Proof and Sizes of Infinity

Rich Schwartz

September 10, 2014

1 What is a Proof?

Assuming that you have accepted some commonly chosen foundations of mathematics, such as first order logic, Zermelo-Frankel set-theory, and the language needed to discuss these things, the notion of a mathematical proof has a precise meaning. One argues straight from the axioms using allowable logical deductions.

Here is an example taken from the last handout.

Theorem 1.1 *No set is a member of itself.*

Proof: Suppose (for the sake of contradiction) there is a set S such that $S \in S$. (This notation means that S is a member of S.) By the Axiom of Pairing, there is a set $\{S\}$ whose only member is S. But then the Axiom of Foundation says that there is a member of $\{S\}$ which has no members in common with $\{S\}$. The only member to choose from is S, so S has no members in common with S. But if $S \in S$, then S is a member common to both S and $\{S\}$. This is a contradiction. This contradiction shows that the situation $S \in S$ is impossible. ♠

Not all mathematical proofs will have this format, but this one has many classic elements. The proof is an *argument by contradiction*. You have some statement that you want to establish as true. So, you assume for the sake of contradiction that the statement is false. Then you derive logical consequences from your assumption and from the axioms. You eventually arrive at a contradiction, and this means that your statement must be true.

In practice, mathematical proofs do not go all the way back to the axioms. What happens is that we have built up a store of constructions and statements and results which follow from the axioms. An example would be

Theorem 1.2 *If a and b are any integers, then $a + b = b + a$.*

The proof is some long chain of logical deductions from the axioms. But the significant thing is that the proof *has already been done by someone else and accepted by the community*. So, in practice (except perhaps if you are taking a set theory class and the point is to test you on your knowledge of this result) you could use this theorem with the same certainty as the rest of the axioms. The theorems proved from the axioms inherit the same certainty that the original axioms have and then in subsequent mathematics you can use these new results as axioms.

2 Some Sample Theorems

Here are some proofs which do not go all the way back to the axioms.

Theorem 2.1 *Every integer greater than* 1 *is the product of finitely many primes.*

Proof: If this result is false, there is some smallest integer N which is not the product of primes. (This statement is a version of the Principle of Induction, which is derived along with the existence of the natural numbers.) Either N is prime or N is not prime.

- Suppose that N is prime. Well, then N is a product of primes. Namely, $N = N$. This is a contradiction.
- Suppose that N is not prime. Then, by definition, $N = AB$ where A and B are two other integers, both greater than 1. By properties of integers derived from the axioms $A < N$ and $B < N$. Since N is the smallest "problem integer", we know that both A and B are the product of primes. That is $A = a_1, ..., a_m$ and $B = b_1, ..., b_n$. But then $N = a_1...a_m b_1...b_n$. We have expressed N as the product of primes, and so we again have a contradiction.

Since both cases lead to a contradiction, we see that no such N exists. That completes the proof. ♠

Theorem 2.2 *There is no rational number p/q such that $(p/q)^2 = 2$.*

Proof: Suppose for the sake of contradiction that there is such a rational number. Let p/q be an example. We can assume that p/q is a reduced fraction (there are no common factors) so that in particular at least one of p or q is odd. We compute

$$(p/q)^2 = p^2/q^2 = 2.$$

Multiplying through by q^2, we get

$$p^2 = 2q^2.$$

This means that p^2 is even. But then p is even as well. Since p is even, we can write $p = 2r$ for some other integer r. But then

$$2q^2 = p^2 = (2r)^2 = 4r^2.$$

Dividing by 2 gives

$$q^2 = 2r^2.$$

But then q^2 is even. Therefore q is even as well. We've proved that both p and q are even, and this is a contradiction to our assumption that p and q are not both even. The only way out of the contradiction is that the situation $(p/q)^2 = 2$ is impossible. ♠

Theorem 2.3 *Suppose that there are* 6 *people in a room. Then either there are* 3 *people who are mutual friends or* 3 *people who are mutual strangers.*

Proof: Let A be one of the people. There are 5 other people in the room, so either A is friends with at least 3 other people or A is a stranger to at least 3 other people. Consider the two possibilities one at a time. Suppose first that A is friends with B and C and D. If B and C are also friends, than (A, B, C) make a triple of mutual friends. The same thing happens if A and C are friends and if B and C are friends. So, either we have 3 mutual friends or B, C, D make a triple of mutual strangers.

The same argument works when A is a stranger to at least 3 people. Just run the same argument interchanging the words *friends* and *strangers*. ♠

This last proof goes even further "off the gold standard". (The gold standard is that you go back to the axioms for every step.) In the previous proofs, we did the whole argument out to the end. In the last proof, we divided the proof in half, did the first half out completely, and then said that the second half was similar. Sometimes this can be a source of errors in proofs. The author will say "this case is similar to the previous one" but then actually there will be some subtle difference which kills the whole thing.

3 Maps Between Sets

Informally, a *map* between two sets A and B is a rule which assigns a member $f(a) \in B$ to each member $a \in A$. You could picture drawing a line from each member $a \in A$ to some member $b \in B$ and then the member connected to a is declared to be $f(a)$. These are really just generalizations of the functions you have been learning about for many years.

One problem with this definition is that you might want some restrictions on the allowable rules. For instance, suppose that A is the integers and B is the set $\{0, 1\}$. Your rule is $f(a) = 0$ if you personally find a boring, and $f(a) = 1$ if you find a interesting. This definition has plenty of problems. One main problem is that it isn't a reproducible rule. Other people might have different opinions. Another main problem is that the definition might change over time. Another problem is that the map might be entirely defined. Maybe you are on the fence about some numbers. And so on.

Here is a formal definition of a map between sets. I'll warn you in advance that you probably won't like it. First of all let $A \times B$ be the product of A and B. So, $A \times B$ consists of all ordered pairs (a, b) with $a \in A$ and $b \in B$. Then a map from A to B is a subset $C \subset A \times B$ such that, for each $a \in A$ there is a unique member $(a, b) \in S = C$. The function value of $f(a)$ is then this particular b. This wierd definition ought be familiar to you: really, we're defining a map between sets in analogy with the familiar *graph of a function*.

If you don't like this definition, you can think of a map between sets as a rule, as above, which is perfectly well defined, and unambiguous, and eternal.

The notation for the map we have been talking about is often written as

$$f : A \to B,$$

and usually I will write it like this. Another common way to say it is: Let f be a map from A to B. Sometimes I'll say both things, for emphasis.

4 Kinds of Maps

So, let $f : A \to B$ be a map from A to B.

Injective Maps: The map f is called *injective* if f takes on different values for different elements of A. That is, if a_1 and a_2 are two different elements of A, then $f(a_1)$ and $f(a_2)$ are two different elements of B. Sometimes one also says in this case that f is *one-to-one*. The words *injective* and *one-to-one* mean the same thing in mathematics. Here are some examples:

- The map $f(x) = 2x$ is an injective map from $\boldsymbol{Z}$ (the integers) to $\boldsymbol{Z}$.
- The map $f(x) = x^2$ is not an injective map from $\boldsymbol{Z}$ to $\boldsymbol{Z}$. The problem is that a and $-a$ are distinct but $f(a) = f(-a)$.

Surjective Maps: The map f is called *surjective* if, for every $b \in B$ there is some a such that $f(a) = b$. In other words, every member of B is "hit by f" so to speak. A surjective map is also called *onto*. Here are some examples.

- The map $f : \boldsymbol{Z} \to \{0, 1, 2, ...\}$ defined by $f(x) = x^2$ is surjective.
- The map $f : \boldsymbol{Z} \to \boldsymbol{Z}$ defined by $f(x) = x^2$ is not surjective. For instance, there is no x so that $f(x) = -1$.

Bijective Maps: A map $f : A \to B$ is *bijective* if it is both injective and surjective. Sometimes it is said that f sets up a *one-to-one correspondence* between A and B. Sometimes a bijective map is also called an *isomorphism*. The bijective maps have a very important property: If $f : A \to B$ is a bijective map and $g : B \to C$ is a bijective map, then we can find a new bijective map $h : A \to C$. It is defined by the rule

$$h(a) = g(f(a)).$$

This definition makes sense because $f(a) \in B$, so that $g(f(a))$ is defined. The map h is often called *the composition* of f and g.

Here is one principle I'll use a lot below: If there is a bijective map from A to C and a bijective map from B to C, then there is a bijective map from A to B. The proof is similar to the proof for the composition.

Exercise 2: Prove the principle I just mentioned. Your proof doesn't have to go back to the axioms.

5 Finite Sets

One way to think about the natural numbers is that they are the specific sets

- $\emptyset$
- $\{\emptyset\}$
- $\{\emptyset, \{\emptyset\}\}$

and so on. The Axiom of Infinite essentially says that these numbers "go on forever".

We say that a set A is *finite* if there is a bijecive map $f : A \to n$ for some natural number n. In particular, each natural number itself is finite. Another way to think about the natural numbers is that each one is a name for all the sets which are bijective with a particular one of the sets "listed" above. I put "listed" in quotes because technically I only listed the first few, but the list is meant to go on and the Axiom of Infinity says that it does. So, in other words 3 is a name all the sets which have a bijective map to the specific set

$$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}.$$

We can also say that two finite sets represent the same number exactly when there is a bijection from each one to the same n. This is true if and only if there is a bijection between the two sets.

For finite sets, the existence of bijective map $f : A \to B$ really is the statement that the number representing A is the same as the number representing B. Often $|A|$ is the notation for "the number representing A" and likewise for B. Here are two other basic theorems:

- $|A| \leq |B|$ if and only if there is an injective map $f : A \to B$.
- $|A| \leq |B|$ if and only if there is a surjective map $f : B \to A$.

I'm not going to prove these things, but the basic idea of the proof is to look at the smallest (supposed) case where the result fails and then observe that neither A nor B is the emptyset, so we can consider the predecessors of A and B obtained by deleting one element from each. So, in short, the same kind of argument from induction works for these results.

6 Infinite Sets

A set is defined to be *infinite* if it is not finite. So, in other words, the set S is infinite if there is no bijective map $f : S \to n$ for any integer n. The set $\boldsymbol{N}$ of positive integers is an example of an infinite set. Likewise, the set of all binary sequences (such as 011010101101010101...) is an infinite set.

Say that two infinite sets A and B are the same *size* if there is a bijective map $f : A \to B$. In this case, it is also said that A and B have the same *cardinality*. The words *size* and *cardinality* mean the same thing in this context. Here is a basic principle: Suppose

- A and B have the same size.
- B and C have the same size.

then A and C have the same size.

Given a set A, we define the *cardinal number* $|A|$ to be the name of all the sets which have the same size as A. With this notation $|A| = |B|$ means that there is a bijection $f : A \to B$. We could also define $|A| \leq |B|$ if there is an injective map $f : A \to B$. You might ask the question: Suppose $|A| \leq |B|$ and $|B| \leq |A|$. Is it true that $|A| = |B|$. This is true. It is called the *Schroeder-Bernstein theorem*. I'll give a proof in class (or in a later set of notes) if there is enough interest in it.

Continuing on with this notation, we write $|A| \neq |B|$ if there is no bijective map $f : A \to B$. Finally, we write $|A| < |B|$ if $A \leq |B|$ and $|A| \neq |B|$. All this seems to be in line with the finite numbers. You might ask the question: Are every two sets comparable? When A and B are finite sets, then exactly one of three statements is true:

- $|A| < |B|$.
- $|A| > |B|$ (that is $|B| < |A|$.)
- $|A| = |B|$.

Is this true when A and B are infinite sets? The answer is: It depends. If you are willing to use the axiom of choice, then the result is true. Otherwise, there is no way to prove it using the other ZF axioms.

7 Cantor's Diagonal Argument

Let $\boldsymbol{N}$ denote the natural numbers and let $\boldsymbol{B}$ denote the set of infinite binary sequences.

Theorem 7.1 $|\boldsymbol{N}| \leq |\boldsymbol{B}|$.

Proof: Define $f : \boldsymbol{N} \to \boldsymbol{B}$ by the rule

- $f(1) = 10000...$
- $f(2) = 01000...$
- $f(3) = 00100...$

and so on. This is an injective map from $\boldsymbol{N}$ to $\boldsymbol{B}$, and that is what we wanted. ♠

Now for the famous result:

Theorem 7.2 $|\boldsymbol{N}| \neq |\boldsymbol{B}|$.

Proof: Suppose that there was a bijective map $f : \boldsymbol{N} \to \boldsymbol{B}$. Then we can make a complete list of the binary sequences, like this

- $f(1) = a_{11}a_{12}a_{13}a_{14}...$
- $f(2) = a_{21}a_{22}a_{23}a_{24}...$
- $f(3) = a_{31}a_{32}a_{33}a_{34}...$

Here each a_{ij} stands for either 0 or 1. We don't know what these values are, but supposedly they are set up to give a complete list. Define $b_{ij} = 1 - a_{ij}$. So, if $a_{ij} = 1$ then $b_{ij} = 0$ and *vice versa*. Consider the sequence

$$b_{11}b_{22}, b_{33}, ...$$

Where is this sequence on the list. The first bit is wrong in the first position, and the second bit is wrong in the second position, and so on. So, this particular sequence is not on the list. This is a contradiction. The only way out is that $|\boldsymbol{N}| \neq |\boldsymbol{B}|$. ♠

Since $|N| \leq |B|$ and $|N| \neq |B|$ we have $|N| < |B|$. The size of the set of natural numbers is smaller than the size of the set of binary sequences.

8 Power Sets

Here are generalizations of the results just proved. Let 2^S denote the power set of a set S. So, 2^S is the set of all subsets of S.

Theorem 8.1 $|S| \leq |2^S|$.

Proof: Define $f : S \to 2^S$ by the formula $f(a) = \{a\}$. That is, $f(a)$ is the set whose only member is a. This map does the job for us: It is an injective map from S to 2^S. ♠

Theorem 8.2 $|S| \neq |2^S|$.

Proof: Suppose that there is a bijective map $f : S \to 2^S$. Define B to be the set $a \in S$ such that $a \notin f(a)$. That is a is not a member of the subset $f(a)$. Since f is a bijective map, there is some $b \in S$ so that $f(b) = B$. Is $b \in B$? There are two cases.

Suppose $b \in B$. Then $b \notin f(b)$. But $f(b) = B$. This means that $b \notin B$. This leads to a contradiction.

Suppose $b \notin B$. Then $b \in f(b)$. But $f(b) = B$. This means that $b \in B$. This leads to a contradiction.

Either case leads to a contradiction. The only way out is that f does not exist. ♠

Since $|S| \leq |2^S|$ and $|S| \neq |2^S|$, we conclude that $|S| < |2^S|$.

The above result suggests that there is an entire hierarchy of infinities: You can start with $S = \boldsymbol{N}|$, the set of natural numbers. Then you can form

- $S_0 = \boldsymbol{N}$.
- $S_1 = 2^{S_0}$.
- $S_2 = 2^{S_1}$.

and so on. The cardinal number S_k is often written as $\aleph_k$. So, all the sizes α_0, α_1, α_2,... are increasingly large sizes of infinity.

9 It Never Ends

Consider the sets $S_0, S_1, S_2, ...$ just constructed. It turns out that there is no bijection from the set S_n to the union

$$T = S_0 \cup S_1 \cup S_2 \cup ...$$

We have $|S_n| < |T|$ for all n. So, the set $|T|$ is a size of infinite larger than the entire infinite hierarchy constructed above. But why stop there, we can go on and define sets:

- $T_0 = T$.
- $T_1 = 2^{T_0}$.
- $T_2 = 2^{T_1}$.

And so on. And so on. And so on. Etc. Etc.