

The Mathematician's Infinities

An introduction to infinity for the thinking layman

In this paper I am trying to strike a balance - avoiding any unnecessary, distracting jargon without sacrificing the logical thoroughness of academic mathematics. In some cases I have replaced mathematical terms with other words whose normal meanings are closer to the intended meaning.

One simplification is made throughout: This presentation completely ignores the existence of the negative numbers. All the results presented are still valid when negative numbers are also considered, but the written explanations of those results would need to be tweaked to account for them.

-Kevin Hughes, July 2009

Part One: Sets and Sizes

For mathematicians, almost everything boils down to sets, or collections of objects. Usually we use curly brackets to symbolize a set: $\{4, 5, 6\}$ is a set containing three numbers. $\{a, b\}$ is a set containing two letters.

Order doesn't matter in sets, so you can rearrange the way the items in the set are listed. $\{6, 5, 4\}$ is considered equal to $\{4,5,6\}$. In other words, this is the same set written two different ways.

What's this: $\{\}$? That's the "null set", or the set containing nothing. The null set shows up a lot, as an exception to the way we phrase different rules, so it is often mentioned separately, as you will see.

Definition: A *subset* of a set is another set made up entirely of members of the other set. For instance, $\{1, 2\}$ is a subset of the set $\{1, 2, 3, 4, 5, 6\}$. $\{4, 6, 8\}$ is a subset of the even numbers. It is also a subset of the counting numbers. Every set is a subset of itself and the null set is (by convention) considered a subset of every other set. We don't make much use of this definition until Part Two, but it is included here because it is very basic to the idea of sets.

Definition: The things inside the set are called the *members*.

Sets can even contain other sets, which we will see in Part Four. We can take the set $\{1, 2, 3\}$ and the set $\{4, 5, 6\}$ and put them into a set as members. This set has only two members: $\{\{1, 2, 3\}, \{4, 5, 6\}\}$ Each member of the set happens to be another set, but that doesn't affect the "size" of the set. Similarly, $\{100, 200\}$ will be considered the same size as $\{1, 2\}$. When mathematicians talk about the size of a set, they are just concerned with how many things are in the set, not with what those things might be.

Ok, so let's talk about size in a more exact manner. The sizes of the sets we've seen so far have been given as numbers, specifically counting numbers; Sets have a size of 1 or 2 or 3, and so on. (The null set, with a size of 0, is an exception to this, because the counting numbers begin with 1.) We say that these numbers are the "size" of the set. A fancier way to set up the idea of size is to say this:

Definition: The *size* of a set is defined to be a counting number "n" if and only if there can exist a "perfect one-to-one correspondence" between the members of the set and the counting numbers 1 through n.

What is a "perfect one-to-one correspondence" between two sets? A perfect one-to-one correspondence (for this paper) satisfies these three properties:

- 1) Each member of one set is (arbitrarily) assigned a correspondence, i.e. paired, with a member of the other set.
- 2) No member of either set is left out.

- 3) No two members of one set correspond to the same member of the other set - in other words we are dealing with unique pairings.

For instance, the set of the letters in the alphabet has a size of 26 because you are able to set up the correspondence (a & 1), (b & 2), (c & 3), ... (z & 26). It is important to note that this isn't the only correspondence you can create between the alphabet and the numbers 1 through 26. Since the order of the set doesn't matter, you could rearrange the letters in any order (z & 1), (y & 2), etc, and every time you will find that there are 26 members. What you cannot do is have both a and b paired with the number 1, or have c be paired with both 2 and 3.

So far, this idea of a perfect one-to-one correspondence may seem like an over-complicated way of talking about the size of the set, but we will see that it is a powerful and important concept later in this presentation.

Finally, not all sets have any such correspondence to the counting numbers 1 through n, no matter which counting number the "n" happens to be. The most trivial example is the null set: Zero is not a counting number. More significantly, some sets are too big. You can start numbering the members of the set with 1, 2, 3, etc, but you never finish, so there isn't a "last" counting number n that tells you what size the set is.

Definition: If a set has a size n, as previously defined, then we say that set is *finite*. All sets (other than the null set) who are not finite by that definition are defined to be *infinite*. (Conventionally, the null set is also considered to be finite.)

Now let's talk about infinity...

Part Two: "Listable" Infinity

Now we can start to talk about infinite sets.

Consider the set of all even numbers, $\{2, 4, 6, 8, 10, \dots\}$. This set is infinite: if we try to create a one-to-one correspondence between this set and the counting numbers, there won't be a counting number n that is the "last" one, which would tell us the size.

Definition: Two sets are said to be the *same size* when there exists a perfect one-to-one correspondence between their members. This applies to both finite and infinite sets!

$\{1, 2, 3\}$ is the same size as $\{2, 4, 6\}$. One rule of correspondence between them might be a factor of two: Multiply each member of the first set to find its unique corresponding member of the second, and divide by two to go the other direction.

Now we're ready to consider the first of many mind-bending concepts about infinite sets. The set of even numbers mentioned above is a limited subset of the counting numbers. In fact, there are an infinite number of members of the counting numbers that are "missing" from the even numbers. In other words, the set of those missing members (the odd numbers) is itself an infinite set, $\{1, 3, 5, 7, 9, \dots\}$. And despite all this, by the definition given, the set of even numbers is in fact the same size as the counting numbers! The rule of multiplying and dividing by two sets up a perfect one-to-one correspondence from one set to the other.

Similarly, the set of counting numbers bigger than 10 is the same size as all the counting numbers: The rule of correspondence for that one is just to subtract 10.

The key thing to understand here is that even though a perfect one-to-one correspondence might exist between one set and some limited subset of the counting numbers, another correspondence might show that they are the exact same size. The fact that it is at all possible to create a rule of correspondence satisfying the same-size principle means that they are the same size, period -- even though other rules of correspondence might seem to show that one is smaller or bigger than the other.

At this point, it might seem that our definition of sets being the "same size" is just flawed or inadequate for infinite sets. Perhaps all infinite sets have such a correspondence. But this is not true. As we will see (and in fact prove) later on, some infinite sets are so fundamentally bigger than the counting numbers that no such correspondence could ever exist.

Let's look at two very big sets:

Definition: The *fractions* are the set of all numbers that can be expressed as a fraction of two whole numbers. So $\frac{1}{2}$ is in this set, but so are all the whole numbers, since 5 can be

expressed as $5/1$. Zero is in this set. The square root of two is not in this set because there don't exist any two integers that you can divide to give you $\sqrt{2}$.

Definition: The **real numbers** are the set of all numbers that can be expressed as an ordinary decimal expansion. That is, any number that can be written out in digits in the normal way we all learn in school when we are young. This includes the whole numbers, the fractions, and some others. Examples:

$1 = 1.000000\dots$ (a counting number and a fraction)
 $1/3 = 0.3333333\dots$ (a fraction but not a counting number)
 $\sqrt{2} = 1.414213\dots$ (neither a counting number nor a fraction)
 $\pi = 3.14159\dots$ (again neither)

Intuitively, these two sets might seem to be bigger than the counting numbers, but about equal in size to each other. They share some interesting properties. For instance, both of them “fill” the number line: For any two fractions, no matter how close they are, there are an infinite number of other fractions between them. The same is true for the real numbers.

Interestingly, it will be demonstrated that the set of fractions is the “same size” as the counting numbers, but the real numbers are not!

Definition: A set is called **listable** if you can make a perfect one-to-one correspondence between that set and either the counting numbers or a subset of the counting numbers.

Finite sets are always listable. Some infinite sets are listable too. The set of even numbers has a perfect correspondence to the counting numbers, so it is both listable and infinite. A good way to understand this is that a listable infinite set can be expressed as a list: It has a first, second, & third member, and so on. Every member of the set is somewhere on the list, even if it is the 235,634,587,324,324th item on the list, it is there somewhere.

Suppose we have a set that is very large and we want to discover whether it is listable. What we would need to do is find a way to list all the members of the set in any particular order without leaving anything out. We will see that the rational numbers are listable and equal in size to the counting numbers.

Each fraction has two pieces of information (the two whole numbers: top and bottom) that we might try to list them by. If we were to start listing them like this, it wouldn't work:

$1/1, 1/2, 1/3, 1/4, 1/5, \dots$

That list uses up all the available spots on the list (first, second, third, etc) for numbers with a top number of 1, without ever getting to numbers like $3/4$. There's a lot of ways we might start to list them, only to discover that it is going to leave something out. The reader is encouraged to attempt this before continuing, to get an idea of how hard it is to come up with a list that doesn't leave anything out.

Here's the solution: What we need to do is sort them by the sum of the top and bottom numbers: If we are dealing with $1/2$, the sum of 1 and 2 is 3. The sum for $3/4$ is 7. For any given "sum", there are a limited number of fractions that result in that sum. We start with 1 and go up from there, so the list looks like the following (with line breaks grouping the sums):

First: $0/1$ (the only fraction that "sums" to 1)

Second: $0/2$ (sums to 2)

Third: $1/1$ (sums to 2)

Fourth: $0/3$ (sums to 3)

Fifth: $1/2$ (sums to 3)

Sixth: $2/1$ (sums to 3)

Seventh: $0/4$ (sums to 4)

...and so on.

You might be able to see from the beginning of the list that there are going to be 4 fractions whose numbers "sum" to 4, 5 who "sum to 5, etc. We can even compute the place of any number on the list. For instance, $5/9$ has a "sum" of 14. I happen to know that the sum of the counting numbers 1 through N always comes out to:

$$N * (N + 1) / 2$$

That means that the items on the list that "sum" to 13 take up a number of places on the list equal to: $13 * 14 / 2 = 91$ places, and $5/9$ is the sixth item that sums to 14 (starting with $0/14$). Then we just say $91 + 6$ is 97, so $5/9$ is the 97th item on the list. Since every fraction's place on the list can be found in this manner, we can conclude that the set of fractions is the same size as the set of counting numbers. Next, we will see that the set of real numbers is not.

But first, one final comment: Did you notice my mistake? This list includes every single fraction, and in fact includes every fraction multiple times: zero appears a lot, as $0/1$, $0/2$, $0/3$, etc. $1/2$ will appear again as $2/4$ and then as $3/6$, etc. These are different ways of writing the same number. Technically, this violates the rule of the perfect one-to-one correspondence! A more rigorous rule of correspondence would have a second step that "collapsed" the list to leave out the duplicates, but that part is trivial in terms of understanding the theory -- we know it can be done.

Part Three: Unlistable Infinity

We now want to prove that the real numbers can't be "listed" - that they can't ever correspond to the counting numbers in a one-to-one fashion. This is a tall order. It's not enough to say that I can't find a way to list them, or that no one I know has found a way. We have to prove that it is impossible! Fortunately, mathematicians often want to prove a negative, and they have devised some handy ways to do so.

The logic will go like this: We will create a rule that generates a real number based on the contents of an infinite list of real numbers. You can think of this as a function where:

input = a list of numbers
output = a new number

We will show that this rule has a special property: The new number is never ever equal to any of the numbers in the list, even though the list is infinite. What does that mean? It means that if you ever thought that you had created a list that included all the real numbers, you could plug it into this function and generate a number not on the list - which would prove you wrong! The new number proves that your list is not complete. And since the rule works for any list at all, the rule proves that every possible list is incomplete.

Suppose you have an input that looks like this: (We've done some underlining in anticipation of the next step.)

0.12349823...
3.37458943...
157.82940572...
1.39803432...
10.23948324...
...and so on

The information the function uses is the digits in the diagonal following the decimal point. They are underlined: 1, 7, 9, 0, 8, ...

That series is then incremented by one, so it becomes 2, 8, 0, 1, 9, ... (We're dealing with single digits so when the nine is incremented, it rolls over and becomes a 0.)

To define a real number we just need to know each digit of its written-out form. The output of the function uses the incremented series after its decimal point. It doesn't matter what the whole number portion is, so we just use zero.

Output = 0.28019...

It should be clear that it doesn't matter what way the numbers are listed: this function will always give you something that is missing from the list. If, to the contrary, the number that comes out of the function were in, say, the 100th place on the list, then that would

contradict the definition of the function: we know that the 100th digit of this number differs from the 100th digit of the 100th number on the list by a factor of one, because the 100th digit of the new number came from the 100th digit of the 100th number in the list, but it was incremented.

In other words, looking at our output number given in the example, we can make the following observations:

- 1) The "2" makes sure it differs from the first number in the list.
- 2) The "8" makes sure it differs from the second number in the list.
- 3) The "0" makes sure it differs from the third number on the list.
...and so on.

Thus we conclude that the set of real numbers is unlistable. According to our definition of "same size" sets, that means that the real numbers are a fundamentally bigger set than the counting numbers, although both sets are infinite. Infinity isn't any more of a "size" than finitude; it is a class of sizes.

We have now considered at least two different sizes of infinity. But are there more? Could anything be bigger than the set of all real numbers?

Part Four: The Infinitude of the Infinities

Definition: A *power set* of a given set is the set of all subsets of the original set. Some examples will clarify what that means:

Consider the set $\{3\}$ containing only one number. There are two possible subsets of this set: one is the null set $\{\}$, which is a subset of every other set. The other is $\{3\}$, because each set can be considered a subset of itself. The power set of $\{3\}$ is therefore:

$\{\{\}, \{3\}\}$

A set containing only one item is a trivial example. Let's look at the power set of $\{a, b, c\}$. It contains all the following eight sets:

$\{\}$
 $\{a\}$
 $\{b\}$
 $\{c\}$
 $\{a, b\}$
 $\{b, c\}$
 $\{a, c\}$
 $\{a, b, c\}$

Remember that order does not matter, so the subset $\{c, a\}$ is already in that list: it is the same as $\{a, c\}$ which is second-to-last in the list.

It should be clear that the power set of any finite set is bigger than the set itself. In fact, if a set has a size of n , the size of the power set is 2^n . A set with 1 item had a power set with $2^1 = 2$ items, and a set with 3 items had a power set with $2^3 = 8$ items. This even works for the null set. The size of the null set $\{\}$ is 0. But the null set is a subset of itself, so the power set of the null set has one member, the null set itself: $\{\{\}\}$. Sure enough, $2^0 = 1$.

It turns out that the same property (not the exact size 2^n , but the fact that the power set is bigger) is true of infinite sets! We will now prove this, using a similar technique as in the argument about the real numbers.

It gets a little complicated, so we're going to start naming things with letters to keep it all clear. We will call our original set S (for "Set") and the power set of S will be called P . We want to prove that they are different sizes. This means we have to conclude that it is impossible for there to be a perfect one-to-one correspondence between P and S . We will do this by again defining a function, where this time:

input = any rule that matches each member of S with something from P (not necessarily a perfect one-to-one correspondence - some members of P could be unused.)

output = a subset of S , which we will call N for "new subset"

This function will have the property that no matter what rule of correspondence between P and S goes in, that same rule of correspondence will not match the output subset N with any member of S. We then know two things:

- 1) N is one the members of P because P is defined as the set of all subsets of S.
- 2) The rule of correspondence between P and S does not pair N with any member of S.

These two things mean that our rule of correspondence between P and S is not a perfect one-to-one correspondence. Since this function can be used to generate that conclusion about any rule of correspondence, we can conclude that it is impossible to create a perfect one-to-one correspondence between P and S. If you thought you had succeeded, you could use this function to prove yourself wrong every time.

In order to define the subset N, we need to know, for each member of S, whether that member is also a member of N or not. The function is simple: We have as our input a rule of correspondence that assigns members of S to subsets of S. Suppose one such member is called m and it is paired with a subset B. Then either m is in B or it isn't. If m is a member of B then our function tells us that m is not a member of N. If m is not in B, then we do put m in N. We repeat this for each member of S and that tells us exactly which members of S are in N.

Now that we have defined the function that creates N, lets look at the result. Suppose we assumed, for some set S, that the power set P was the same size as S. That would mean we could create a perfect one-to-one correspondence between P and S. That correspondence could be fed into our function to generate a subset N. N is already in P, so our rule of perfect one-to-one correspondence will pair up N with some member of S, which we can call m. If that is the case, we then ask whether m is in N or not?

Consider that N existed in P as a subset even before we used the function. N, essentially, is one of those subsets we talked about as "B". If m is in N, then our function would have generated a new subset that didn't include m. If m is not in N, then the output would have included m. This is a blatant contradiction: If m is in N, then it isn't. If it isn't, then it is. The underlying assumption that led to this contradiction must have been false: the assumption that P is the same size as S. Since this function can always be used for any set, finite or infinite, listable or unlistable, our generalized conclusion is that a power set can never be the same size as the original set.

It should be clear that it can't be smaller, either. We won't bother to give a rigorous proof of this here, (we technically would need to define what "bigger" or "smaller" even mean for infinite sets) but consider that the power set contains, for each member of the original set, the subset containing only that member.

And now, having concluded that the power set of a set is always bigger than the original set, we can answer the question posed: Are the counting numbers and the real numbers the only two sizes of infinity that exist? If not, how many different sizes of infinity exist?

In fact, there are infinitely many different sizes of infinity. No infinite set is the “biggest” infinite set, because you can always find a bigger one - specifically, you can always look to its power set. And then you can look to the power set of the power set, and so on. (This is similar to saying that the counting numbers are infinite because if you think you have the biggest one, you can always add 1 to it and get a bigger counting number.)

And if that doesn't blow your mind, ask yourself this: We know that the number of different sizes of infinity is infinite. But which infinity is it? Is there only a listably infinite number of infinities, or are there many, many more ? ...

Part Five: The Real Mathematical Jargon

In case anyone is interested, here is a quick run-down on the real way mathematicians talk about these ideas.

First, "listable" isn't a real term. This concept is called "countable". So the counting numbers are called "countably infinite" and the real numbers are "uncountably infinite". I thought that "countable" sounded too much like "finite" for the general reader.

Mathematicians don't usually say "fractions" in this sort of context. Ratios of whole numbers are called "rational" numbers. All the other real numbers are called "irrational".

There is no such term as a "perfect one-to-one correspondence". This is actually a "bijection", and a bijection is technically a function that is only defined in one direction. (It's usually easy to find its inverse, though.) The phrase "one-to-one" is used to describe functions that don't map two inputs to the same output, but not all possible outputs have to be reserved for an input. If all possible outputs are used by some input, the function is called "onto". A function that is both "one-to-one" and "onto" is always a bijection.

The "size" of a set is usually referred to as its "cardinality". Infinite sets get cardinalities too, but they are not counting numbers (obviously). Instead, the Hebrew letter aleph (\aleph) is used with a subscript. For instance, the counting numbers are said to have a size of \aleph_0 .

The following words are genuine terms meaning the same thing as in this presentation: set, member, subset, power set, finite, infinite, the counting numbers, the real numbers.

The counting numbers are also called the natural numbers or the cardinal numbers. All three meanings exclude negatives (and zero). The real numbers and the rational numbers include negatives. If we include the negative whole numbers with the counting numbers, that is called either the whole numbers or the integers.