POL502: Foundations

Kosuke Imai
Department of Politics, Princeton University

October 10, 2005

Our first task is to develop the foundations that are necessary for the materials covered in this course.

1 Sets

We start with the notion of sets. An intuitive definition of a set is sufficient for our purposes.

Definition 1 *A set is a collection of objects, and these objects are called elements of the set.*

We also introduce some notations of sets.

Definition 2 *Let A, and B be sets.*

1. *$x \in A$: an element x belongs to A.*
2. *$x \notin A$: an element x does not belong to A.*
3. *$\emptyset$: the empty set has no elements.*
4. *$A = B$: two sets A and B are equal if every $x \in A$ belongs to B and every $y \in B$ belongs to A.*
5. *$A \subset B$: A is a subset of B if $x \in B$ whenever $x \in A$. If $A \subset B$ and $A \neq B$, then A is a proper subset of B.*
6. *$A \cap B = \{x : x \in A \text{ and } x \in B\}$: the intersection of A and B.*
7. *$A \cup B = \{x : x \in A \text{ or } x \in B\}$: the union of A and B.*
8. *$A \setminus B = \{x : x \in A \text{ and } x \notin B\}$: the complement of B relative to A. When all sets under consideration are subsets of some set S, we often write $B^C = S \setminus B$.*

Important sets include some number systems we use.

Example 1 *Four important number systems:*

1. *$\mathbf{N} = \{1, 2, 3, \ldots\}$: the set of natural numbers.*
2. *$\mathbf{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$: the set of integers.*
3. *$\mathbf{Q} = \{p/q : p \text{ and } q \in \mathbf{Z} \text{ and } q \neq 0\}$: the set of rational numbers.*

4. $\mathbf{R}$*: the set of real numbers. What is a real number? We will discuss this question soon.*

Many mathematical objects can be written as a set. An interval is one of them.

Example 2 *Let* $a, b \in \mathbf{R}$.

1. $[a, b] = \{x : x \in \mathbf{R}$ *and* $a \leq x \leq b\}$*: closed interval.*
2. $[a, b) = \{x : x \in \mathbf{R}$ *and* $a \leq x < b\}$*: half closed (half open) interval.*
3. $(a, b] = \{x : x \in \mathbf{R}$ *and* $a < x \leq b\}$*: half closed (half open) interval.*
4. $(a, b) = \{x : x \in \mathbf{R}$ *and* $a < x < b\}$*: open interval.*

In this course, we spend most of time understanding and writing mathematical proofs. The only way to improve our ability to write elegant proofs is to write many proofs on our own! Our first attempt is to prove something that looks very obvious. Note that every proof we do in this course does NOT require additional knowledge you obtained elsewhere.

Theorem 1 (Equality of Sets) *Let* A *and* B *be sets. Then,* $A = B$ *if and only if* $A \subset B$ *and* $B \subset A$.

The next theorem is about various operations of sets. After proving Theorem 1, you should have no difficulty of proving these.

Theorem 2 (Operations of Sets) *Let* $A, B,$ *and* C *be sets. Then,*

1. $(A \cap B) \cap C = A \cap (B \cap C)$.
2. $(A \cup B) \cup C = A \cup (B \cup C)$.
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
5. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.
6. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

An important theorem about set operations is De Morgan's laws.

Theorem 3 (De Morgan's Laws) *Let* A *and* B *be sets.*

1. $(A \cap B)^C = A^C \cup B^C$.
2. $(A \cup B)^C = A^C \cap B^C$.

It is often convenient to create a set of sets that contains other sets as its elements.

Definition 3 *Let* A *be a set and* B_α *be a subset of a set for each* $\alpha \in A$. *Then,* $\{B_\alpha\}_{\alpha \in A}$*, the collection of sets* A_α*, is called an indexed family of subsets with* A *as the index set.*

Example 3 *$\{A_n\}_{n\in\mathbf{N}}$ defined below is an indexed family of sets where $\mathbf{N}$ is an indexed set.*

$$\begin{aligned} A_1 &= \mathbf{N}, \\ A_2 &= \{2,3,4,\ldots\}, \\ A_3 &= \{3,4,5,\ldots\}, \\ &\ldots \\ A_n &= \{n, n+1, n+2, \ldots\} \end{aligned}$$

The notion of the union and intersection can be generalized to an indexed family of subsets.

Definition 4 *Let $\{B_\alpha\}_{\alpha\in A}$ be an indexed family of subsets with the index set A given that for every $\alpha \in A$.*

1. $\bigcap_{\alpha\in A} B_\alpha = \{x : x \in B_\alpha \text{ for every } \alpha \in A\}$.
2. $\bigcup_{\alpha\in A} B_\alpha = \{x : x \in B_\alpha \text{ for some } \alpha \in A\}$.

Example 4 *Let's return to Example 3. We see that $\bigcup_{n=1}^{\infty} A_n = A_1$ and $\bigcap_{n=1}^{\infty} A_n = \emptyset$.*

Example 5 *More examples:*

1. $\bigcup_{n=1}^{\infty} \left(\frac{1}{n}, 1\right) = (0,1)$.
2. $\bigcap_{n=1}^{\infty} \left(0, \frac{1}{n}\right) = \emptyset$.

We can also generalize De Morgan's Laws in Theorem 3

Theorem 4 (Generalized De Morgan's Laws) *Let $\{B_\alpha\}_{\alpha\in A}$ be an indexed family of subsets of some set S. Then,*

1. $\left(\bigcup_{\alpha\in A} B_\alpha\right)^C = \bigcap_{\alpha\in A} B_\alpha^C$.
2. $\left(\bigcap_{\alpha\in A} B_\alpha\right)^C = \bigcup_{\alpha\in A} B_\alpha^C$.

Proof by contradiction Although direct proof is always preferred, one might sometimes have to use an indirect proof, often called *proof by contradiction*. A direct proof begins with the hypothesis of the theorem and shows how the hypothesis logically leads to the conclusion. An indirect proof starts with the negation of what we would like to prove and then shows how this leads to the contradiction of hypotheses or other accepted facts.

Proof by induction Another important technique often used for mathematical proof is induction. If $S(n)$ is a statement containing the variable $n \in \mathbf{N}$ such that $S(1)$ is a true statement, and for any $k \in \mathbf{N}$ if $S(k)$ is true, then $S(k+1)$ is also true. This implies that $S(n)$ is true for all n.

Example 6 *Some examples using these proof techniques*

1. *There is no rational number whose square is 2.*
2. *$n^3 + 5n$ is divisible by 6 for each $n \in \mathbf{N}$.*
3. *For any $n \in \mathbf{N}$, $1 + x + x^2 + \ldots + x^n = (1 - x^{n+1})/(1 - x)$.*

2 Functions

In the previous section, we said that many mathematical objects are sets. Here, we use sets to define functions. Before we give the definition of functions, we define ordered pairs. So far, we did not care about the order of elements of a set. However, in many situations the order in which elements appear is important.

Definition 5 *The ordered pair (x, y) is the set $\{x, \{x, y\}\}$ where x and y are referred to as the first and second coordinate, respectively.*

From this definition, it is clear that $(x, y) = (z, w)$ if and only if $x = z$ and $y = w$. This naturally leads to the construction of the Cartesian product, which is defined as a collection of ordered pairs.

Definition 6 *The Cartesian product of two sets X and Y, denoted as $X \times Y$, is the set of all ordered pairs (x, y) with $x \in X$ and $y \in Y$.*

An example may help to understand this definition.

Example 7 *Let $X = \{1, 2\}$ and $Y = \{2, 4\}$. Then $X \times Y = \{(1, 2), (1, 4), (2, 2), (2, 4)\}$.*

Before defining a function, we consider a broader concept that includes functions as a special case; i.e., a relation from a set to another.

Definition 7 *Let X and Y be sets. A relation R from X to Y is a subset of $X \times Y$. We often write xRy if $(x, y) \in R$. The converse of a relation R is $R^{-1} = \{(x, y) : (y, x) \in R\}$.*

Now, we are ready to give the definition of a function. A function is basically a mapping from one set to another. It takes an element from one set as an input and produces an element as an output which belongs to another set. In the language of ordered pairs, the first coordinate is an input of a function and the second coordinate is its output.

Definition 8 *Let X and Y be sets. A function $f : X \mapsto Y$ is a relation from its domain X to its codomain Y. That is, for each $x \in X$, $(x, y_1) \in f$ and $(x, y_2) \in f$ imply $y_1 = y_2 = f(x)$. The set $f(X) = \{f(x) : x \in X\}$ is called the image (or range) of f.*

It is important to note that the image of a function is not necessarily equal to its codomain. There are different types of functions. We first give some basic categories based on this definition.

Definition 9 *Let $f : X \mapsto Y$ be a function. Then,*

1. *f is onto (or surjective) if $Y = f(X)$.*
2. *f is one-to-one (or injective) if $f(x_1) = f(x_2)$ implies $x_1 = x_2$.*
3. *f is bijective if it is onto and one-to-one.*

In words, a function is onto if the image of the function equals to its codomain. An one-to-one function defines an one-to-one relationship between all the elements of the domain and their counterparts of the codomain. Let's see some examples to make sure we understand this definition.

Example 8 *Three simple examples:*

1. *The function* $f : \mathbf{Z} \mapsto \mathbf{Z}$ *defined as* $f(x) = x^3$ *is one-to-one, but not onto.*
2. *The function* $g : \mathbf{Z} \mapsto \{0, 1, 2, 3, \ldots\}$ *defined as* $g(x) = |x|$ *is onto but not one-to-one.*
3. *The function* $h : \mathbf{R} \mapsto \mathbf{R}$ *defined as* $h(x) = 2x - 1$ *is bijective.*

Next, we consider two types of functions that we encounter often in practice. The first is an inverse function.

Theorem 5 (Inverse Functions) *Let* f *be a function. The converse of* f*, denoted by* f^{-1}*, is a function if and only if* f *is one-to-one. We call* f^{-1} *the inverse of* f*.*

This theorem also implies that if f^{-1} is a function, then its domain and image are equal to the image and domain of f, and that if f is bijective, so it f^{-1}.

Example 9 *Let's find inverse functions of the functions in Example 8.*

1. $f^{-1}(x) = \sqrt[3]{x}$.
2. $g^{-1}(x)$ *does not exist.*
3. $h^{-1}(x) = \frac{x+1}{2}$.

Another useful type of functions is a composite function, which nests one function within another,

Definition 10 *Let* X *and* Y *be sets. If* $f : X \mapsto Y$ *and* $g : Y \mapsto Z$*, then the composition of* g *by* f*, denoted by* $g \circ f$*, is the set,* $\{(x, z) : \exists\, y \in Y$ *such that* $(x, y) \in f$ *and* $(y, z) \in g\}$*.*

Example 10 *Again, we return to Example 8.*

1. $g \circ f = |x^3|$.
2. $h \circ g = 2|x| - 1$.
3. $f \circ h = (2x - 1)^3$.

We end this section with the proof of some theorems about composite functions.

Theorem 6 (Composite Functions) *Let* $f : X \mapsto Y$ *and* $g : Y \mapsto Z$*.*

1. *Then,* $g \circ f$ *is a function and* $g \circ f : X \mapsto Z$*.*
2. *If* f *and* g *are onto, then* $g \circ f$ *is also onto.*
3. *If* f *and* g *are one-to-one, then* $g \circ f$ *is also one-to-one.*

3 Real Numbers

In this section, we study the real number system, which is the main number system we use in this course. Before getting to the details, recall that in Example 6 we proved that there is no rational number whose square is 2. This implies that there are holes in the rational number system. That is, the rational number system cannot cover all the numbers. This motivates us to define the real number system, which is more "dense" than the rational number system. Rather than constructing the real number system (which we can if we want to), we assume its existence with the following properties.

Axiom 1 (Real Numbers) *The set of real numbers,* $\mathbf{R}$*, has two functions,* $+ : \mathbf{R} \times \mathbf{R} \mapsto \mathbf{R}$ *and* $\cdot : \mathbf{R} \times \mathbf{R} \mapsto \mathbf{R}$*, called addition and multiplication, as well as a relation* $<$ *such that for all* $a, b, c \in \mathbf{R}$*,*

1. *(associative laws)* $(a + b) + c = a + (b + c)$ *and* $(ab)c = a(bc)$*.*
2. *(commutative laws)* $a + b = b + a$ *and* $ab = ba$*.*
3. *(distributive laws)* $a(b + c) = ab + ac$*.*
4. *(existence of 0) There exists a unique element* $0 \in \mathbf{R}$ *such that* $a + 0 = a$ *for all* $a \in \mathbf{R}$*.*
5. *For all* $a \in \mathbf{R}$*, there exists a unique element* $b \in \mathbf{R}$ *such that* $a + b = 0$*, and we write* $b = -a$*.*
6. *(existence of 1) There exists a unique element* $1 \in \mathbf{R}$ *such that* $a1 = a$ *for all* $a \in \mathbf{R}$*.*
7. *For all* $a \in \mathbf{R}$ *with* $a \neq 0$*, there exists a unique element* $b \in \mathbf{R}$ *such that* $ab = 1$*, and we write* $b = a^{-1}$ *or* $b = \frac{1}{a}$*.*
8. *For any* $a, b \in \mathbf{R}$*, either* $a < b$*,* $b < a$*, or* $a = b$*.*
9. *If* $a < b$*, then* $a + c < b + c$ *for all* $c \in \mathbf{R}$*.*
10. *If* $a < b$ *and* $b < c$*, then* $a < c$*.*
11. *If* $a < b$ *and* $0 < c$*, then* $ac < bc$*.*

Given these axioms of the real number system, let's prove its elementary properties.

Theorem 7 (Real Numbers) *Let* $a, b, c \in \mathbf{R}$*. Then,*

1. *If* $a < b$*, then* $-b < -a$*.*
2. $0 < 1$*.*
3. *If* $0 < a < b$*, then* $0 < \frac{1}{b} < \frac{1}{a}$*.*
4. *If* $a < b$ *and* $c < 0$*, then* $bc < ac$*.*
5. $0 \leq a^2$*.*

Very important concepts of the real number system are those of maximum/minimum and (least) upper / (greatest) lower bounds.

Definition 11 *Let S be a subset of $\mathbf{R}$. Then,*

1. *$m \in \mathbf{R}$ is an upper bound for S if $a \leq m$ for every $a \in S$.*
2. *$m \in \mathbf{R}$ is a lower bound for S if $m \leq a$ for every $a \in S$.*
3. *$m \in \mathbf{R}$ is a maximal element of S if $m \in S$ and there exists no $a \in S$ such that $m < a$.*
4. *$m \in \mathbf{R}$ is a minimal element of S if $m \in S$ and there exists no $a \in S$ such that $a < m$.*
5. *$m \in \mathbf{R}$ is the least upper bound for S if m is an upper bound for S and $m \leq m'$ for any upper bound m' for S. We write $m = \sup S$ and call it the supremum of S.*
6. *$m \in \mathbf{R}$ is the greatest lower bound for S if m is a lower bound for S and $m' \leq m$ for any lower bound m' for S. We write $m = \inf S$ and call it the infinimum of S.*

The definition implies that although a set can have many upper bounds, it can have only one least upper bound. To make sure we understand these definitions, we again return to the examples.

Example 11 *Consider the following three sets.*

1. *$A = \{\frac{1}{n} : n \in \mathbf{N}\}$.*
2. *$B = (0, 1)$.*
3. *$C = [0, 1]$.*

All three sets have 1 as the supremum and 0 as the infinimum. However, A and B do not have the maximum or minimum. This means that the supremum and infinimum of a set does not necessarily belong to that set while the maximum and minimum always do.

Now, we are ready to investigate the key properties about the real number system. The main difference between the real numbers and the rational numbers is that the former does not contain the gap like the one the latter has. That is, the real numbers are complete.

Axiom 2 (Completeness) *If $A \subset \mathbf{R}$ is a non-empty set and is bounded from above, then A has a least upper bound.*

Although it is clear that we do not always have the maximum, this axiom is somewhat difficult to understand.

Example 12 *Consider the set, $\{q \in \mathbf{Q} : q^2 < 2\}$. Does this set have the least upper bound in $\mathbf{Q}$?*

The completeness property works with a least lower bound, too.

Theorem 8 (Completeness) *If $A \subset \mathbf{R}$ is a non-empty set and is bounded from below, then A has a greatest lower bound.*

Another look at the least upper bound is the following theorem

Example 13 (Least Upper Bound) *Let $m \in \mathbf{R}$ is an upper bound for a set $A \subset \mathbf{R}$. Then, $m = \sup A$ if and only if for every $\epsilon > 0$ there exists an element $a \in A$ such that $m - \epsilon < a$.*

We now study the consequences of this completeness property. The next two theorems tells you how $\mathbf{N}$ and $\mathbf{Q}$ sit inside of $\mathbf{R}$. These are the key properties of the real numbers. The next theorem says that $\mathbf{N}$ is not bounded above.

Theorem 9 (Archimedian Property) *For any $a \in \mathbf{R}$ with $a > 0$,*

1. *there exists $n \in \mathbf{N}$ such that $n > a$.*
2. *there exists $n \in \mathbf{N}$ such that $\frac{1}{n} < a$.*

Before proving that $\mathbf{Q}$ is dense in $\mathbf{R}$, we define the formal meaning of "dense". Intuitively speaking, for any rational number there are real numbers that are arbitrarily close to it. This means that the real numbers can fill in the holes that exist in the rational number system.

Definition 12 *Let $S \subset \mathbf{R}$ be a set. Then, S is dense in $\mathbf{R}$ if for any $a \in \mathbf{R}$ there exists $b \in S$ such that $a - \epsilon < b < a + \epsilon$ for any $\epsilon \in \mathbf{R}$ with $\epsilon > 0$.*

Theorem 10 (Density of Q in R) *For any $a, b \in \mathbf{R}$ with $a < b$, there exists $r \in \mathbf{Q}$ such that $a < r < b$.*

Earlier, we saw that $\sqrt{2}$ is not a rational number. Now we prove that it is a real number.

Theorem 11 (Existence of $\sqrt{2}$) *There exists $a \in \mathbf{R}$ such that $a^2 = 2$.*

We end this section with an important function on $\mathbf{R}$.

Theorem 12 (Absolute Value) *For any $x \in \mathbf{R}$, define the function $|x|$ as follows*

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

Then,

1. *$|ab| = |a||b|$.*
2. *If $\epsilon > 0$, then $|a| \leq \epsilon$ if and only if $-\epsilon \leq a \leq \epsilon$.*
3. *(triangle inequality) $|a + b| \leq |a| + |b|$.*

4 Countable and Uncountable Sets

As the final topic of the chapter on foundations, we study the size of a set, called *cardinality*. Intuitively, the cardinality of a finite set can be computed by attaching a natural number to the set. For example, the cardinality of $A = \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\}$ is 3. Such a set is called "countable" and indeed all finite sets are countable. But, how about infinite sets? Before answering this question, let's formalize the distinction between finite and infinite sets.

Definition 13 *Let X be a set.*

1. *X is finite if X is empty or there exists a bijective function from $\{1, 2, \ldots, n\}$ to X for some $n \in \mathbf{N}$.*
2. *X is infinite if it is not finite.*

Note that n is called the cardinality of the finite set X. Remind ourselves that a bijective function is a function which is onto as well as one-to-one. Our intuition about countability can then be extended to infinite sets

Definition 14 *Let X be an infinite set.*

1. *X is countable if there exists a bijective function from $\mathbf{N}$ to X.*
2. *X is uncountable if it is not countable.*

This implies that a set is countable if it is either finite or countably infinite. To get used to this idea, let's first look at an example of countable sets.

Example 14 *The set of even natural numbers is countable.*

Some key theorems about countable sets are

Theorem 13 (Coutable Sets) *Two key statements about countable sets.*

1. *Every subset of $\mathbf{N}$ is countable.*
2. *Every subset of a countable set is countable.*

Now, we prove the fact that the set of rational numbers is countable.

Theorem 14 (Countability of Q) **Q** *is countable.*

As we expect, the set of real numbers, unlike the set of rational numbers, is uncountable.

Theorem 15 (Uncountability of R) **R** *is uncountable.*

These theorems suggest that there are far more irrational numbers than rational numbers. Recall that if we use decimal expansion to express rational numbers, it recurs, e.g,. $\frac{1}{7} = 1428571428\ldots$. Now suppose that we pick a number between 0 and 1 by rolling a ten-sided die independently for each decimal. How likely do you think the resulting decimal expansion is to recur? Now, you should not be surprised about the fact that the real number system has a lot more numbers than the rational number system.