VII. SETS, LOGIC, PROOFS, AND RELATIONS

If you continue on to grad school, you will do a lot of proofs. If you haven’t had a proof-based math class before, these will seem weird and awkward. They are simply a way of making an argument, just as essays are. Like essays, good proofs use a certain language and follow a certain formula. You will see these logical symbols:

\[ \therefore \text{ since} \quad \therefore \text{ therefore} \]
\[ \land \text{ and} \quad \lor \text{ or} \]
\[ \neg \text{ not} \quad \Rightarrow \text{ implies} \]
\[ \leftrightarrow \text{ is equivalent to} \quad \iff \text{ contradiction} \]

For the most part, I discourage the “since” and “therefore” symbols because they’re confusing (at least for me). I’ve never been able to find a way to keep them straight, which one means what. There are other symbols to express all of these (including about fourteen different ways to express a contradiction). I also dislike using the symbols \( \land \) and \( \lor \) for “or” and “and,” since they are also used to represent the minimum and the maximum of two things, or the join and meet of two partitionings. In logic, their meaning is clear, but outside of that, there may be some ambiguity.

An useless piece of trivia is that the symbol \( \lor \) comes from the first letter of the Latin conjunction \( vel \), which is an inclusive or. In logic, “Or” is always used in the inclusive sense: saying that “all first year grad students take prob/stats or metrics” does not rule out that some might take both. It only rules out people taking neither. In fact, the negation of “and” and “or” are:

\[ \neg (P \land Q) \iff (\neg P) \lor (\neg Q) \]
\[ \neg (P \lor Q) \iff (\neg P) \land (\neg Q) \]

A proposition is a statement that is either true or false. It might be like “the moon is in the seventh house.” You can go consult an astrology chart and verify whether it is. A conditional statement like “if the moon is in the seventh house and Jupiter aligns with Mars, then peace will rule the planets” consists of a hypothesis and a conclusion. This is often written \( P \Rightarrow Q \) and is read “\( P \) implies \( Q \)” or “\( P \) is sufficient for \( Q \).” The converse of a conditional is \( Q \Rightarrow P \) or \( P \Leftarrow Q \), and does not need to be true even if the original statement is. The converse is often read “\( P \) only if \( Q \)” or “\( P \) is necessary for \( Q \).” The contrapositive \( (\neg Q) \Rightarrow (\neg P) \) is. If a conditional and its converse are both true, the statement is known as a biconditional statement, \( P \iff Q \). The biconditional is sometimes written “\( P \) iff \( Q \)” and often read “\( P \) if and only if \( Q \)” or “\( P \) is necessary and sufficient for \( Q \).”

It hard to come up with a better definition of a set than just “a collection of things.” These are called elements or members. Common set notation includes:
\( \emptyset \) the empty set
\( \forall \) “for all” or “for any” or “for each” or “for every”
\( \exists \) “there exists” or “for some”
\( \in \) is an element of the set
\( \subseteq \) is a subset of the set
\( \subset \) is a (proper) subset of the set

(Many people intend the last of these to allow weak inclusion as well.) For one set \( A \) to be a subset of another set \( B \) means that every element of \( A \) is also an element of \( B \). If every element in \( A \) is also in \( B \), and \( B \) elements that are not in \( A \), then \( A \) is a proper subset of \( B \). Alternatively, we can say that \( A \) is contained in \( B \) when it is a subset and properly contained or strictly contained when it is a proper subset. When two sets are each a (weak) subset of each other, then they are equivalent, denoted with an equal sign. The complement of a set \( A \) is everything in the universe that is not in \( A \), usually denoted by \( A' \) or \( A^c \). The cardinality of a set is the number of elements in it, usually denoted by \( |A| \) or \( |A| \). (Think of the set of all people in the classroom; the cardinality of this set is the number of people in the classroom.) The set without any elements is the empty set, denoted with the nifty letter \( \emptyset \). Two sets are disjoint when their intersection is the empty set; that is, when they share no elements.

Elements of a set can often be characterized by the common features that they share, at the exclusion of all elements not in the set, like “the set of all squirrels in the classroom” or “the set of all vegetables that have a purple exterior.” There is a simple notation for expressing this:

\[
\{ x \in \mathbb{R}_+ : 10x \in \mathbb{N} \}
\]

This is much easier than actually naming all nonnegative real numbers that are some integer divided by ten, a list which starts with 0, 0.1, 0.2, 0.3, 0.4, and goes on for quite a while.

Recall the definition of a function continuous at some \( x \): for any \( \varepsilon > 0 \) there exists a \( \delta > 0 \) such that if a point \( y \) is less than distance \( \delta \) from \( x \), then \( f(y) - f(x) \) is less than \( \varepsilon \). In mathematical shorthand, this becomes:

\[
(\forall \varepsilon > 0)(\exists \delta > 0) \text{ s.t. } \left[ |y - x| < \delta \Rightarrow \left| f(y) - f(x) \right| < \varepsilon \right]
\]

This is much more compact though harder to read.

**Example:** The utility function \( U_i : X_i \rightarrow \mathbb{R} \) is **locally non-satiated** if \( \forall x_i \in X_i \) and \( \forall \varepsilon > 0 \), \( \exists x'_i \in X_i \) s.t. \( \left\| x'_i - x_i \right\| \leq \varepsilon \) and \( U_i(x'_i) > U_i(x_i) \).
Example: An individual's demand correspondence is defined as:

\[ x_i(p, w_i) \equiv \left\{ x \in X_i : U_i(x) \geq U_i(\hat{x}) \quad \forall \hat{x} \in B(p, w_i) \right\} \]

where his budget correspondence is

\[ B(p, w_i) \equiv \left\{ x \in X_i, X_i \subseteq \mathbb{R}^L_+ : \sum_{l=1}^L x_{il}p_l \leq w_i \right\} . \]

Can you translate this into words, and then, can you translate it into something that makes sense?

There is some connection between the logic above and sets. If \( A \) is the set of events for which statement \( P \) is true (for example, the set of days on which the moon is in the seventh house), and \( B \) is the set of events for which statement \( Q \) is true, then we have the following relationships:

\[
\begin{align*}
(P \Rightarrow Q) & \iff (A \subseteq B) \\
(P \Leftarrow Q) & \iff (A \supseteq B) \\
(P \iff Q) & \iff (A = B) \\
(\neg P) & \iff (A^C) \\
(P \land Q) & \iff (A \cap B) \\
(P \lor Q) & \iff (A \cup B)
\end{align*}
\]

Often, these relationships are expressed using a Venn diagram.

In economics, we are often asked to provide a formal proof of some assertion. Writing a proof is like writing an essay. There is a “proper” way to do each, there are several standard formats, and there are stylistic details which make your argument easier to follow.

Generally, this is the format I like to follow for a proof. On the first line, I state exactly what I will be proving. On the next line, I begin my proof. In the body of the proof, I first announce the method of the proof (if nothing is stated, direct proof is understood). Then I state my suppositions, followed by relevant definitions. This is like an introduction in an essay. Having clearly spelled out the background, I give the argument.

Ultimately, the proof comes to some “punchline.” In a direct proof, this is the conclusion; in a proof by contradiction, it is the critical contradiction and its implication. This is like the conclusion of the essay. Traditionally, this is followed by a black box or the letters \( \text{QED} \) to announce that the proof is complete.

A direct proof is the most basic way to show “if \( P \), then \( Q \).” You start off by supposing that \( P \) is true, and then you show the logical sequence that implies \( Q \) must also be true. To do some of the following proofs, I first need to define some terms:
Definition: An integer \( n \in \mathbb{Z} \) is **even** if there exists some \( k \in \mathbb{Z} \) such that \( n = 2k \). An integer \( n \in \mathbb{Z} \) is **odd** if there exists some \( k \in \mathbb{Z} \) such that \( n = 2k + 1 \).

Property: If \( m, n \in \mathbb{Z} \), then \( (m + n) \in \mathbb{Z} \) and \((m \cdot n) \in \mathbb{Z}\). (The set of integers is closed under addition and multiplication.)

Now for our first formal proofs.

Prove: If \( m, n \in \mathbb{Z} \), then \( (m - n) \in \mathbb{Z} \).

Proof: Suppose that \( m \) and \( n \) are both integers. Since the set of integers is closed under multiplication and since \(-1 \in \mathbb{Z}\), it must be the case that \(-n = (-1) \cdot n \) is also an integer. Since the set of integers is closed under addition and since \( m \) and \(-n \) are integers, then \( m + (-n) = m - n \) must also be an integer. QED.

Prove: If \( n \) is an odd integer, then \( n^2 \) is odd.

Proof: Suppose that \( n \) is an odd integer. By definition of odd, this means that \( \exists k \in \mathbb{Z} \) s.t. \( n = 2k + 1 \). Then \( n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 \). Set \( m = 2k^2 + 2k \). We know that \( m \) is an integer, since \( \mathbb{Z} \) is closed under multiplication and addition. Thus, \( n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2m + 1 \) for \( m \in \mathbb{Z} \), and so by definition, \( n^2 \) is odd. QED.

A **proof by contradiction** or a **proof by contrapositive** is sometimes the easiest way to show how “\( P \) implies \( Q \).” You start by supposing that \( P \) is true but \( Q \) is not, and show how this leads to two conclusions which are mutually exclusive. These proofs often start with the statement, “The proof is by contradiction,” or, “Suppose not.” When you arrive at the two contradictory statements, denote that this is an impossibility with one of several common symbols or the words “which is a contradiction.” Then you state that this contradiction means that \( P \) must imply \( Q \).

Prove: If \( r \) is a real number such that \( r^2 = 2 \), then \( r \) is irrational.

Proof: Suppose not: suppose \( r \) is rational and \( r^2 = 2 \). By definition, \( r \) is rational if \( \exists m, n \in \mathbb{Z} \) such that \( r = m / n \). We can assume without loss of generality that \( m \) and \( n \) have no common divisor greater than one (if they did, we could simply find \( \exists m' \in \mathbb{Z} : m' = m / \gcd(m, n) \), similarly \( n' \), to reduce the fraction). Then \( r = m / n \Rightarrow r^2 = m^2 / n^2 \Rightarrow m^2 = r^2 n^2 = 2n^2 \). Since \( m^2 \) is even, we know that \( m \) is even (as shown in the previous proof). Then by definition, \( \exists k \in \mathbb{Z} \) s.t. \( m = 2k \). Then \( m^2 = 4k^2 = 2n^2 \Rightarrow 2k^2 = n^2 \). Thus, \( n^2 \) is even and as a consequence, \( n \) is even. Then two divides both \( m \) and \( n \), but they have no common divisor greater than one. This is a contradiction. Therefore, when \( r^2 = 2 \), it must be the case that \( r \) is irrational. QED.

A third method is **proof by induction**. This works extremely well in certain cases, generally, when the assertion looks something like “\( P \) implies \( Q \) for all \( n \) greater than
or equal to \( N \).” There are two steps here. First, show that “\( P \) implies \( Q \) for \( N \).” Then you just assume that “\( (P \Rightarrow Q) \) is true for some \( n \geq N \)” and show that this leads to “\( (P \Rightarrow Q) \) must be true for \( n + 1 \).” It works just like knocking over dominoes: if the \( n \)-th domino falls, it pushes over the \((n + 1)\)-th domino, provided that you started the chain at the \( N \)-th one.

These proofs often start with the words, “The proof is by induction.” They announce the two steps. Finally, they state that the assertion has been shown for the first case, and that one case implies the next, and thus is must be true for all cases greater than the first.

**Prove:** Let \( x \in \mathbb{R}, x \neq 1 \). Then \( \forall n \in \mathbb{Z}_{++}, \sum_{k=1}^{n} x^{k-1} = \frac{x^n - 1}{x - 1} \).

**Proof:** The proof is by induction. Let us take any \( x \in \mathbb{R}, x \neq 1 \). In the first step, we want to show that \( \sum_{k=1}^{n} (x^{k-1}) = (x^n - 1)/(x - 1) \) for \( n = 1 \). The left-hand side of this is \( \sum_{k=1}^{1} x^{k-1} = x^0 = 1 \). The right-hand side is \((x^1 - 1)/(x - 1) = 1 \). Thus, it is true when \( n = 1 \).

Now suppose that \( \sum_{k=1}^{n} (x^{k-1}) = (x^n - 1)/(x - 1) \) is true for some arbitrary \( n \in \mathbb{Z}_{++} \). We want to show that this implies that the relationship also holds for \( n + 1 \). For \( n + 1 \),

\[
\sum_{k=1}^{n+1} (x^{k-1}) = x^n + \sum_{k=1}^{n} (x^{k-1}) = x^n + (x^n - 1)/(x - 1) = x(x^n - 1)/(x - 1) + (x^n - 1)/(x - 1) = (x^{n+1} - x^n)/(x - 1)
\]

This is what we needed to show: \( \sum_{k=1}^{n+1} (x^{k-1}) = (x^{n+1} - 1)/(x - 1) \). Since \( \sum_{k=1}^{n} (x^{k-1}) = (x^n - 1)/(x - 1) \) for \( n = 1 \), and since \( \sum_{k=1}^{n} (x^{k-1}) = (x^n - 1)/(x - 1) \Rightarrow \sum_{k=1}^{n+1} (x^{k-1}) = (x^{n+1} - 1)/(x - 1) \), we can conclude that \( \forall n \in \mathbb{Z}_{++}, \sum_{k=1}^{n} (x^{k-1}) = (x^n - 1)/(x - 1) \). \( \Box \)

This relationship shows up in discounting future payoffs—now you’ll know the derivation when you see it. Here is the sort of thing that you’ll be asked to prove in micro classes this next year, and what a good proof looks like:

**Prove:** If \( x(p, w) \) is homogeneous of degree one with respect to \( w \) and satisfies Walras’ law (ignore this for now!), then \( \forall \ell \in L \) the elasticity of demand with respect to wealth, \( \varepsilon_{\ell w}(p, w) \), equals one.

**Proof:** Suppose that the demand function \( x(p, w) \) is homogenous of degree one with respect to \( w \) and satisfies Walras’ law. Income elasticity for good \( \ell \) is defined as \( \varepsilon_{\ell w}(p, w) \equiv \frac{\partial x_{\ell}(p, w)}{\partial w} \cdot \frac{w}{x_{\ell}(p, w)} \). By definition of homogeneity, \( \forall \alpha > 0, x(p, \alpha w) = \alpha x(p, w) \).

Differentiating with respect to \( \alpha \),

\[
\frac{\partial}{\partial \alpha} x(p, \alpha w) = \frac{\partial}{\partial \alpha} \alpha x(p, w) = \alpha \frac{\partial x(p, w)}{\partial w} = x(p, w).
\]

Evaluating this at \( \alpha = 1 \), for each good \( \forall \ell \in L \frac{\partial x_{\ell}(p, w)}{\partial w} = \frac{x_{\ell}(p, w)}{w} \Rightarrow \frac{\partial x_{\ell}(p, w)}{\partial w} \cdot \frac{w}{x_{\ell}(p, w)} = 1 \). Thus, \( \varepsilon_{\ell w}(p, w) = 1 \). \( \Box \)
In your graduate-level micro, you will work with preference relations before you are allowed to play with utility functions. For some set $X$, a **binary relation** gives you a true or false statement about any pair $(x, y)$, with $x \in X$ and $y \in Y$. For example, think about a group of five kids (Annabel, Beladonna, Clarabel, Isabel, and Gargamel). “Person $x$ liking person $y$” is a binary relation: you can pull out any pair, such as $(Annabel, Clarabel)$, and ask: “does Annabel like Clarabel?” The answer is yes or no. The order of the pair $(x, y)$ matters: this might not be the same as the answer to “does Clarabel like Annabel?” Similarly, you can pull out the pair $(Clarabel, Gargamel)$ or $(Isabel, Isabel)$ or any others, and the relation gives you an answer to all of these.

Because each of these statements is a true or false proposition, sometimes a relation will be called a subset of $X \times X$. The set $X \times X$ is all possible pairs from $X$, and the relation is the subset on which this is true.

On a set of names, a possible relation is “$x$ comes earlier than or at the same place as $y$ in alphabetic order.” On a set of potential bundles of goods, a possible relation is “I like $x$ at least as much as I like as $y$. The last of these is a **preference relation**. Here is almost the right definition of a nice property of some orderings:

**Definition:** A binary relation $R$ on a set $X$ is (close to) a **partial ordering** if, $\forall x, y, z \in X$, the following three properties hold:

1. **reflexive:** $(x, x) \in R$
2. **complete:** $(x, y) \in R \lor (y, x) \in R$
3. **transitive:** $[(x, y) \in R \land (y, z) \in R] \Rightarrow (x, z) \in R$.

You can verify that the alphabetic order rule above does in fact fit the definition of a partial ordering. (Something very similar to alphabetical order will arise later under the name of the **lexicographic ordering**.)

When we have a partial ordering, it is more intuitive to write $x \succeq y$ for $(x, y) \in R$. This makes sense when you compare it to a certain partial ordering on the real numbers, the “greater than or equal to” ordering.

**Definition:** A partial ordering $R$ on a set $X$ is a **well-ordering** if every nonempty subset $S$ has a first element $s$ in $X$; that is, $(x, s) \in R \forall x \in A$.

What this gets at is the problem that in some sets it’s very hard to name exactly what the largest or smallest element is. Under a few circumstances it is:

**Principle:** Any partial ordering on a finite set is a well-ordering.
This relates to a topic mentioned previously. The relations ≤ and ≥ are partial orderings on the real numbers. The real numbers are not well-ordered under these relations, though. As mentioned, not all sets have a maximum and a minimum. However, all bounded sets have upper bounds (who would have thought?); that is, numbers that are above everything in the set. They also have lower bounds. The \textbf{least upper bound} of a set, also known as a \textbf{supremum}, is the smallest number with the property that it is greater than or equal to any element in the set. Unlike a maximum, it does not need to be in the set—however, if the set has a maximum, it is also the supremum. On the other hand, when a set doesn’t have a maximum, like the interval \((0,1)\), the supremum is the next best thing. The \textbf{greatest lower bound} or \textbf{infimum} of a set is the largest number such that it is less than or equal to anything in the set. If a minimum exists, it is the infimum.

Suprema of sets are denoted in one of these ways:

\[
\sup\{x,y,z\} \quad \text{or:} \quad \lub\{y,x,z\}
\]

while infima are represented as:

\[
\inf\{x,y,z\} \quad \text{or:} \quad \glb\{x,y,z\}
\]