THE BASIC TRICHOTOMY: FINITE, COUNTABLE, UNCOUNTABLE

PETE L. CLARK

1. Introducing equivalence of sets, countable and uncountable sets

We assume known the set $\mathbb{Z}^+$ of positive integers, and the set $\mathbb{N} = \mathbb{Z}^+ \cup \{0\}$ of natural numbers. For any $n \in \mathbb{Z}^+$, we denote by $[n]$ the set $\{1, \ldots, n\}$. We take it as obvious that $[n]$ has n elements, and also that the empty set $\emptyset$ has 0 elements. Just out of mathematical fastidiousness,[1] let's define $[0] = \emptyset$ (why not?).

It is pretty clear what it means for an arbitrary set S to have 0 elements: it must be the empty set. That is – and this is a somewhat curious property of the empty set – $\emptyset$ as a set is uniquely characterized by the fact that it has 0 elements.

What does it mean for an arbitrary set S to have n elements? By definition, it means that there exists a bijection $\iota : S \to [n]$, i.e., a function which is both injective and surjective; or, equivalently, a function for which there exists an inverse function $\iota' : [n] \to S$.[2]

Let us call a set *finite* if it has n elements for some $n \in \mathbb{N}$, and a set *infinite* if it is not finite.

Certainly there are some basic facts that we feel should be satisfied by these definitions. For instance:

Fact 1. *The set $\mathbb{Z}^+$ is infinite.*

Proof: It is certainly nonempty, so we would like to show that for no $n \in \mathbb{Z}^+$ is there a bijection $\iota : [n] \to \mathbb{Z}^+$. This seems obvious. Unfortunately, sometimes in mathematics we must struggle to show that the obvious is true (and sometimes what seems obvious is not true!). Here we face the additional problem of not having formally axiomatized things, so it's not completely clear what's "fair game" to use in a proof. But consider the following: does $\mathbb{Z}^+$ have one element? Absolutely not: for any function $\iota : [1] = \{1\} \to \mathbb{Z}^+$, ι is not surjective because it does not hit $\iota(1) + 1$. Does $\mathbb{Z}^+$ have two elements? Still, no: if ι is not injective, the same argument as before works; if ι is injective, its image is a 2 element subset of $\mathbb{Z}^+$. Since $\mathbb{Z}^+$ is totally ordered (indeed well-ordered), one of the two elements in the image is larger than the other, and then that element plus one is not in the image of our map. We could prove it for 3 as well, which makes us think we should probably

[1] Well, not really: this will turn out to be quite sensible.

[2] I am assuming a good working knowledge of functions, injections, surjections, bijections and inverse functions. This asserts at the same time (i) a certain amount of mathematical sophistication, and (ii) a certain amount of metamathematical informality.

work by induction on n. How to set it up properly? Let us try to show that for all n and all $\iota : [n] \to \mathbb{Z}^+$, there exists $N = N(\iota)$ such that $\iota([n]) \subset [N]$. If we can do this, then since $[N]$ is clearly a proper subset of $\mathbb{Z}^+$ (it does not contain $N+1$, and so on) we will have shown that for no n is there a surjection $[n] \to \mathbb{Z}^+$ (which is in fact stronger than what we claimed). But carrying through the proof by induction is now not obvious but (much better!) very easy, so is left to the reader.

Remark: What did we use about $\mathbb{Z}^+$ in the proof? Some of the Peano axioms for $\mathbb{Z}^+$, most importantly that it satisfies the principle of mathematical induction (POMI). Since it is hard to imagine a rigorous proof of a nontrivial statement about $\mathbb{Z}^+$ that does not use POMI, this is a good sign: things are proceeding well so far.

What about $\mathbb{Z}$: is it too infinite? It should be, since it contains an infinite subset. This is logically equivalent to the following fact:

Fact 2. *A subset of a finite set is finite.*

Proof: More concretely, it suffices to show that for any $n \in \mathbb{N}$ and and subset $S \subset [n]$, then for some $m \in \mathbb{N}$ there exists a bijection $\iota : S \to [m]$. As above, for any specific value of n, it straightforward to show this, so again we should induct on n. Let's do it this time: assume the statement for n, and let $S \subset [n+1]$. Put $S' = S \cap [n]$, so by induction there exists a bijection $\iota' : [m] \to S'$ for some $m' \in \mathbb{N}$. Composing with the inclusion $S' \subset S$ we get an injection $\iota : [m] \to S$. If $n+1$ is not an element of S, then $S' = S$ and ι is a bijection. If $n+1 \in S$, then extending ι' to a map from $[m+1]$ to S by sending $m+1$ to $n+1$ gives a bijection. Done.

Again, by contraposition this shows that many of our most familiar sets of numbers – e.g. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ – are infinite.

There is one more thing we should certainly check: namely, we have said that a set S has n elements if it can be put in bijection with $[n]$ for some n. But we have not shown that this n is unique: perhaps a set can have n elements and also $n + 691$ elements? Of course not:

Fact 3. *For distinct natural numbers n, n', there is no bijection from $[n]$ to $[n']$.*

Of course, we even know a more precise result:

Fact 4. *Let S be a set with m elements and T a set with n elements.*
a) If there exists a surjection $\varphi : S \to T$, then $m \geq n$.
b) If there exists an injection $\varphi : S \to T$, then $m \leq n$.

Exercise 1: Give a proof of Fact 4 which is rigorous enough for your taste.

Remark: For instance, part b) is the famous "Pigeonhole" or "Dirichlet's box" principle, and is usually regarded as obvious. Of course, if we play the game of formalized mathematics, then "obvious" means "following from our axioms in a way which is so immediate so as not to deserve mention," and Fact 4 is not obvious in this sense. (But one can give a proof in line with the above induction proofs, only a bit longer.)

Exercise 2: Show that for sets S and T, the following are equivalent:

a) There exists a surjection $S \to T$.
b) There exists an injection $T \to S$.

Let us press on to study the properties of *infinite* sets.

Basic Definition (Cantor): We say that S and T as *equivalent*, and write $S \cong T$ if there exists a bijection $\iota : S \to T$.

Historical Remark: When there exists a bijection between S and T, Cantor first said that S and T have the same *power*.[3] As is often the case in mathematics, this forces us to play a linguistic-grammatical game – given that a definition has been made to have a certain part of speech, write down the cognate words in other parts of speech.[4] Thus a faithful rendition of Cantor's definition in adjectival form would be something like *equipotent*. The reader should be warned that it would be more common to use the term *equinumerous* at this point.

However, we have our reasons for choosing to use "equivalent." The term "equinumerous," for instance, suggests that the two sets have the same number of elements, or in other words that there is some numerical invariant we are attaching to a single set with the property that two sets can be put in bijection exactly when both have the same value of this numerical invariant. But we would like to view things in exactly the opposite way. Let us dilate a bit on this point.

It was Cantor's idea that we should regard two sets as "having the same size" iff they are equivalent, i.e., iff their elements can be paired off via a one-to-one correspondence. Certainly this is consistent with our experience from finite sets. There is, however, a brilliant and subtle twist: colloquially one thinks of counting or measuring something as a process which takes as input one collection of objects and outputs a "number." We therefore have to have names for all of the "numbers" which measure the sizes of things: if you like, we need to count arbitrarily high. Not every civilization has worked out such a general counting scheme: I have heard tell that in a certain "primitive tribe" they only have words for numbers up to 4 and anything above this is just referred to as "many." Indeed we do not have proper names for arbitrarily large numbers in the English language (except by recourse to iteration, e.g., million million for a trillion).

But notice that we do not have to have such an elaborate "number knowledge" to say whether two things have the same size or not. For instance, one may presume that shepherding predates verbal sophistication, so the proto-linguistic shepherd needs some other means of making sure that when he takes his sheep out to graze in the countryside he returns with as many as he started with. The shepherd can do this as follows: on his first day on the job, as the sheep come in, he has ready some sort of sack and places stones in the sack, one for each sheep. Then in the future he counts his sheep, not in some absolute sense, but in relation to these stones. If one day he runs out of sheep before stones, he knows that he is missing some sheep (at least if he has only finitely many sheep!).

Even today there are some situations where we test for equivalence rather than

[3]Or rather, he said something in German that gets translated to this. Such pedantic remarks will be omitted from now on!

[4]This is a game that some play better than others, viz.: generization, sobrification, unicity.

count in an absolute sense. For instance, if you come into an auditorium and everyone is sitting in a (unique!) seat then you know that there are at least as many seats as people in the room without counting both quantities.

What is interesting about infinite sets is that these sorts of arguments break down: the business of taking away from an infinite set becomes much more complicated than in the finite case, in which, given a set S of n elements and any element $x \in S$, then $S \setminus x$ has $n-1$ elements. (This is something that you can establish by constructing a bijection and is a good intermediate step towards Fact 4.) On the other hand, $\mathbb{Z}^+$ and $\mathbb{N}$ are equivalent, since the map $n \mapsto n-1$ gives a bijection between them. Similarly $\mathbb{Z}^+$ is equivalent to the set of even integers ($n \mapsto 2n$). Indeed, we soon see that much more is true:

Fact 5. *For any infinite subset $S \subset \mathbb{Z}^+$, S and $\mathbb{Z}^+$ are equivalent.*

Proof: Using the fact that $\mathbb{Z}^+$ is well-ordered, we can define a function from S to $\mathbb{Z}^+$ by mapping the least element s_1 of S to 1, the least element s_2 of $S \setminus \{s_1\}$ to 2, and so on. If this process terminates after n steps then S has n elements, so is finite, a contradiction. Thus it goes on forever and clearly gives a bijection.

It is now natural to wonder which other familiar infinite sets are equivalent to $\mathbb{Z}^+$ (or $\mathbb{N}$). For this, let's call a set equivalent to $\mathbb{Z}^+$ *countable*.[5] A slight variation of the above argument gives

Fact 6. *Every infinite set has a countable subset.*

(Indeed, for infinite S just keep picking elements to define a bijection from $\mathbb{Z}^+$ to some subset of S; we can't run out of elements since S is infinite!) As a first example:

Fact 7. *The two sets $\mathbb{Z}$ and $\mathbb{Z}^+$ are equivalent.*

We define an explicit bijection $\mathbb{Z} \to \mathbb{Z}^+$ as follows: we map $0 \mapsto 1$, then $1 \mapsto 2$, $-1 \mapsto 3$, $2 \mapsto 4$, $-2 \mapsto 5$ and so on. (If you are the kind of person who thinks that having a formula makes something more rigorous, then we define for positive n, $n \mapsto 2n$ and for negative n, $n \mapsto 2|n| + 1$.)

The method proves something more general, a "splicing" result.

Fact 8. *Suppose that S_1 and S_2 are two countable sets. Then $S_1 \bigcup S_2$ is countable.*

Indeed, we can make a more general splicing construction:

Fact 9. *Let $\{S_i\}_{i \in I}$ be an indexed family of pairwise disjoint nonempty sets; assume that I and each S_i is at most countable (i.e., countable or finite). Then $S := \bigcup_{i \in I} S_i$ is at most countable. Moreover, S is finite iff I and all the S_i are finite.*

We sketch the construction: since each S_i is at most countable, we can order the elements as s_{ij} where either $1 \leq j \leq \infty$ or $1 \leq j \leq N_j$. If everything in sight is finite, it is obvious that S will be finite (a finite union of finite sets is finite). Otherwise, we define a bijection from $\mathbb{Z}^+$ to S as follows: $1 \mapsto s_{11}$, $2 \mapsto s_{12}$, $3 \mapsto s_{22}$, $4 \mapsto s_{13}$, $5 \mapsto s_{23}$, $6 \mapsto s_{33}$, and so on. Here we need the convention that when

[5]Perhaps more standard is to say "countably infinite and reserve "countable" to mean countably infinite or finite. Here we suggest simplifying the terminology.

s_{ij} does not exist, we omit that term and go on to the next element in the codomain.

Fact 9 is used very often in mathematics. As one immediate application:

Fact 10. *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof: Each nonzero rational number α can be written uniquely as $\pm\frac{a}{b}$, where $a,\ b \in \mathbb{Z}^+$. We define the height $h(\alpha)$ of α to be $\max a, b$ and also $h(0) = 0$. It is clear that for any height $n > 0$, there are at most $2n^2$ rational numbers of height n,[6] and also that for every $n \in \mathbb{Z}^+$ there is at least one rational number of height n, namely the integer $n = \frac{n}{1}$. Therefore taking $I = \mathbb{N}$ and putting some arbitrary ordering on the finite set of rational numbers of height n, Fact 9 gives us a bijection $\mathbb{Z}^+ \to \mathbb{Q}$.

In a similar way, one can prove that the set $\overline{\mathbb{Q}}$ of algebraic numbers is countable.

Fact 11. *If A and B are countable, then the Cartesian product $A \times B$ is countable.*

Exercise 3: Prove Fact 11. (Strategy 1: Reduce to the case of $\mathbb{Z}^+ \times \mathbb{Z}^+$ and use the diagonal path from the proof of Fact 9. Strategy 2: Observe that $A \times B \cong \bigcup_{a \in A} B$ and apply Fact 9 directly.)

The buck stops with $\mathbb{R}$. Let's first prove the following theorem of Cantor, which is arguably the single most important result in set theory. Recall that for a set S, its power set 2^S is the set of all subsets of S.

Theorem 12. *(First Fundamental Theorem of Set Theory)*
There is no surjection from a set S to its power set 2^S.

Remark: When S is finite, this is just saying that for all $n \in \mathbb{N}$, $2^n > n$, which is, albeit true, not terribly exciting. On the other hand, taking $S = \mathbb{Z}^+$ Cantor's Theorem provides us with an uncountable set $2^{\mathbb{Z}^+}$. In fact it tells us much more than this, as we shall see shortly.

Proof of Cantor's Theorem: It is short and sweet. Suppose that $f : S \to 2^S$ is any function. We will produce an element of 2^S which is not in the image of f. Namely, let T be the set of all $x \in S$ such that x is not an element of $f(x)$, so T is some element of 2^S. Could it be $f(s)$ for some $s \in S$? Well, suppose $T = f(s)$ for some $s \in S$. We ask the innocent question, "Is $s \in T$?" Suppose first that it is: $s \in T$; by definition of T this means that s is not an element of $f(s)$. But $f(s) = T$, so in other words s is not an element of T, a contradiction. Okay, what if s is not in T? Then $s \in f(s)$, but again, since $f(s) = T$, we conclude that s *is* in T. In other words, we have managed to define, in terms of f, a subset T of S for which the notion that T is in the image of f is logically contradictory. So f is not surjective!

What does this have to do with $\mathbb{R}$? Let us try to show that the interval $(0, 1]$ is uncountable. By Fact 5 this implies that $\mathbb{R}$ is uncountable. Now using binary expansions, we can identify $(0, 1]$ with the power set of $\mathbb{Z}^+$. Well, almost: there is the standard slightly annoying ambiguity in the binary expansion, that

$$.a_1a_2a_3\cdots a_n01111111111\ldots = .a_1a_2a_2\cdots a_n1000000000\ldots.$$

[6] I will resist the temptation to discuss how to replace the 2 with an asymptotically correct constant.

There are various ways around this: for instance, suppose we agree to represent every element of $(0,1]$ by an element which does not terminate in an infinite string of zeros. Thus we have identified $(0,1]$ with a certain subset T of the power set of $\mathbb{Z}^+$, the set of *infinite* subsets of $\mathbb{Z}^+$. But the set of finite subsets of $\mathbb{Z}^+$ is countable (Fact 9 again), and since the union of two countable sets would be countable (and again!), it must be that T is uncountable. Hence so is $(0,1]$, and so is $\mathbb{R}$.

There are many other proofs of the uncountability of $\mathbb{R}$. For instance, we could contemplate a function $f : \mathbb{Z}^+ \to \mathbb{R}$ and, imitating the proof of Cantor's theorem, show that it cannot be surjective by finding an explicit element of $\mathbb{R}$ not in its image. We can write out each real number $f(n)$ in its decimal expansion, and then construct a real number $\alpha \in [0,1]$ whose nth decimal digit α_n is different from the nth decimal digit of $f(n)$. Again the ambiguity in decimal representations needs somehow to be addressed: here we can just stay away from 9's and 0's. Details are left to the reader.

A more appealing, albeit more advanced, proof comes from a special case of the Baire category theorem: in any complete metric space, the intersection of a countable number of dense open subsets remains dense (although not necessarily open, of course). Dualizing (i.e., taking complements), we get that in any complete metric space, the union of a countable number of closed subsets with empty interior also has empty interior. Thus:

Corollary 13. *A complete metric space without isolated points is uncountable.*

Proof: Apply the dual form of Baire's theorem to the one-point subsets of the space.

Thus, since $\mathbb{R}$ is by definition the completion of $\mathbb{Q}$ with respect to the standard Euclidean metric, and has no isolated points, $\mathbb{R}$ must be uncountable. For that matter, even $\mathbb{Q}$ has no isolated points (which is strictly stronger: no element of the completion of a metric space minus the space itself can be isolated, since this would contradict the density of a space in its completion), so since we know it is countable, we deduce that it is incomplete without having to talk about $\sqrt{2}$ or any of that sort of thing. Indeed, the same argument holds for $\mathbb{Q}$ endowed with a p-adic metric: there are no isolated points, so $\mathbb{Q}_p$ is uncountable and not equal to $\mathbb{Q}$.

The above was just one example of the importance of distinguishing between countable and uncountable sets. Let me briefly mention some other examples:

Example 2: Measure theory. A measure is a $[0,\infty]$-valued function defined on a certain family of subsets of a given set; it is required to be countably additive but not uncountably additive. For instance, this gives us a natural notion of size on the unit circle, so that the total area is π and the area of any single point is 0. The whole can have greater measure than the sum of the measures of the parts if there are uncountably many parts!

Example 3: Given a differentiable manifold M of dimension n, then any submanifold of dimension $n-1$ has, in a sense which is well-defined independent of any particular measure on M, measure zero. In particular, one gets from this that a countable family of submanifolds of dimension at most $n-1$ cannot "fill out" an n-dimensional manifold. In complex algebraic geometry, such stratifications occur

naturally, and one can make reference to a "very general" point on a variety as a point lying on the complement of a (given) countable family of lower-dimensional subvarieties, and be confident that such points exist!

Example 4: Model theory is a branch of mathematics which tends to exploit the distinction between countable and uncountable in rather sneaky ways. Namely, there is the Lowenheim-Skolem theorem, which states in particular that any theory (with a countable language) that admits an infinite model admits a countable model. Moreover, given any uncountable model of a theory, there is a countable submodel which shares all the same "first order" properties, and conversely the countable/uncountable dichotomy is a good way to get an intuition on the difference between first-order and second-order properties.

2. Some further basic results

2.1. **Dedekind's characterization of infinite sets.**

Fact 14. *A set S is infinite iff it is equivalent to a proper subset of itself.*

Proof: One direction expresses an obvious fact about finite sets. Conversely, let S be an infinite set; as above, there is a countable subset $T \subset S$. Choose some bijection ι between T and $\mathbb{N}$. Then there is a bijection ι' between $T' := T \setminus \iota^{-1}(0)$ and T (just because there is a bijection between $\mathbb{N}$ and $\mathbb{Z}^+$. We therefore get a bijection between $S' := S \setminus \iota^{-1}(0$ and S by applying ι' from T' to T and the identity on $S \setminus T$.

This characterization of infinite sets is due to Dedekind. What is ironic is that in some sense it is cleaner and more intrinsic than our characterization of finite sets, in which we had to compare against a distinguished family of sets $\{[n] \mid n \in \mathbb{N}\}$. Thus perhaps we should define a set to be finite if it cannot be put in bijection with a proper subset of itself! (On the other hand, this is not a "first order" property, so is not in reality that convenient to work with.)

2.2. **An uncountable set not of continuum type.** Notice that in making the definition "uncountable," i.e., an infinite set which is not equivalent to $\mathbb{Z}^+$, we have essentially done what we earlier made fun of the "primitive tribes" for doing: giving up distinguishing between very large sets. In some sense, set theory begins when we attempt to classify uncountable sets up to equivalence. This turns out to be quite an ambitious project – we will present the most basic results of this project in the next installment – but there are a few further facts that one should keep in mind throughout one's mathematical life.

Let us define a set S to be *of continuum type* (or, more briefly, a continuum[7]) if there is a bijection $\iota : S \to \mathbb{R}$. One deserves to know the following:

Fact 15. *There exists an uncountable set not of continuum type, namely $2^{\mathbb{R}}$.*

Proof: By Theorem 12 there is no surjection from $\mathbb{R}$ to $2^{\mathbb{R}}$, so $2^{\mathbb{R}}$ is certainly not of continuum type. We must however confirm what seems intuitively plausible: that $2^{\mathbb{R}}$ is indeed uncountable. It is certainly infinite, since via the natural injection $\iota : \mathbb{R} \to 2^{\mathbb{R}}$, $r \mapsto \{r\}$, it contains an infinite subset. But indeed, this also shows

[7]This has a different meaning in general topology, but no confusion should arise.

that $2^{\mathbb{R}}$ is uncountable, since if it were countable, its subset $\iota(\mathbb{R}) \cong \mathbb{R}$ would be countable, which it isn't.

2.3. **Some sets of continuum type.** For any two sets S and T, we define T^S as the set of all functions $f : S \to T$. When $T = [2]$, the set of all functions $f : S \to [2]$ is naturally identified with the power set 2^S of S (so the notation is *almost* consistent: for full consistency we should be denoting the power set of S by $[2]^S$, which we will not trouble ourselves to do).

Fact 16. *The sets* $(0,1]$, $2^{\mathbb{Z}^+}$ *and* $\mathbb{R}^{\mathbb{Z}^+}$ *are of continuum type.*

Proof: Earlier we identified the unit interval $(0,1]$ in $\mathbb{R}$ with the infinite subsets of $\mathbb{Z}^+$ and remarked that, since the finite subsets of $\mathbb{Z}^+$ form a countable set, this implies that $(0,1]$ hence $\mathbb{R}$ itself is uncountable. Let us refine this latter observation slightly:

Lemma 17. *Let S be an uncountable set and $C \subset S$ an at most countable subset. Then $S \setminus C \cong S$.*

Proof: Suppose first that C is finite, say $C \cong [n]$. Then there exists an injection $\iota : \mathbb{Z}^+ \to S$ such that $\iota([n]) = C$ (as follows immediately from Fact 6). Let $C_\infty = \iota(\mathbb{Z}^+)$. Now we can define an explicit bijection β from $S \setminus C$ to S: namely, we take β to be the identity on the complement of C_∞ and on C_∞ we define $\beta(\iota(k)) = \iota(k-n)$.

Now suppose C is countable. We do something rather similar. Namely, taking $C_1 = C$, since $S \setminus C_1$ is uncountable, we can find a countably infinite subset $C_2 \subset S \setminus C_1$. Proceeding in this way we can find a family $\{C_i\}_{i \in \mathbb{Z}^+}$ of pairwise disjoint countable subsets of S. Let us identify each of these subsets with $\mathbb{Z}^+$, getting a doubly indexed countable subset $C_\infty := \bigcup_i C_i = \{c_{ij}\}$ – here c_{ij} is the jth element of C_i. Now we define a bijection β from $S \setminus C_1$ to S by taking β to be the identity on the complement of C_∞ and by putting $\beta(c_{ij}) = c_{(i-1)j}$. This completes the proof of the lemma.

Thus the collection of infinite subsets of $\mathbb{Z}^+$ – being a subset of $2^{\mathbb{Z}^+}$ with countable complement – is equivalent to $2^{\mathbb{Z}^+}$, and hence $(0,1] \cong 2^{\mathbb{Z}^+}$. So let us see that $(0,1]$ is of continuum type. One way is as follows: again by the above lemma, $[0,1] \cong (0,1)$, and $\mathbb{R}$ is even homeomorphic to $(0,1)$: for instance, the function

$$\arctan(\pi(x - \frac{1}{2})) : (0,1) \xrightarrow{\sim} \mathbb{R}.$$

For the case of $(\mathbb{Z}^+)^{\mathbb{R}}$: since $\mathbb{R} \cong 2^{\mathbb{Z}^+}$, it is enough to find a bijection from $(\mathbb{Z}^+)^{2^{\mathbb{Z}^+}}$ to $2^{\mathbb{Z}^+}$. This is in fact quite easy: we are given a sequence a_{ij} of binary sequences and want to make a single binary sequence. But we can do this just by choosing a bijection $\mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{Z}^+$.

A little more abstraction will make this argument seem much more reasonable:

Lemma 18. *Suppose A, B and C are sets. Then there is a natural bijection*

$$(A^B)^C \cong A^{C \times B}.$$

Proof of the Lemma: Indeed, given a function F from C to A^B and an ordered pair $(c,b) \in C \times B$, $F(c)$ is a function from B to A and so $F(c)(b)$ is an element of a.

Conversely, every function from $C \times B$ to A can be viewed as a function from C to the set A^B of functions from B to A, and these correspondences are evidently mutually inverse.[8] So what we said above amounts to

$$2^{\mathbb{Z}^+} \cong 2^{\mathbb{Z}^+ \times \mathbb{Z}^+} \cong (2^{\mathbb{Z}^+})^{\mathbb{Z}^+}.$$

Exercise 4: A subinterval of $\mathbb{R}$ containing more than one point is of continuum type.

It is also the case that $(\mathbb{Z}^+)^{\mathbb{Z}^+}$ is of continuum type. At the moment I do not see a proof of this within the framework we have developed. What we can show is that there exists an injection $\mathbb{R} \hookrightarrow (\mathbb{Z}^+)^{\mathbb{Z}^+}$ – indeed, since $\mathbb{R} \cong 2^{\mathbb{Z}^+}$, this is obvious – and also that there exists an injection $(\mathbb{Z}^+)^{\mathbb{Z}^+} \hookrightarrow 2^{\mathbb{Z}^+} \cong \mathbb{R}$.

To see this latter statement: given any sequence of positive integers, we want to return a binary sequence – which it seems helpful to think of as "encoding" our original sequence – in such a way that the decoding process is unambiguous: we can always reconstruct our original sequence from its coded binary sequence. The first thought here is to just encode each positive integer a_i in binary and concatenate them. Of course this doesn't quite work: the sequence 2, 3, 1, 1, 1 ... gets coded as 1011 followed by an infinite string of ones, as does the sequence 11, 1, 1, 1 ..... But this can be remedied in many ways. One obvious way is to retreat from binary notation to *unary* notation: we encode a_i as a string of i ones, and in between each string of a_i ones we put a zero to separate them. This clearly works (it seems almost cruelly inefficient from the perspective of information theory, but no matter).

Roughly speaking, we have shown that $(\mathbb{Z}^+)^{\mathbb{Z}+}$ is "at least of continuum type" and "at most of continuum type," so if equivalences of sets do measure some reasonable notion of their size, we ought to be able to conclude from this that $(\mathbb{Z}^+)^{\mathbb{Z}+}$ is itself of continuum type. This is true, a special case of the important Schröder-Bernstein theorem whose proof we defer until the next installment.

2.4. **Lots of inequivalent uncountable sets.** From the fundamental Theorem 12 we first deduced that not all infinite sets are equivalent to each other, because the set $2^{\mathbb{Z}^+}$ is not equivalent to the countable infinite set $\mathbb{Z}^+$. We also saw that $2^{\mathbb{Z}^+} \cong \mathbb{R}$ so called it a set of continuum type. Then we noticed that Cantor's theorem implies that there are sets not of continuum type, namely $2^{\mathbb{R}} \cong 2^{2^{\mathbb{Z}+}}$. By now one of the most startling mathematical discoveries of all time must have occurred to the reader: we can keep going!

To simplify things, let us use (and even slightly abuse) an obscure[9] but colorful notation due to Cantor: instead of writing $\mathbb{Z}^+$ we shall write $\beth_0$. For $2^{\mathbb{Z}+}$ we shall write $\beth_1$, and in general, for $n \in \mathbb{N}$, having defined $\beth_n$ (informally, as the n-fold iterated power set of $\mathbb{Z}^+$), we will define $\beth_{n+1}$ as $2^{\beth_n}$. Now hold on to your hat:

Fact 19. *The infinite sets $\{\beth_n\}_{n \in \mathbb{N}}$ are pairwise inequivalent.*

Proof: Let us first make the preliminary observation that for any nonempty set S, there is a surjection $2^S \to S$. Indeed, pick your favorite element of S, say x; for every $s \in S$ we map $\{s\}$ to s, which is "already" a surjection; we extend the mapping to all of 2^S by mapping every other subset to x.

[8] This is canonical bijection is sometimes called "adjunction."

[9] At least, *I* didn't know about it until recently; perhaps this is not your favorite criterion for obscurity.

Now we argue by contradiction: suppose that for some $n > m$ there exists even a surjection $s : \beth_m \to \beth_n$. We may write $n = m + k$. By the above, by concatenating (finitely many) surjections we get a surjection $\beta : \beth_{m+k} \to \beth_{m+1}$. But then $\beta \circ s : \beth_m \to \beth_{m+1} = 2^{\beth_m}$ is a surjection, contradicting Cantor's theorem.

Thus there are rather a lot of inequivalent infinite sets. Is it possible that the $\beth_n$'s are all the infinite sets? In fact it is *not*: define $\beth_\omega := \bigcup_{n \in \mathbb{N}} \beth_n$. This last set $\beth_\omega$ is certainly not equivalent to $\beth_n$ for any n, because it visibly surjects onto $\beth_{n+1}$. Are we done yet? No, we can keep going, defining $\beth_{\omega+1} := 2^{\beth_\omega}$.

To sum up (!!), we have a two-step process for generating a mind-boggling array of equivalence classes of sets. The first step is to pass from a set to its power set, and the second stage is to take the union over the set of all equivalence classes of sets we have thus far considered. Inductively, it seems that each of these processes generates a set which is not surjected onto by any of the sets we have thus far considered, so it gives a new equivalence class. Does the process ever end?!?

Well, the above sentence is an example of the paucity of the English language to describe the current state of affairs, since even the sequence $\beth_0$, $\beth_1$, $\beth_2 \ldots$ does not end in the conventional sense of the term. Better is to ask whether or not we can reckon the equivalence classes of sets even in terms of infinite sets. At least we have only seen countably many equivalence classes of sets[10] thus far: is it possible that the collection of all equivalence classes of sets is countable?

No again, and in fact that's easy to see. Suppose $\{S_i\}_{i \in \mathbb{N}}$ is any countable collection of pairwise inequivalent sets. Then – playing both of our cards at once! – one checks immediately that there is no surjection from any S_i onto $2^{\bigcup_{i \in \mathbb{N}} S_i}$. In fact it's even stranger than this:

Fact 20. *For no set I does there exists a family of sets $\{S_i\}_{i \in I}$ such that every set S is equivalent to S_i for at least one i.*

Proof: Again, take $S_{\text{bigger}} = 2^{\bigcup_{i \in I} S_i}$. There is no surjection from $\bigcup_{i \in I} S_i$ onto S_{bigger}, so for sure there is no surjection from any S_i onto S_{bigger}.

3. Some Final Remarks

Fact 20 is a truly amazing result. Once you notice that it follows readily from Cantor's Theorem 12, you may believe, as I do, that this theorem is the single most amazing result in all of mathematics.

There is also the question of whether this result is disturbing, or paradoxical. Can we then not speak of the set of all equivalence classes of sets (let alone, the set of all sets)? Evidently we cannot. There are too many sets to wrap all of them up into a single set. Some people have referred to this as **Cantor's Paradox**, although I do not favor this terminology: as far as I am aware, Cantor did not regard his results as paradoxical, nor do I. It does destroy the "ultranaive" notion of a set, namely, that given any "property" P, there is a set $S_P = \{x \mid P(x)\}$: according to Cantor's result, we cannot take P to be the property $x = x$. This was surprising in the late 19th century. But now we know of such things as Russell's paradox, which

[10]The day you ever "see" uncountably many things, let me know.

shows that the property $P(x)$ given by $x \notin x$ does not give rise to a set: the set of all sets which are not members of itself is a logical contradiction.[11]

But in truth it is hard to find anyone in the 21st century who has thought for more than a few hours about sets and is this naive, i.e., who thinks that every "property" of "objects" should give rise to a set. Indeed, as you can see from the quotation marks in the previous sentence, the idea that "all mathematical objects" is well-defined and meaningful has itself come to be regarded as problematic: what is the definition of a "mathematical object"? In some sense our idea of what sets are has come to be more dynamic and iterative following Cantor's work: we start with some simple sets and some operations (like union, subsets, and power sets), and by applying various procedures these operations allow us to create new and more complicated sets.

It is certainly true that deciding what "procedures" are legal is a difficult point: none of these procedures are of the sort that the truly finitistic mind need admit to as meaningful or possible. One can only say that in order to do mathematics the vast majority of us are willing to admit (indeed, unwilling to deny) the existence of certain infinite structures and processes: note that we began by saying "[w]e assume known the set $\mathbb{Z}^+$," i.e., we assumed the existence of an infinite set. If you decide to press on to read about a more explicit examination of what properties we think sets should satisfy, you will see that one of them baldly asserts the existence of infinite sets (of a certain kind). If we remove this axiom from the list, then the collection of sets $\{[n] \mid n \in \mathbb{N}\}$ becomes a model (in the sense of mathematical logic) for all the remaining axioms: that is, it is entirely consistent and logical to believe that sets of n elements exist for every n and not to believe that the collection of *all* n's makes sense as a set. It just happens to be extraordinarily useful and interesting – and, apparently, noncontradictory – to believe in the existence of infinite sets. When contemplating the "legality" of certain abstruse-looking set-theoretic constructions, it seems wise to keep in mind the leap of faith we make even to entertain $\mathbb{Z}^+$.

[11] I apologize for springing this so casually on the unfamiliar reader, but surely you've seen it before, no?