: The proof is left as an exercise.

6.45 Corollary: *Let A be an infinite set. For each $u \in A$ let B_u be a finite set. Then*

$$\Big|\bigcup_{u \in A} B_u\Big| \leq |A|.$$

Proof: For each $u \in A$, choose a surjective map $f_u : A \to B_u$. Define $f : A \times A \to \bigcup_{u \in A} B_u$ by $g(u, v) = f_u(v)$. Note that g is surjective and so we have

$$\Big|\bigcup_{u \in A} B_u\Big| \geq |A \times A| = |A|.$$

The Dimension of an Infinite Dimensional Vector Space

6.46 Theorem: *Let R be a ring and let U be a free R-module. Then any two infinite bases for U have the same cardinality.*

Proof: Let $\mathcal{A}$ and $\mathcal{B}$ be infinite bases for U. For each $u \in \mathcal{A}$, let $c = c(u) : \mathcal{A} \to R$ be the (unique) function, with $c(u)_v = 0$ for all but finitely many elements $v \in \mathcal{B}$, such that $u = \sum_{v \in \mathcal{B}} c(u)_v \cdot v$, and then let B_u be the set of all elements $v \in \mathcal{B}$ for which $c(u)_v \neq 0$. Note that each set B_u is a nonempty finite subset of $\mathcal{B}$. Let $\mathcal{C} = \bigcup_{u \in \mathcal{A}} B_u$. Note that $\mathcal{C}$ spans U because given any $x \in U$ we can write $x = \sum_{i=1}^{n} t_i u_i$ with each $u_i \in \mathcal{A}$, and then for each index i we can write $u_i = \sum_{j=1}^{m_i} s_{i,j} v_{i,j}$ with each $v_{i,j} \in B_{u_i}$, and then we have $x = \sum_{i,j} (t_i s_{i,j}) v_{i,j} \in \text{Span}\big(\bigcup_{i=1}^{n} B_{u_i}\big) \subseteq \text{Span}\,(\mathcal{C})$. Since $\mathcal{B}$ is linearly independent and $\mathcal{C} \subseteq \mathcal{B}$, it follows that $\mathcal{C}$ is linearly independent. Since $\mathcal{C}$ is linearly independent and spans U, it is a basis for U. Since $\mathcal{C}$ and $\mathcal{B}$ are bases for U with $\mathcal{C} \subseteq \mathcal{B}$ it follows that $\mathcal{C} = \mathcal{B}$ because if we had $\mathcal{C} \subsetneqq \mathcal{B}$ then we could choose $v \in \mathcal{B} \setminus \mathcal{C}$ then write v as a linear combination $v = \sum_{i=1}^{n} t_i v_i$ with each $v_i \in \mathcal{C}$, but then we would have $0 = 1 \cdot v - \sum_{i=1}^{n} t_i v_i$ which is a linear combination of elements in $\mathcal{B}$ with not all coefficients equal to 0. By the above theorem, we have

$$|\mathcal{B}| = |\mathcal{C}| = \Big| \bigcup_{u \in \mathcal{A}} B_u \Big| \leq |\mathcal{A}|.$$

By interchanging the rôles of $\mathcal{A}$ and $\mathcal{B}$ in the above proof, we see that $|\mathcal{A}| \leq |\mathcal{B}|$. Thus we have $|\mathcal{A}| = |\mathcal{B}|$ by the Cantor-Schroeder-Bernstein Theorem.

6.47 Definition: Let R be a ring and let U be a free R-module with an infinite basis. We define the **rank** of R to be $\text{rank}(R) = |\mathcal{A}|$ where $\mathcal{A}$ is any basis for U. When F is a field and U is a vector space over F which has an infinite basis, the rank of U is also called the **dimension** of U.