Math 751 **Introductory Topology** **Fall 2010**

0. Mathematical Background

Many theorems in this chapter are stated without proof because their proofs are routine and well known. If a student is not familiar with a proof, he should regard it as an exercise.

A. Sets

Abbreviations.

$\Rightarrow$	implies
$\Leftrightarrow$	if and only if
$\forall$	for every
$\exists$	there exists
$\exists!$	there exists a unique
s.t.	such that
$x \in X$	x is an element of X

Definitions. Let X and Y be sets.
X is a *subset* of Y, denoted $X \subset Y$, means $x \in X \Rightarrow x \in Y$.
X *equals* Y, denoted $X = Y$, means $x \in X \Leftrightarrow x \in Y$.
$x \notin X$, $X \not\subset Y$, and $X \neq Y$ denote the negations of $x \in X$, $X \subset Y$, and $X = Y$, respectively.

Theorem 0.1. $X = Y \Leftrightarrow X \subset Y$ and $Y \subset X$.

Definition. A set is *empty* if it has no elements.

Theorem 0.2. a) If $\varnothing$ is an empty set and X is any set, then $\varnothing \subset X$.
b) If $\varnothing$ and $\varnothing'$ are empty sets, then $\varnothing = \varnothing'$.

Definition. Let $\varnothing$ denote the (unique) empty set.

Definitions. If $\Phi(x)$ is a statement about the object x, let

$$\{ x : \Phi(x) \}$$

denote the set of all objects x such that $\Phi(x)$ is true. For example, $\{ x : x \neq x \} = \varnothing$.
If X is a set and $\Phi(x)$ is a statement about x, let

$$\{ x \in X : \Phi(x) \} = \{ x : x \in X \text{ and } \Phi(x) \}.$$

If $x_1, x_2, \ldots, x_n$ are objects, let $\{ x_1, x_2, \ldots, x_n \} = \{ x : x = x_1, \text{ or } x = x_2, \text{ or } \cdots, x = x_n \}$.

Definitions. Let X and Y be sets.
The *union* of X and Y is $X \cup Y = \{ x : x \in X \text{ or } x \in Y \}$.
The *intersection* of X and Y is $X \cap Y = \{ x : x \in X \text{ and } x \in Y \}$.
The *difference* of X and Y or the *complement* of Y in X is $X - Y = \{ x : x \in X \text{ and } x \notin Y \}$.

Theorem 0.3. Let X, Y and Z be sets.
a) (Associativity) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ and $(X \cap Y) \cap Z = X \cap (Y \cap Z)$.
b) (Commutativity) $X \cup Y = Y \cup X$ and $X \cap Y = Y \cap X$.
c) (Distributativity) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ and
$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$.
d) (Identity) $\varnothing \cup X = X$ and $\varnothing \cap X = \varnothing$.
e) (De Morgan Laws) $X - (Y \cup Z) = (X - Y) \cap (X - Z)$ and
$X - (Y \cap Z) = (X - Y) \cup (X - Z)$.
f) (Idempotence) $X \cup X = X = X \cap X$.
g) (Absorption) $X \cup (X \cap Y) = X = X \cap (X \cup Y)$.
h) $X \subset Y \Leftrightarrow X \cup Y = Y \Leftrightarrow X \cap Y = X$.

Remark. In general, the words "collection" and "family" mean the same as "set". However, in these notes, we reserve the term "collection" to mean a set whose elements are sets. Thus we speak of a "collection of sets" instead of a "set of sets".

Definitions. Let $\mathscr{X}$ be a collection of sets.
The *union* of $\mathscr{X}$ is $\bigcup \mathscr{X} = \{ x : x \in X \text{ for some } X \in \mathscr{X} \}$.
The *intersection* of $\mathscr{X}$ is $\bigcap \mathscr{X} = \{ x : x \in X \text{ for every } X \in \mathscr{X} \}$.

Theorem 0.4. (Distributive Laws) If $\mathscr{X}$ and $\mathscr{Y}$ are collections of sets, then
a) $(\bigcup \mathscr{X}) \cap (\bigcup \mathscr{Y}) = \bigcup \{ X \cap Y : X \in \mathscr{X} \text{ and } Y \in \mathscr{Y} \}$, and
b) $(\bigcap \mathscr{X}) \cup (\bigcap \mathscr{Y}) = \bigcap \{ X \cup Y : X \in \mathscr{X} \text{ and } Y \in \mathscr{Y} \}$.

Theorem 0.5. (De Morgan's Laws) If X is a set and $\mathscr{Y}$ is a collection of sets, then
a) $X - (\bigcup \mathscr{Y}) = \bigcap \{ X - Y : Y \in \mathscr{Y} \}$, and
b) $X - (\bigcap \mathscr{Y}) = \bigcup \{ X - Y : Y \in \mathscr{Y} \}$.

It is sometimes convenient to introduce an *indexed* collection of sets of the form $\{ X_\gamma : \gamma \in \Gamma \}$, Γ being the *index set.* (The indexed collections of sets $\{ X_\gamma : \gamma \in \Gamma \}$ is really a function with domain Γ which assigns the set X_γ to the element $\gamma \in \Gamma$. The definition of “function” is given in section B.) We denote the union of $\{ X_\gamma : \gamma \in \Gamma \}$ by $\bigcup_{\gamma \in \Gamma} X_\gamma$, and we denote the intersection of $\{ X_\gamma : \gamma \in \Gamma \}$ by $\bigcap_{\gamma \in \Gamma} X_\gamma$. If $\{ X_\gamma : \gamma \in \Gamma \}$ and $\{ Y_\delta : \delta \in \Delta \}$ are indexed collections of sets and X is a set, then the Distributive Laws stated above take on the form

a) $(\bigcup_{\gamma \in \Gamma} X_\gamma) \cap (\bigcup_{\delta \in \Delta} Y_\delta) = \bigcup_{\gamma \in \Gamma, \delta \in \Delta} (X_\gamma \cap Y_\delta)$ and

b) $(\bigcap_{\gamma \in \Gamma} X_\gamma) \cup (\bigcap_{\delta \in \Delta} Y_\delta) = \bigcap_{\gamma \in \Gamma, \delta \in \Delta} (X_\gamma \cup Y_\delta)$

and De Morgan's Laws take on the form

a) $X - (\bigcup_{\delta \in \Delta} Y_\delta) = \bigcap_{\delta \in \Delta} (X - Y_\delta)$ and

b) $X - (\bigcap_{\delta \in \Delta} Y_\delta) = \bigcup_{\delta \in \Delta} (X - Y_\delta)$.

Definition. If X is a set, then the *power set* of X is $\mathscr{P}(X) = \{ A : A \subset X \}$.

Definition. For any two objects x and y, there is an object called an *ordered pair,* which is denoted (x,y), and which has the following characteristic property:

$$(x,y) = (x',y') \quad \Leftrightarrow \quad x = x' \text{ and } y = y'.$$

For our purposes, this is an adequate definition of "ordered pair". However, if we wish to think of everything in our mathematical universe as a set, then we can interpret every ordered pair as a set. For instance, if we identify the ordered pair (x,y) with the set $\{ \{x\}, \{x,y\} \}$, then the characteristic property of ordered pairs will be satisfied.

Definition. The *Cartesian product* of two sets X and Y is

$$X \times Y = \{ (x,y) : x \in X \text{ and } y \in Y \}.$$

B. Functions

Definition. A *function* from a set X to a set Y is an object which assigns to each element x belonging to the set X a unique element f(x) that belongs to the set Y. If f is a function from the set X to the set Y, we denote this by writing $f : X \to Y$. In this situation, we call X the *domain* of f, and for each $x \in X$, we call f(x) the *value* of f at x.

While this definition suffices for our purposes, if we wish to think of all the objects in our mathematical universe as sets, we can interpret functions as sets. To accomplish this, we identify every function with its graph. Thus, we define f to be a function from the set X to the set Y if and only if f is a subset of the Cartesian product $X \times Y$ with the property that of each $x \in X$, there is a unique $y \in Y$ such that $(x,y) \in f$. In this situation, for each $x \in X$, we identify f(x) with the unique $y \in Y$ such that $(x,y) \in f$. This approach identifies f with the subset $\{ (x,y) \in X \times Y : y = f(x) \}$ of $X \times Y$ which is usually called the *graph* of f.

Note that according to the preceding definition, $\varnothing$ is a function from $\varnothing$ to any set.

Theorem 0.6. Two functions f and g are equal if and only if their domains are equal and f(x) = g(x) for every element x of the domain of f.

Theorem 0.6 allows us to define a function by specifying its domain and by specifying its value at every point of its domain. The following two definitions define functions in exactly this way.

Definition. For a set X, the *identity function* on X, denoted $id_X : X \to X$, is the function with domain X satisfying $id_X(x) = x$ for $x \in X$.

Definition. If $f : X \to Y$ and $g : Y \to Z$ are functions, then their *composition,* denoted $g \circ f : X \to Z$, is the function with domain X satisfying $g \circ f(x) = g(f(x))$ for $x \in X$.

Theorem 0.7. a) If $f : X \to Y$ is a function, then $f \circ id_X = f = id_Y \circ f$.
b) If $f : W \to X$, $g : X \to Y$ and $h : Y \to Z$ are functions, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Definitions. Let $f : X \to Y$ be a function.
f is *injective* or *one-to-one* if for all $x, x' \in X$, $x \neq x' \Rightarrow f(x) \neq f(x')$ (or equivalently, $f(x) = f(x') \Rightarrow x = x'$).
f is *surjective* or *onto* if for every $y \in Y$, there is an $x \in X$ such that $f(x) = y$.
f is *bijective* or a *one-to-one correspondence* if f is both injective and surjective.

Note that $\varnothing : \varnothing \to Y$ is injective for any set Y, and $\varnothing : \varnothing \to \varnothing$ is bijective.

Theorem 0.8. Let $f : X \to Y$ be a function.
a) f is injective $\Leftrightarrow$ there is a function $g : Y \to X$ such that $g \circ f = id_X$. (In this case, g is called a *left inverse* of f.)
b) f is surjective $\Leftrightarrow$ there is a function $g : Y \to X$ such that $f \circ g = id_Y$. (In this case, g is called a *right inverse* of f.)
c) f is bijective $\Leftrightarrow$ there is a unique function $g : Y \to X$ such that $g \circ f = id_X$ and $f \circ g = id_Y$. (In this case, g is called the *inverse* of f and is denoted f^{-1}.)

Proof. a) Assume f is injective. Pick $x_0 \in X$. Define $g : Y \to X$ as follows. For $y \in Y$: (1) if there is an $x \in X$ such that $y = f(x)$, then set $g(y) = x$; and (2) if $y \neq f(x)$ for every $x \in X$, then set $g(y) = x_0$. Since f is injective, there is at most one $x \in X$ such that $f(x) = y$. So this process unambiguously defines a function $g : Y \to X$. It is clear from the definition of g that if $x \in X$, then $g(f(x)) = x$. So $g \circ f = id_X$.

Assume there is a function $g : Y \to X$ such that $g \circ f = id_X$. If $x, x' \in X$ and $f(x) = f(x')$, then $x = g(f(x)) = g(f(x')) = x'$. This proves f is injective.

b) Assume f is surjective. Then for every $y \in Y$, the set $f^{-1}(\{y\}) = \{ x \in X : f(x) = y \}$ is a non-empty subset of X. Hence, a function $g : Y \to X$ can be defined by requiring that for every $y \in Y$, $g(y) \in f^{-1}(\{y\})$. (Here we are using the *Axiom of Choice.* See the remark following this proof.) Clearly $f(g(y)) = y$ for every $y \in Y$. So $f \circ g = id_Y$.

Assume there is a function $g : Y \to X$ such that $f \circ g = id_Y$. If $y \in Y$, then $g(y) \in X$ and $f(g(y)) = y$. So f is surjective.

c) First, we prove:

Lemma. If $g, g' : Y \to X$ are functions such that $g \circ f = id_X$ and $f \circ g' = id_Y$, then $g = g'$.

Proof of Lemma. $g = g \circ id_Y = g \circ (f \circ g') = (g \circ f) \circ g' = id_X \circ g' = g'$. ☐

Now assume f is bijective. Then a) and b) provide functions $g, g' : Y \to X$ such that $g \circ f = id_X$ and $f \circ g' = id_Y$. Then the lemma implies $g = g'$. Hence, $g \circ f = id_X$ and $f \circ g = id_Y$. Therefore g is an inverse of f. If $g'' : Y \to X$ also satisfies $g'' \circ f = id_X$ and $f \circ g'' = id_Y$, then the lemma implies $g = g''$. Thus g is the *unique* inverse of f.

Conversely assume there is a function $g : Y \to X$ such that $g \circ f = id_X$ and $f \circ g = id_Y$. Then a) and b) imply f is bijective. ☐

Remark. The proof of b) requires the following fundamental set theoretic principle.

The Axiom of Choice. If $\{ X_\gamma : \gamma \in \Gamma \}$ is an indexed collection of non-empty sets, then $\exists$ a function $g : \Gamma \to \bigcup_{\gamma \in \Gamma} X_\gamma$ called a *choice function,* such that $g(\gamma) \in X_\gamma$ for every $\gamma \in \Gamma$.

The Axiom of Choice is a strong and useful set theoretic principle. While the founders of set theory tended to believe that the Axiom of Choice is true, they did not find it glaringly self evident. They were not sure whether they should be able to prove it from more obvious set theoretic principles, or whether they would have to assume it as an axiom of set theory without proof. It is now known that neither the Axiom of Choice nor its negation can be proved from more self evident set theoretic principles. Hence, when the Axiom of Choice is needed in a proof (as it is in the proof of Theorem 0.8 b), it must be assumed as an axiom. In this course we will encounter other significant theorems that require either the Axiom of Choice or one of two other powerful set theoretic principles that are logically equivalent to the Axiom of Choice. These other two principles are known as *The Well Ordering Principle* and *Zorn's Lemma.* The standard approach of most mathematicians to the usage of principles like the Axiom of Choice is one of economy: don't use such principles when they can easily be avoided; and when they must be used in a proof, assume they are true, and remember that the proof depends on them.

Definition. Let $f : X \to Y$ be a function. If $A \subset X$, the set

$$f(A) \;=\; \{\, f(x) : x \in A \,\}$$

is called the *image of A under f.* $f(X)$ is simply called the *image of f.* If $B \subset Y$, the set

$$f^{-1}(B) \;=\; \{\, x \in X : f(x) \in B \,\}$$

is called the *preimage* or *inverse image of B under f.* Note: $f^{-1}(B)$ is *always* defined, even if f is not injective and has no well-defined inverse. If $A \subset X$, the *restriction of f to A* is the function

$$f \mid A : A \to Y$$

with domain A satisfying $(f \mid A)(x) = f(x)$ for $x \in A$.

Theorem 0.9. Let $f : X \to Y$ be a function.
a) If $A \subset A' \subset X$, then $f(A) \subset f(A')$.
Let $\mathscr{A}$ be a collection of subsets of X.
b) $f(\bigcup\mathscr{A}) \;=\; \bigcup\{\, f(A) : A \in \mathscr{A}\}$.
c) $f(\bigcap\mathscr{A}) \;\subset\; \bigcap\{\, f(A) : A \in \mathscr{A}\}$, and equality holds if f is injective.
d) If $A \subset A' \subset X$, then $f(A') - f(A) \;\subset\; f(A - A')$; and equality holds if f is injective.
e) If $A \subset X$, then $A \subset f^{-1}(f(A))$; and equality holds if f is injective.

Theorem 0.10. Let $f : X \to Y$ be a function.
a) If $B \subset B' \subset Y$, then $f^{-1}(B) \subset f^{-1}(B')$.
Let $\mathscr{B}$ be a collection of subsets of Y.
b) $f^{-1}(\bigcup\mathscr{B}) \;=\; \bigcup\{\, f^{-1}(B) : B \in \mathscr{B}\,\}$.
c) $f^{-1}(\bigcap\mathscr{B}) \;=\; \bigcap\{\, f^{-1}(B) : B \in \mathscr{B}\,\}$.
d) If $B \subset B' \subset Y$, then $f^{-1}(B' - B) = f^{-1}(B') - f^{-1}(B)$.
e) If $B \subset Y$, then $f(f^{-1}(B)) \subset B$; and equality holds if f is surjective.
f) If $A \subset X$ and $B \subset Y$, then $f(A \cap f^{-1}(B)) = f(A) \cap B$.

Theorem 0.11. If $f : X \to Y$ and $g : Y \to Z$ are functions and $C \subset Z$, then $(g\circ f)^{-1}(C) \;= f^{-1}(g^{-1}(C))$.

Definition. For sets X and Y, let Y^X denote the set of all functions from X to Y. (The motivation for this notation is that if X has m elements and Y has n elements, then Y^X has n^m elements.) For a positive integer n, let $Y^n = Y^{\{1, 2, \cdots, n\}}$. (Usually, Y^n is defined to be $\{ (y_1, y_2, \ldots , y_n) : y_i \in Y \text{ for } 1 \le i \le n \}$. This is consistent an because each function $y : \{ 1, 2, \cdots , n \} \to Y$ can be thought of as the n-tuple $(y(1), y(2), \cdots , y(n))$.)

C. Infinity

To give brief and simple proofs of the theorems in this section, we exploit the following two elementary properties of the prime numbers.

1) There are infinitely many prime numbers.

2) *The uniqueness of prime decompositions:* If $p_1^{m_1}p_2^{m_2}\ldots p_r^{m_r} = q_1^{n_1}q_2^{n_2}\ldots q_s^{n_s}$ where $p_1, p_2, \ldots , p_r$ are distinct prime numbers, $q_1, q_2, \ldots , q_s$ are distinct prime numbers, and $m_1, m_2, \ldots , m_r, n_1, n_2, \ldots , n_s$ are positive integers, then r = s and, after reindexing, $p_i = q_i$ and $m_i = n_i$ for $1 \le i \le r$.

Definition. Let X and Y be sets. X is *equivalent* to Y, denoted $X \approx Y$, if there is a bijection from X to Y. Let $X \not\approx Y$ denote the negation of $X \approx Y$.

Theorem 0.12. $\approx$ is an equivalence relation.
In other words, if X, Y and Z are sets then
a) $X \approx X$,
b) $X \approx Y \Rightarrow Y \approx X$, and
c) $X \approx Y$ and $Y \approx Z \Rightarrow X \approx Z$.

Theorem 0.13. If W, X, Y and Z are sets such that $W \approx Y$ and $X \approx Z$, then
a) $\mathscr{P}(W) \approx \mathscr{P}(Y)$,
b) $W \times X \approx Y \times Z$, and
c) $W^X \approx Y^Z$.

Definition. Let X and Y be sets. Write $X \preceq Y$ if there is an injection from X to Y. Write $X \prec Y$ if $X \preceq Y$ and $X \not\approx Y$. (Note: $\varnothing \preceq Y$ for any set Y.)

Theorem 0.14. If X, Y and Z are sets, then
a) $X \preceq X$ and
b) $X \preceq Y$ and $Y \preceq Z \Rightarrow X \preceq Z$.

The Schroder-Bernstein Theorem 0.15. If X and Y are sets, then
$X \preceq Y$ and $Y \preceq X \Rightarrow X \approx Y$.

The proof of the Schroder-Bernstein Theorem uses:

The Absorption Lemma. If X is a set, $h : X \rightarrow X$ is an injection, and $A \subset X - h(X)$, then $X \approx X - A$.

Proof of the Absorption Lemma. Set $A_1 = h(A)$, $A_2 = h(A_1)$, $A_3 = h(A_2)$, … , $A_{n+1} = h(A_n)$, …; and set $B = A_1 \cup A_2 \cup A_3 \cup \ldots$. Then $h(A \cup B) = B$. Also $A \cap B = \emptyset$ because $B \subset h(X)$ and $A \cap h(X) = \emptyset$. Now a bijection $k : X \rightarrow (X - A)$ is defined by setting $k | A \cup B = h | A \cup B$ and $k | X - (A \cup B) = \mathrm{id}_{X - (A \cup B)}$. □

Proof of the Schroder-Bernstein Theorem. Assume $X \preceq Y$ and $Y \preceq X$. Then there are injections $f : X \rightarrow Y$ and $g : Y \rightarrow X$. An injection $h : X \rightarrow X$ is defined by $h = g \circ f$. Set $A = X - g(Y)$. Since $h(X) = g(f(X)) \subset g(Y)$ and $A \cap g(Y) = \emptyset$, then $A \cap h(X) = \emptyset$. Therefore, the Absorption Lemma implies $X \approx X - A$. Also $Y \approx X - A$, because $g : Y \rightarrow g(Y)$ is a bijection and $g(Y) = X - A$. We conclude that $X \approx Y$. □

Definition. Let $\mathbb{N} = \{ 1, 2, 3, \cdots \}$, the set of all natural number (positive integers).
Let $\mathbb{Z} = \{ \cdots -2, -1, 0, 1, 2, \cdots \}$, the set of all integers.
Let $\mathbb{Q} = \{ {}^m/_n : m, n \in \mathbb{Z} \text{ and } n \neq 0 \}$, the set of all rational numbers.
Let $\mathbb{R}$ denote the set of all real numbers.

Definition. Let X be a set.
X is *finite* if either $X = \emptyset$ or $X \approx \{ 1, 2, \cdots , n \}$ for some $n \in \mathbb{N}$.
X is *infinite* if it is not finite.
X is *countable* if $X \preceq \mathbb{N}$.
X is *uncountable* if it is not countable.

Theorem 0.16. If X is an infinite set, then $\mathbb{N} \preceq X$. (The proof of this theorem requires the *Countable Axiom of Choice.*)

Proof. An injection $f : \mathbb{N} \rightarrow X$ is constructed inductively as follows. Since $X \neq \emptyset$, we can choose $f(1) \in X$. Let $n \in \mathbb{N}$ and inductively assume we have chosen distinct points $f(1), f(2), \cdots , f(n)$ in X. Since X is infinite, then $X - \{ f(1), f(2), \cdots , f(n) \} \neq \emptyset$, and we can choose $f(n+1) \in X - \{ f(1), f(2), \cdots , f(n) \}$. □

Corollary. If X is countable and infinite, then $X \approx \mathbb{N}$.

Theorem 0.17. If X and Y are countable, then so is $X \times Y$.

Proof. If $f : X \rightarrow \mathbb{N}$ and $g : Y \rightarrow \mathbb{N}$ are injections, then an injection $h : X \times Y \rightarrow \mathbb{N}$ is defined by $h(x,y) = 2^{f(x)}3^{g(y)}$. □

Corollary. $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$.

Theorem 0.18. $\mathbb{Q}$ is countable.

Proof. Each element of $\mathbb{Q}$ can be expressed uniquely as $(-1)^{\varepsilon}({}^{m}/_{n})$ where $\varepsilon \in \{0,1\}$, $m \in \{0\} \cup \mathbb{N}$, and $n \in \mathbb{N}$ is as small as possible. Hence, an injection $f : \mathbb{Q} \to \mathbb{N}$ is defined by $f((-1)^{\varepsilon}({}^{m}/_{n})) = 2^{\varepsilon}3^{m}5^{n}$. □

Theorem 0.19. If X_n is a countable set for every $n \in \mathbb{N}$, then $\bigcup_{n \in \mathbb{N}} X_n$ is countable.

Proof. For every $n \in \mathbb{N}$, there is an injection $f_n : X_n \to \mathbb{N}$. Define an injection $g : \bigcup_{n \in \mathbb{N}} X_n \to \mathbb{N}$ as follows. Let $x \in \bigcup_{n \in \mathbb{N}} X_n$. Then there is a unique *smallest* $n \in \mathbb{N}$ such that $x \in X_n$. Set $g(x) = 2^n 3^{f_n(x)}$. □

Theorem 0.20. If X is a set, then $\mathscr{P}(X) \approx \{ 0, 1 \}^X$.

Proof. For $A \subset X$, define the *characteristic function* $\chi_A : X \to \{ 0, 1 \}$ by $\chi_A(x) = 0$ if $x \notin A$ and $\chi_A(x) = 1$ if $x \in A$. Now define functions $\Phi : \mathscr{P}(X) \to \{ 0, 1 \}^X$ and $\Psi : \{ 0, 1 \}^X \to \mathscr{P}(X)$ by $\Phi(A) = \chi_A$ and $\Psi(f) = f^{-1}(\{1\})$. Then $\Psi \circ \Phi = \mathrm{id}$ and $\Phi \circ \Psi = \mathrm{id}$. So $\Phi : \mathscr{P}(X) \to \{ 0, 1 \}^X$ is a bijection. □

Theorem 0.21. If X is a set, then $X \prec \mathscr{P}(X)$.

Proof. An injection $f : X \to \mathscr{P}(X)$ is defined by $f(x) = \{x\}$. So $X \preceq \mathscr{P}(X)$. Assume $X \approx \mathscr{P}(X)$. Then there is a bijection $g : X \to \mathscr{P}(X)$. Set $A = \{ x \in X : x \notin g(x) \}$; and observe that for each $x \in X$, $x \in g(x) \Rightarrow x \notin A$, and $x \notin g(x) \Rightarrow x \in A$. Since $A \in \mathscr{P}(X)$ and since $g : X \to \mathscr{P}(X)$ is a surjection, then there is an $a \in X$ such that $g(a) = A$. But then: $a \in A \Rightarrow a \in g(a) \Rightarrow a \notin A$, and $a \notin A \Rightarrow a \notin g(a) \Rightarrow a \in A$. We have reached a contradiction. We conclude that $X \not\approx \mathscr{P}(X)$. This proves $X \prec \mathscr{P}(X)$. □

Corollary. If X is a set, then $X \prec \{0,1\}^X$.

Corollary. $\mathscr{P}(\mathbb{N})$ and $\{0,1\}^{\mathbb{N}}$ are uncountable.

Theorem 0.22. $\mathbb{R} \approx \{ 0, 1 \}^{\mathbb{N}}$.

Proof. We will justify every relation in the sequence $\{ 0, 1 \}^{\mathbb{N}} \preceq \mathbb{R} \preceq \mathscr{P}(\mathbb{Q}) \approx \mathscr{P}(\mathbb{N}) \approx \{ 0, 1 \}^{\mathbb{N}}$. Then $\mathbb{R} \approx \{ 0, 1 \}^{\mathbb{N}}$ will follow by the Schroder-Bernstein Theorem.

To justify the first inequality, observe that an injection $f : \{ 0, 1 \}^{\mathbb{N}} \to \mathbb{R}$ is defined by $f(\sigma) = \Sigma_{n \in \mathbb{N}} {}^{2\sigma(n)}/_{3^n}$ for $\sigma \in \{0,1\}^{\mathbb{N}}$. To justify the second inequality, observe that an injection $g : \mathbb{R} \to \mathscr{P}(\mathbb{Q})$ is defined by $g(x) = (-\infty,x) \cap \mathbb{Q}$ for $x \in \mathbb{R}$. $\mathscr{P}(\mathbb{Q}) \approx \mathscr{P}(\mathbb{N})$ because $\mathbb{Q} \approx \mathbb{N}$. Theorem 0.20 implies $\mathscr{P}(\mathbb{N}) \approx \{ 0, 1 \}^{\mathbb{N}}$. □

Corollary. $\mathbb{R}$ is uncountable.

Theorem 0.23. If X, Y and Z are sets, then $(X^Y)^Z \approx X^{(Y \times Z)}$.

Proof. Define functions $\Phi : (X^Y)^Z \to X^{(Y \times Z)}$ and $\Psi : X^{(Y \times Z)} \to (X^Y)^Z$ as follows. For $f \in (X^Y)^Z$, $\Phi(f)(y,z) = (f(z))(y)$; and for $g \in X^{(Y \times Z)}$, $(\Psi(g)(z))(y) = g(y,z)$. Then $\Psi \circ \Phi = \mathrm{id}$ and $\Phi \circ \Psi = \mathrm{id}$. Hence, $\Phi : (X^Y)^Z \to X^{(Y \times Z)}$ is a bijection. □

Theorem 0.24. $\mathbb{R}^{\mathbb{N}} \approx \mathbb{R}$.

Proof. The preceding theorems justify each equivalence in the following sequence. $\mathbb{R}^{\mathbb{N}} \approx (\{ 0, 1 \}^{\mathbb{N}})^{\mathbb{N}} \approx \{ 0, 1 \}^{(\mathbb{N} \times \mathbb{N})} \approx \{ 0, 1 \}^{\mathbb{N}} \approx \mathbb{R}$. □

Definition. For each set X, let $\mathscr{F}(X) = \{ A \subset X : A \text{ is finite} \}$.

Theorem 0.25. If X is a countable set, then so is $\mathscr{F}(X)$.

Proof. There is an injection $f : X \to \{ p \in \mathbb{N} : p \text{ is prime} \}$. Hence, an injection $g : \mathscr{F}(X) \to \mathbb{N}$ is defined by $g(A) = \prod_{x \in A} f(x)$ for $A \in \mathscr{F}(X)$. □

Definition. For each set X, let $\mathscr{C}(X) = \{ A \subset X : A \text{ is countable} \}$.

Theorem 0.26. $\mathscr{C}(\mathbb{R}) \approx \mathbb{R}$.

Proof. $\mathbb{R} \preceq \mathscr{C}(\mathbb{R})$, because an injection $e : \mathbb{R} \to \mathscr{C}(\mathbb{R})$ is defined by $e(x) = \{x\}$ for $x \in \mathbb{R}$.

Clearly, a surjection $\Phi : \mathbb{R}^{\mathbb{N}} \to \mathscr{C}(\mathbb{R})$ is defined by $\Phi(f) = f(\mathbb{N})$ for $f \in \mathbb{R}^{\mathbb{N}}$. Therefore, Theorem 0.8 b) provides a right inverse for Φ. In other words, there is a function $\Psi : \mathscr{C}(\mathbb{R}) \to \mathbb{R}^{\mathbb{N}}$ such that $\Phi \circ \Psi = \mathrm{id}$. Then Theorem 0.8 a) implies that $\Psi : \mathscr{C}(\mathbb{R}) \to \mathbb{R}^{\mathbb{N}}$ is injective. So $\mathscr{C}(\mathbb{R}) \preceq \mathbb{R}^{\mathbb{N}}$. Hence, by Theorem 0.24, $\mathscr{C}(\mathbb{R}) \preceq \mathbb{R}$.

Finally, the Schroder-Bernstein Theorem implies $\mathscr{C}(\mathbb{R}) \approx \mathbb{R}$. □

D. Vector Spaces, Norms and Inner Products

Definition. A *vector space (over $\mathbb{R}$)* is a set V together with two operations:

vector addition: $(v,w) \mapsto v + w : V \times V \to V$, and

scalar multiplication: $(a,v) \mapsto av : \mathbb{R} \times V \to V$

satisfying the following properties.

a) $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
b) There is an element of V denoted 0 such that $v + 0 = 0 + v = v$ for all $v \in V$.
c) For every $v \in V$, there is an element of V denoted $-v$ such that $v + (-v) = (-v) + v = 0$.
d) $v + w = w + v$ for all $v, w \in V$.
e) $a(v + w) = (av) + (aw)$ for all $a \in \mathbb{R}$ and all $v,w \in V$.
f) $(a + b)v = (av) + (bv)$ for all $a, b \in \mathbb{R}$ and all $v \in V$.
g) $a(bv) = (ab)v$ for all $a, b \in \mathbb{R}$ and all $v \in V$.
h) $1v = v$ for all $v \in V$.

Definition. A subset W of a vector space V is a *vector subspace* of V if W becomes a vector space when the vector addition and scalar multiplication operations on V are restricted to W. Thus, W is a vector subspace of V if and only if W is closed under vector addition and scalar multiplication (i.e., $v, w \in W \Rightarrow v + w \in W$, and $v \in W$ and $a \in \mathbb{R} \Rightarrow av \in W$).

Example. $\mathbb{R}^n$ is a vector space with respect to the following operations:

vector addition: $(x_1, x_2, \ldots , x_n) + (y_1, y_2, \ldots , y_n) = (x_1 + y_1, x_2 + y_2, \ldots , x_n + y_n)$ for $(x_1, x_2, \ldots , x_n), (y_1, y_2, \ldots , y_n) \in \mathbb{R}^n$; and

scalar multiplication: $a(x_1, x_2, \ldots , x_n) = (ax_1, ax_2, \ldots , ax_n)$ for $a \in \mathbb{R}$ and $(x_1, x_2, \ldots , x_n) \in \mathbb{R}^n$.

Example. Let X be a set and V a vector space. Then the set V^X of all functions from X to V is a vector space with respect to the following operations:

vector addition: $(f + g)(x) = f(x) + g(x)$ for $x \in X$, for $f, g \in V^X$; and

scalar multiplication: $(af)(x) = a(f(x))$ for $x \in X$, for $a \in \mathbb{R}$ and $f \in V^X$.

Definition. A *norm* on a vector space V is a function $\| \ \| : V \to [0,\infty)$ satisfying the following properties.

a) $\| v \| = 0 \Leftrightarrow v = 0$ for every $v \in V$.

b) $\| av \| = | a | \| v \|$ for all $a \in \mathbb{R}$ and all $v \in V$.

c) $\| v + w \| \leq \| v \| + \| w \|$ for all $v, w \in V$.

If $\| \ \|$ is a norm on a vector space V, then the pair $(V, \| \ \|)$ is called a *normed vector space.*

Definition. An *inner product* on a vector space V is a function $\langle \ , \ \rangle : V{\times}V \to \mathbb{R}$ with the following properties.

a) $\langle au + bv,w \rangle = a\langle u,w \rangle + b\langle v,w \rangle$ and $\langle u,av + bw \rangle = a\langle u,v \rangle + b\langle u,w \rangle$ for all $a, b \in \mathbb{R}$ and all $u, v, w \in V$.

b) $\langle v,w \rangle = \langle w,v \rangle$ for all $v, w \in V$.

c) $\langle v,v \rangle \geq 0$, and $\langle v,v \rangle = 0 \Leftrightarrow$ if $v = 0$, for all $v \in V$.

If $\langle \ , \ \rangle$ is an inner product on a vector space V, then the pair $(V, \langle \ , \ \rangle)$ is called in *inner product space.*

Theorem 0.27. If $(V, \langle \ , \ \rangle)$ is an inner product space, then a norm $\| \ \|$ is defined on V by the formula $\| v \| = (\langle v,v \rangle)^{1/2}$ for $v \in V$.

Proof. We must prove that $\| \ \|$ satisfies properties a), b) and c) in the definition of a norm.

Exercise. Prove that $\| \ \|$ satisfies properties a) and b).

To prove property c) we need the following fact.

Lemma 0.28. The Cauchy Schwartz Inequality. $|\langle v,w \rangle | \leq \| v \| \| w \|$ for $v, w \in V$.

Proof of Lemma 0.28. If $\langle v,w \rangle = 0$, the inequality is obvious. So assume $\langle v,w \rangle \neq 0$. Then $\| v \| \neq 0 \neq \| w \|$. Set $\varepsilon = | \langle v,w \rangle |/\langle v,w \rangle$. Then $\varepsilon^2 = 1$. So

$$0 \leq \| (\| w \|v - \varepsilon\| v \|w) \|^2 = \langle (\| w \|v - \varepsilon\| v \|w) , (\| w \|v - \varepsilon\| v \|w) \rangle$$

$$= \| w \|^2 \langle v,v \rangle - 2\varepsilon\| v \| \| w \| \langle v,w \rangle + \varepsilon^2\| v \|^2 \langle w,w \rangle$$

$$= 2\| v \|^2 \| w \|^2 - 2\varepsilon\| v \| \| w \| \langle v,w \rangle.$$

Hence, $2\varepsilon\| v \| \| w \| \langle v,w \rangle \leq 2\| v \|^2 \| w \|^2$.

Therefore, $|\langle v,w\rangle| = \varepsilon\langle v,w\rangle \leq \| v \| \| w \|$. □

We now prove property c). Let v, w $\in$ V. Then, using the Cauchy Schwartz inequality, we have

$$\| v + w \|^2 = \langle v + w, v + w\rangle = \langle v,v\rangle + 2\langle v,w\rangle + \langle w,w\rangle = \| v \|^2 + 2\langle v,w\rangle + \| w \|^2 \leq$$

$$\| v \|^2 + 2|\langle v,w\rangle| + \| w \|^2 \leq \| v \|^2 + 2\| v \| \| w \| + \| w \|^2 = (\| v \| + \| w \|)^2.$$

Hence, $\| v + w \| \leq \| v \| + \| w \|$. □

Example. The standard inner product on $\mathbb{R}^n$ is defined by

$$\langle \mathbf{x},\mathbf{y}\rangle = \sum_{i=1}^{n} x_i y_i$$

for $\mathbf{x} = (x_1, x_2, \dots , x_n)$ and $\mathbf{y} = (y_1, y_2, \dots , y_n) \in \mathbb{R}^n$.

Exercise. Verify that the standard inner product on $\mathbb{R}^n$ is an actual inner product.

Example. For $n \geq 1$, we define three norms on $\mathbb{R}^n$.

a) The *taxicab* or *1-norm:* $\| \mathbf{x} \|_1 = \sum_{i=1}^{n} | x_i |$.

b) The *Euclidean* or *2-norm:* $\| \mathbf{x} \|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} = (\langle \mathbf{x},\mathbf{x}\rangle)^{1/2}$.

c) The *supremum* or *∞-norm:* $\| \mathbf{x} \|_\infty = \max \{ | x_i | : 1 \leq i \leq n \}$.

Here, $\mathbf{x} = (x_1, x_2, \dots , x_n) \in \mathbb{R}^n$.

Exercise. Verify that these three formulas actually define norms on $\mathbb{R}^n$. For b), use Theorem 0.27.

Definition. Let X be a set, and let (V, $\| \ \|$) be a normed vector space. Recall that the set V^X of all functions from X to V is a vector space. A function $f \in V^X$ is *bounded* if $\sup \{ \| f(x) \| : x \in X \} < \infty$. Set $B(X,V) = \{ f \in V^X : f \text{ is bounded} \}$. Define the *supremum* or *∞-norm* $\| \ \|_\infty$ on B(X,V) by the formula $\| f \|_\infty = \sup \{ \| f(x) \| : x \in X \}$ for $f \in B(X,V)$.

Exercise. Prove that B(X,V) is a vector subspace of V^X, and verify that the supremum norm $\| \ \|_\infty$ is indeed a norm on B(X,V).