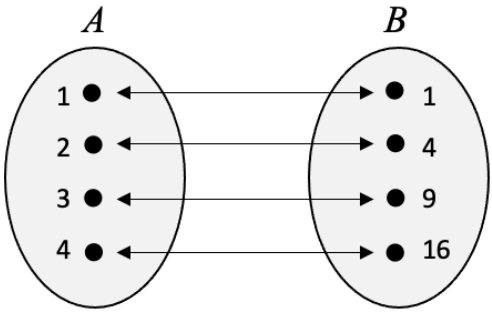
9. CARDINAL NUMBERS

§9.1. Equivalence of Sets

Most facts in the last few chapters will be well-known to you. The novelty was in redeveloping them within ZF set theory. We now turn our attention to infinite set theory itself. We begin with infinite cardinal numbers.

We can't count infinite sets in the same way as we do finite sets. We need a definition of the size of a set which, on the one hand agrees with our existing concept for finite sets, but which applies to infinite sets as well.

Do we simply invent a number, ∞, that we assign to all infinite sets? (Of course there remains the question of how we formally define 'finite' and 'infinite'.) We're perfectly entitled to do this but we would miss out on the interesting theory of transfinite numbers. We'd be in the same position of a certain tribe of aborigines who, it has been falsely claimed, had no word for numbers after 'three'. They were supposed to count "one, two, three, many".



Equivalent sets

Turn your mind back to the days when you first learnt to count. In kindergarten, we

pointed to things, or pictures of things, as we said aloud "one, two, three …". Essentially we were setting up a 1-1 correspondence between the things and a set of numbers that we learnt. If we got up to 'five' then we said that the number of things was five.

The concept of same-number-as is more fundamental than numbers themselves. A glance around a classroom of able-bodied people can reveal quickly that there's the same number of left arms as right arms. We don't need to count the left arms and the right arms and say "there are 27 left arms and 27 right arms so there must be the same number of each".

Recall that a **bijection** is a function that is 1-1 and onto. Two sets X, Y are **equivalent** ($\mathbf{X \approx Y}$) if there is a bijection $F:X \rightarrow Y$.

Clearly $\approx$ is an equivalence relation since it is reflexive, symmetric and transitive. However it's actually a generalised relation rather than a relation in the sense of a set of ordered pairs on the set of all sets, because there's no such thing as the set of all sets. Such an entity would give rise to the Russell Paradox.

Given a set S, how can we define its size, or cardinal number? As we saw in the last chapter, equivalence classes are out. We might think of choosing one particular representative of each size. That's what we'll do to begin with for finite numbers and the smaller

cardinal numbers, but as a general technique it has the problem of choosing a representative. We will need an additional axiom to do this. Let's postpone this problem and explore the basic facts about equivalence of sets.

Theorem 1: If $A \approx C$ and $B \approx D$ and A, B are disjoint and C, D are disjoint then $A \cup B \approx C \cup D$.
Proof: Let $F:A \to C$ and $G:B \to D$ be bijections. Define $H:A \cup B \to C \cup D$ by:

$$H(x) = \begin{cases} F(x) \text{ if } x \in A \\ G(x) \text{ if } x \in B \end{cases}.$$

The fact that A, B are disjoint means that H is well-defined. The fact that C, D are disjoint means that H is 1-1. It's clearly onto. ☺

Theorem 2: If $A \approx C$ and $B \approx D$ then $A \times B \approx C \times D$.
Proof: Suppose F, G are as above.
Define $H:A \times B \to C \times D$ by:
$H((x, y)) = (F(x), G(y))$. It's easy to check that this is a bijection. ☺

Theorem 3: $A \approx B \to \wp(A) \approx \wp(B)$.
Proof: Let $F:A \to B$ be a bijection.
Define $H: \wp(A) \to \wp(B)$ by:
$H(S) = \{F(x) \mid x \in S\}$ for all $S \in \wp(A)$. ☺

Size need not be preserved under intersections, however, as the following example shows.

Example 1: Let A = {1, 2, 3}, B = {2, 3, 5},
C = {1, 2, 3}, D = {3, 4, 5}.
Then A ≈ C and B ≈ D but A ∩ B = {2, 3} is not equivalent to C ∩ D = {3}.

It's intuitively obvious that if we have two equivalent sets, and take one element from each, the remaining sets will be equivalent. It's certainly obvious for finite sets but we want to show that it holds for infinite sets as well.

Theorem 4: If A ≈ B and $a \in$ A, $b \in$ B then
A – {a} ≈ B – {b}.
Proof: Let F: A→B be a bijection.
Case 1: F(*a*) = *b*: Then the restriction of F to A – {a} is a bijection between A – {a} and B – {b}.
Case 2: F(a) = c ≠ b: Let F(d) = b.
Then G: A–{a} → B–{b} defined by:
G(d) = c and
G(x) = x if $x \neq$ d, is a bijection. ✋☺

As we all know, if you remove one element from a finite set you have fewer elements. At this stage we haven't defined 'fewer' so we must be content to say that we have a different number of elements if we remove one element from a finite set.

Theorem 5: No natural number is equivalent to a proper subset of itself.

Proof: Let $P = \{n \in \mathbb{N} \mid n \approx$ a proper subset of $n\}$ and let $S = \mathbb{N} - P$.

So S is the set of all natural numbers that are not equivalent to a proper subset of themselves.

Clearly $0 \in S$ since the empty set has no proper subsets.

Suppose $n \in S$ and $n^+ \notin S$.

Then n^+ contains a proper subset m such that $m \approx n^+$.

Let $r \in n^+ - m$. Now since $n^+ = n \cup \{n\}$, m may or may not contain n.

Case 1: $n \in m$: Since $r \notin m$, $r \neq n$ and so $r \in n$.

Now $n^+ \approx m$ and $n \in n^+$ and $n \in m$ so by Theorem 4,

$n = n^+ - \{n\} \approx m - \{n\} \subseteq n$.

Now $r \notin m - \{n\}$ so $m - \{n\} \subset n$.

But this means that n is equivalent to one of its proper subsets and so $n \notin S$, a contradiction.

Case 2: $n \notin m$: Then $m \subseteq n$.

Let F be a bijection from n^+ to m.

Since $n \in n^+$ and $F(n) \in m$ we conclude from Theorem 4 that $n = n^+ - \{n\} \approx m - \{F(n)\}$.

But $m - \{F(n)\} \subset m \subseteq n$ and so $m - \{F(n)\}$ is a proper subset of n.

Again it follows that $n \notin S$, a contradiction.

It follows from Peano Axiom 5 that $\mathbb{N} = S$ and so no natural number is equivalent to one of its proper subsets. ✋☺

However, removing one element from an infinite set doesn't change its size. This is true for all infinite sets but here we only prove it here for $\mathbb{N}$, the set of all finite numbers.

Theorem 6: $\mathbb{N}$ is equivalent to a proper subset of itself.
Proof: $\mathbb{N} \approx \{n \in \mathbb{N} \mid n \neq 0\}$ under the bijection $F(n) = n^+$. ✋☺

Likewise, adding one element to an infinite set does not change its size. Again we only prove it for $\mathbb{N}$ at this stage.

Theorem 7: $\mathbb{N} \approx \mathbb{N}^+$.
Proof: Define $F(n^+) = n$, $F(0) = \mathbb{N}$.

It seems intuitively obvious that if $S \approx T$ then $S^+ \approx T^+$ but we can't prove it from the ZF axioms. We'd need the Axiom of Foundation. For suppose there was a set S for which $S = \{S\}$. Then $S^+ = S$ and although $S \approx \{0\} = 1$ it would not be the case that $S^+ \approx 1^+ = 2$.

§9.2. Inequalities With Sizes of Sets

Intuitively we have a notion, not just of two sets having the same size, but also of one set being smaller than another. Now by smaller we don't just mean 'proper subset' otherwise we couldn't compare the sizes of two disjoint sets. "Equivalent to a proper subset" would be better, and it is certainly adequate for finite sets.

Suppose we were in a classroom we saw that every student was seated and there was at least one empty student chair left over. We could say that there were more chairs than students. But if we made this concept the basis for 'less than' for infinite sets we would have to conclude that $\mathbb{N}$ is smaller than $\mathbb{N}^+$ despite them being the same size.

Instead we define 'less than or equals' first, by defining $S \leq T$ if S is equivalent to a subset of T and then defining $S < T$ to mean that $S \leq T$ and $S \neq T$.

We define $\mathbf{X \leq Y}$ if there is a 1-1 function $F:X \to Y$ and
$\mathbf{X < Y}$ if $X \leq Y$ and $X \neq Y$.

For natural numbers we now have two definitions of 'less-than-or-equals'. There is the one given in Chapter 4, where $m \leq n$ means '$m \in n$ or $m = n$', (or equivalently, '$m \subseteq n$'), and the one given here in terms of 1-1 functions. Naturally it would be very confusing if these gave a different ordering of the natural numbers.

Fortunately they are equivalent definitions, in the sense that $m \leq n$ under one definition if and only if $m \leq n$ under the other. To begin with we need to show that it's impossible to have a 1-1 function from n^+ to n.

Theorem 8: For all $n \in \mathbb{N}$ there is no 1-1 function from n^+ to n.
Proof: Suppose $n \in \mathbb{N}$ and suppose that $F:n^+ \to n$ is 1-1. Then $n^+ \approx \text{im}F \subseteq n \subset n^+$. This contradicts Theorem 5. ✋☺

Theorem 9: If m, n are natural numbers then $m \subseteq n$ if and only if there is a 1-1 map from m to n.

Proof: Suppose $m, n \in \mathbb{N}$.

Clearly, if $m \subseteq n$ there's a 1-1 function from m to n.

Suppose that there's a 1-1 function $F: m \to n$.

Either $m \in n$, $m = n$ or $n \in m$.

Suppose that $n \in m$.

Then $n^+ \subseteq m$.

Hence there exists a 1-1 function $G: n^+ \to m$.

Then $GF: n^+ \to n$ is a 1-1 function,

contradicting Theorem 10.

Hence $m \in n$ or $m = n$. In either case $m \subseteq n$. ✋☺

Theorem 10: Every proper subset of a natural number is equivalent to a smaller one.

Proof: We prove this by induction on n, the natural number. The statement holds for $n = 0$ vacuously since 0 has no proper subset.

Suppose it is true for n. If $S \subset n^+$ then

$$S \subset n \text{ or } S = n \text{ or } n \in S.$$

The conclusion is obvious in the first two cases.

Suppose $n \in S$. Then $S - \{n\} \approx m < n$ by induction.

Hence $S \approx m^+ < n^+$. ✋☺

§9.3. Finite and Infinite Sets

You'd probably say that most of the theorems in this section are obvious. For example, the set of natural numbers is infinite and any set that contains the natural numbers is infinite. But remember that we're trying not to rely on intuition. The purpose of these proofs is not to convince you, but rather convince a theorem-checking computer program that mechanically processes statements from the axioms.

How would you define the statements 'the set S is infinite' and 'the set S is finite'. Clearly we need only define one of these with one being the negation of the other. Here are some possible answers:

• S is infinite if $S \approx \mathbb{N}$: That's no good because $\mathbb{R}$ would not be infinite under this definition.

• S is infinite if $S \approx S^+$. That's better, and it would be adequate for developing all of standard mathematics. But suppose there was a set S, where $S = \{S\}$. Then $S^+ = S \cup \{S\} = S \cup S$. By this definition such as set would be infinite even though it clearly has only one element!

Now such sets play no role in mathematics and we could outlaw them by adopting an additional axiom to add to the standard ZF axioms.

• S is infinite if $\forall[x \in S \rightarrow S - \{x\} \approx S]$ In this case a set S, where $S = \{S\}$, is clearly finite, as we'd expect. Note,

also, that $\varnothing$ is finite vacuously because $x \in S$ is never TRUE.

• S is finite if $S \approx n$ for some $n \in \mathbb{N}$.

We'll adopt the last definition.

S is **finite** if $S \approx n$ for some $n \in \mathbb{N}$.
S is **infinite** if $\forall n \in \mathbb{N}[S \not\approx n]$.

Theorem 11: $\mathbb{N}$ is infinite
Proof: If $\mathbb{N} \approx n$ then $n \approx \mathbb{N} \approx \mathbb{N}^+ \approx n^+$, contradicting Theorem 5. ✋☺

Theorem 12: If S is infinite then $S \geq n$ for all $n \in \mathbb{N}$.
Proof: Suppose S is infinite.
We prove that $\forall n[n \in \mathbb{N} \rightarrow S \geq n]$ by induction (Peano axiom 5).
Let $T = \{n \in \mathbb{N} \mid S \geq n\}$.
Clearly $S \geq 0$. This is because the empty set can be viewed as a 1-1 function from 0 to S.
Hence $0 \in T$.
Suppose $n \in T$.
Then there exists a 1-1 function $F:n \rightarrow S$.
If F is onto then $S \approx n$ and so S is finite.
Hence there exists $s \in S$ such that $s \notin \text{im } F$.

Define G:$n^+ \to$ S by:

$$G(x) = \begin{cases} F(x) \text{ if } x \in n \\ s \text{ if } x = n \end{cases}.$$

Since F is 1-1 and $s \notin$ im F it follows that G is 1-1.
Hence $n^+ \in$ T.
Thus T = ℕ and so S ≥ n for all $n \in$ ℕ. ✋☺

Theorem 13: If S is infinite then S ≥ ℕ.
Proof: Suppose that S is infinite.
Then, for each $n \in$ ℕ, there exists a 1-1 function

$$F_n: n^+ \to S.$$

It would be tempting to say that F: ℕ → S defined by F(n) = F_n(n) is 1-1, but just because each F_n is 1-1 doesn't mean that F is. This is because the values of F_m(n) can vary with m. What we need is for F_{n+1} to be an extension of F_n for all $n \in$ ℕ.

Consider a triangular array:

$$\begin{array}{l} a_{11} \\ a_{21}\ a_{22} \\ a_{31}\ a_{32}\ a_{33} \\ \ldots\ \ldots\ \ldots\ \ldots\ldots \\ a_{n1}\ a_{n2}\ a_{n3}\ldots a_{nn} \end{array}$$

where (i) each $a_{mn} \in$ S;
(ii) $\forall n \forall m \geq n[a_{mn} = a_{mm}]$ and
(iii) $\forall m \forall r \leq m \forall t < r[a_{mr} \neq a_{ms}]$.

The elements in each column are equal and those in each row are distinct.

The last row gives a 1-1 function from n^+ to S.
Call such an array an S-triangle of depth n.

It's easy to see that if T_n is an S-triangle of depth n then there exists an S-triangle T_{n+1}, of depth $n + 1$, in which the first n rows are identical to T_n, for we define $a_{n+1,r} = a_{nr}$ if $r \leq n$.

Since S is infinite there exists $s \in S - \{a_{nr} \mid r \leq n\}$. Define $a_{n+1,n+1} = c$. Consequently we have an S-triangle T of infinite depth, with t_{ij} in the i-th row and j-th column.

Define $F: \mathbb{N} \to S$ by $t(n) = t_{nn}$ that is, the n'th entry on the diagonal of T.
It remains to show that F is 1-1. Suppose that $m < n$.
Being an S-triangle $t_{mm} = t_{nm} \neq t_{nn}$ and so $F(m) \neq F(n)$. ✋☺

Theorem 14: Subsets of finite sets are finite.
Proof: Suppose T is an infinite subset of the finite set S.
Suppose $S \approx n$ and let $F:S \to n$ be a bijection.
Since T is infinite, $T \geq n^+$.
Hence there exists a 1-1 function $G:n^+ \to T$.
Since $T \subseteq S$ there exists a 1-1 function $H:T \to S$.
Hence $GHF:n^+ \to n$ is a bijection so $n^+ \leq n$, a contradiction. ✋☺

Theorem 15: If S, T are disjoint finite sets then S + T is finite.
Proof: Suppose $S \approx m$ and $T \approx n$ where $m, n \in \mathbb{N}$.
Then there exist bijections $F:S \to m$ and $G:T \to n$.

Define the bijection $H:S + T \to m + n$ by:

$$H(x) = \begin{cases} F(x) \text{ if } x \in S \\ G(x) + m \text{ if } x \in T \end{cases}$$.✋☺

Theorem 16: If S, T are finite then so is $S \cup T$.
Proof: Suppose S, T are finite.
Then $S \cup T = S + (T - S)$ is finite by Theorems 14 and 15. ✋☺

Theorem 17: If S, T are finite then so is $S \times T$.
Proof: Suppose $T \approx n$. Then $S \times T \approx S \times n$.
We prove by induction on n that $S \times n$ is finite
for all $n \in \mathbb{N}$.
The inductive step relies on the fact that:

$$S \times n^+ \approx (S \times n) + S.$$ ✋☺

Theorem 18: If S, T are finite then so is S^T.
Proof: Suppose $T \approx n$. Then $S^T \approx S^n$.
We prove by induction on n that S^n is finite
for all $n \in \mathbb{N}$.
The inductive step relies on the fact that $S^{n^+} \approx S^n \times S$. ✋☺

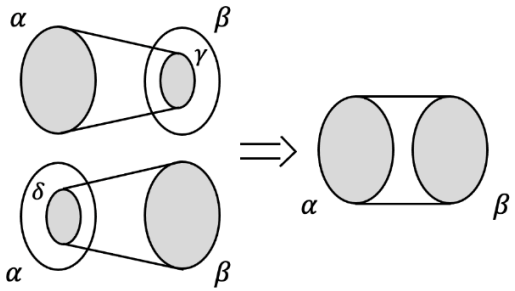
Theorem 19: $\wp(S) \approx 2^S$
Proof: Remember that 2^S is the set of functions from S to 2. As sets of ordered pairs these functions have elements of the form $(x, 0)$ or $(x, 1)$.

Let $F: \wp(S) \to 2^S$ be defined by $F(x) = \{(x, 1) \mid x \in S\}$ (this maps elements of X to 1 and all other elements of S to 0).
Let $G: 2^S \to \wp(S)$ be defined by $G(a) = \{x \mid a(x) = 1\}$.
It's easy to see that these are inverse functions and hence are bijections. ✋☺

Theorem 20: $S < \wp(S)$.
Proof: Clearly $S \leq \wp(S)$ since $F(x) = \{x\}$ is 1-1.
Suppose $F:X \to \wp(X)$ is a bijection.
Let $z = F^{-1}(\{x \mid x \notin F(x)\})$.
Then $z \in F(z)$ if and only if $z \notin F(z)$, a contradiction.
Corollary: $\omega < \wp(\omega) < \wp^2(\omega) < \ldots$

§9.4. The Schröder-Bernstein Theorem

So we've established that there are infinitely many sizes of infinite sets, although we haven't yet defined the corresponding infinite numbers. When we do that we'll expect that they'll be **comparable**, just like the natural numbers. That is, if α, β are any two infinite numbers we expect that either $\alpha < \beta$, $\alpha = \beta$ or $\beta < \alpha$, or equivalently that if $\alpha \leq \beta$ and $\beta \leq \alpha$ then $\alpha = \beta$. But for the time being we must be content to express this in terms of sets rather than their sizes. It's



known as the Schröder-Bernstein Theorem. Suppose there's a 1-1 function from X to Y and another from Y to X. From these two we have to somehow manufacture a bijection between these two sets. It's not easy.

Theorem 21 (Schröder-Bernstein):
If $X \leq T$ and $Y \leq X$ then $X \approx Y$.
Proof: Suppose that $F:X \to Y$ and $M:Y \to X$ are 1-1 functions.
Let $y \in Y$.
If $x \in$ im M then $M^{-1}(x)$ is defined and is in Y.
If $M^{-1}(x) \in$ im F then $F^{-1}M^{-1}(x) \in Y$ is defined.
So either $(F^{-1}M^{-1})^n(x)$ is defined for all n or there exists a largest n such that $(F^{-1}M^{-1})^n(x)$ is defined. In the latter case, either $(F^{-1}M^{-1})^n(x) \in$ im M or it is not.
This gives rise to three cases:
(1) $(F^{-1}M^{-1})^n(x)$ is defined for all n;
(2) There is a largest n such that $(F^{-1}M^{-1})^n(x)$ is defined and *is not* in im M;
(3) There is a largest n such that $(F^{-1}M^{-1})^n(x)$ is defined and *is* in im M.

In case (2) we say that $(F^{-1}M^{-1})^n(x)$ is called the **ultimate ancestor** of x. We denote it by $\boldsymbol{\alpha(x)}$ and $\alpha(x) \in X$.
In case (3) we say that $M^{-1}(F^{-1}M^{-1})^n(x)$ is the **ultimate ancestor** of x.
Again we denote it by $\boldsymbol{\alpha(x)}$ and $\alpha(x) \in Y$.
In case (1) x has no ultimate ancestor.

Let $X_1 = \{x \in X \mid x \text{ has no ultimate ancestor}\}$.
This corresponds to case (1).
Let $X_2 = \{x \in X \mid \alpha(x) \in X\}$.
This corresponds to case (2).
Let $X_3 = \{x \in X \mid \alpha(x) \in Y\}$.
This corresponds to case (3).

Similarly we define ultimate ancestors of elements of Y, by swapping X and Y and swapping M and F.

Let $Y_1 = \{y \in Y \mid y \text{ has no ultimate ancestor}\}$.
Let $Y_2 = \{y \in Y \mid \alpha(y) \in Y\}$.
Let $Y_3 = \{y \in Y \mid \alpha(y) \in X\}$.

Then $M(X_1) = Y_1$,
$M(X_2) = Y_3$ and
$F(Y_2) = X_3$.
Moreover, M restricted to X_1 and M restricted to X_2 are both 1-1 and F restricted to Y_2 is 1-1.
Define $H:X \to Y$ by:

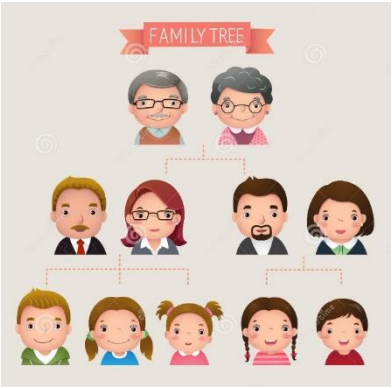
$$H(x) = \begin{cases} M(x) \text{ if } x \in X_1 \\ M(x) \text{ if } x \in X_2 \\ F^{-1}(x) \text{ if } x \in X_3 \end{cases}.$$

Then H is a bijection from X to Y and its inverse is

$$H^{-1}(y) = \begin{cases} M^{-1}(y) \text{ if } y \in Y_1 \\ M^{-1}(y) \text{ if } y \in Y_2 \\ F(y) \text{ if } y \in Y_3 \end{cases}.$$

You may find difficulty in following this argument. Consider the following analogy involving family trees. Suppose X is the set of all males who were ever born in Australia and suppose Y is the set of all females ever born in Australia (now living or dead). We'll assume that every father and every mother has exactly one son and one daughter. But not every person in X or Y has children. (Ignore the younger sister in the following picture.)

Let M:X $\rightarrow$Y take a male to his mother and let F:Y $\rightarrow$ X take a female to her father. Under our assumptions, F and G are 1-1 functions. You can't have two males having the same mother because we're assuming that each family with children only has one son and one daughter.



We'll also assume that these sons and daughters are born in Australia. So once somebody comes to Australia all their descendants will be born in Australia. It is, after all, a good country to live in! Why would anyone want to leave it?!

As a model for Australian families it isn't very good. But as an analogy for the proof of the Schröder-Bernstein Theorem it works well. So MF will take a male

to his maternal grandfather and FM will take a female her paternal grandmother.

Now we're going to go back up the family tree. Take a man. His mother might have been born in Australia and her father might have also been born in Australia. Going back up the family tree we may eventually reach an ancestor who wasn't born in Australia. Because Australia is a young country, I suppose we'd always reach an ancestor who was born overseas. Even indigenous people are not really '*aboriginal*' because if you could trace back their family tree for forty thousand years or more you'd probably reach someone born in Indonesia.

But to make this analogy work we'll also need to assume that there have been people living in Australia forever. Not just for 40,000 years but forever! So, under this assumption, there would be Australians (no doubt they'd be the *real* aboriginals) where, if you traced back their family tree, in the alternating way I've described, their infinitely many ancestors were *all* born in Australia. (I'm not casting any aspersions on the indigenous population of Australia. Although they were once immigrants themselves, they've lived here far longer than Europeans have been in Europe!) For other Australians, tracing back in this way, we'd reach someone who was born overseas. Call such an ancestor, an **ultimate ancestor**. Now such an ultimate ancestor will be either a man or a woman.

Let X_1 be the set of all males with no ultimate ancestor.
Let X_2 be the set of all males whose ultimate ancestor was male.
Let X_3 be the set of all males whose ultimate ancestor was female.
Let Y_1 be the set of all females with no ultimate ancestor.
Let Y_2 be the set of all females whose ultimate ancestor was female.
Let Y_3 be the set of all females whose ultimate ancestor was male.
Now follow the argument in the theorem.

§9.5. Infinite Cardinal Numbers

We still have some way to go before we are in a position to define cardinal numbers, the numbers that describe the sizes of sets. For the time being we'll regard them as objects that are associated with sets so that equivalent sets have the same number. We denote the **cardinal number** (size) of a set S by the symbol **#S**.

We can still prove theorems about these, as yet undefined, numbers because each of those theorems can be expressed in terms of the relations $\leq$ and $\approx$.

We follow the notation of Georg Cantor, who first considered infinite cardinal numbers, and denote the size of $\mathbb{N}$ by the symbol $\aleph_0$.

We define $\#\wp\mathbb{N}$ to be $\aleph_1$. This notation suggests that this is the next cardinal number after $\aleph_0$ and hence that there is no cardinal number between $\aleph_0$ and $\aleph_1$. If there was we might have to write it as something like $\aleph_{0.5}$, which we would be a bit clumsy.

In fact most books define $\aleph_1$ as the next cardinal number after $\aleph_0$ and the question is asked "Is $\aleph_1 = \#\wp\mathbb{N}$?" But we're defining $\aleph_1 = \#\wp\mathbb{N}$ and are asking "is there a cardinal number between $\aleph_0$ and $\aleph_1$?"

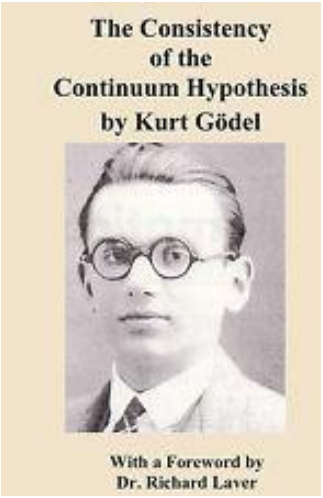
It can be shown that neither question can be answered. The **Continuum Hypothesis** states that the answers to the above questions are "yes" and "no" respectively.

For we who define $\aleph_1 = \#\wp\mathbb{N}$ the Continuum Hypothesis states that there is no cardinal number between $\aleph_0$ and $\aleph_1$. For those who define $\aleph_1$ to be the next number after $\aleph_0$ it states that $\aleph_1 = \#\wp\mathbb{N}$. We'll state it in terms of our definition of $\aleph_1$.

Continuum Hypothesis: There is no cardinal number between $\aleph_0$ and $\aleph_1$.

The name of this hypothesis arose from is fact that we can prove that the set of real numbers (the continuum) is equivalent to $\wp(\mathbb{N})$. Then $\#\mathbb{N} = \aleph_0$ and $\#\mathbb{R} = \aleph_1$ and

the continuum hypothesis asserts that there is no cardinal number between them.



Now this statement has never been proved. It *will* never be proved. It *can* never be proved. There's a theorem that proves that the Continuum Hypothesis is unprovable. It's too long and too technical to present here. But it's important that you understand what this theorem is saying.

Theorem 22: The Continuum Hypothesis is consistent with, and independent of, the ZF axioms.
Proof: As we said we won't be presenting a proof, but here is some discussion of what such a proof involves. ✋

A set of axioms is consistent if no contradiction can arise from accepting them. The standard way of proving that a set of axioms is consistent is to construct a model that satisfies them. The group axioms are consistent because we can construct explicit examples of groups.

Now the annoying thing is that the ZF axioms have never been proved to be consistent. While it has never been proved that their consistency is unprovable, it seems reasonable to believe that it is. The problem is that in constructing models you have to make them out of

something, and if we can't use sets it's not clear what we could use.

The consequence of the fact that the ZF axioms haven't been proved consistent is that there's the possibility that some new paradox, like the Russell Paradox, might one day come to light. The whole of mathematics would seem to be built on a shaky foundation that could come tumbling down at any moment.

But mathematicians aren't worried by such a possibility. If such a calamity should ever arise all that would happen would be that the set theory axioms would be modified to fix the problem. The edifice of mathematics would still stand and most practising mathematicians wouldn't even notice the change. That is indeed what happened when the Russell Paradox was pointed out.

To prove consistency we'd need to construct a model that satisfies the ZF axioms as well as the Continuum Hypothesis. We'd begin by assuming we have a model for the ZF axioms (thereby assuming that the ZF axioms are consistent). Then we construct special sets, and a membership relation between them, that is similar to the ones in ordinary set theory but which is modified in some way. We would then prove that in this modified model the Continuum Hypothesis holds.

The upshot is a sort of *relative* consistency. If ZF is consistent then ZF + CH is consistent. If a paradox ever

arises when using the Continuum Hypothesis there must be an inconsistency in the ZF axioms themselves.

The second part of the proof involves constructing a second model, again assuming that the ZF axioms alone are consistent. This second model is constructed in such a way that it's possible to find a cardinal number between what corresponds, in this model, to $\aleph_0$ and $\aleph_1$. In other words we prove that if ZF is consistent then so is ZF + not(CH). If ever a paradox arises when assuming the existence of a cardinal number between $\aleph_0$ and $\aleph_1$ there must be an inconsistency in the ZF axioms alone.

Such models showing the consistency and independence of the Continuum Hypothesis have been constructed. ✋

It comes down to what, in the religious world, would be called a matter of faith! We are logically free to either assert or deny the Continuum Hypothesis. But, while either possibility is equally valid, there's a practical reason for accepting it.

Certainly, while we can't *prove* that there are no cardinal numbers between $\aleph_0$ and $\aleph_1$ we'll never actually be able to *find* one. Clearly, if we could find such an explicit example we would, as a consequence, have a proof that the Continuum Hypothesis is false and this would contradict the proof that it is independent from the ZF axioms.

In other words, we apply *Occam's Razor*, a principle that "of all the hypotheses that fit the facts, choose the simplest". Throughout the rest of the notes we'll assume the Continuum Hypothesis in addition to the ZF axioms.

We define the infinite cardinal numbers $\aleph_0$, $\aleph_1$, $\aleph_2$, … inductively as follows:

$$\aleph_0 = \#\mathbb{N};$$
$$\aleph_{n+1} = \#\wp(\aleph_n) \text{ for } n \in \mathbb{N}.$$

§9.6. The Arithmetic of Cardinal Numbers

I define **addition**, **multiplication**, and **exponentiation** of cardinal numbers as follows:

Suppose $a = \#A$ and $b = \#B$, where A, B are disjoint. Then:

$\boldsymbol{a + b}$ is defined to be $\#(A + B)$;

$\boldsymbol{ab}$ is defined to be $\#(A \times B)$;

$\boldsymbol{a^b}$ is defined to be $\#(A^B)$.

The set of functions from $2 = \{0, 1\}$ to A is equivalent to $\wp(A)$ so $\#\wp(A) = 2^a$.

The following arithmetic properties can be proved readily by setting up suitable bijections.

Theorem 23:

(1) $a + b = b + a$
(2) $ab = ba$
(3) $a + (b + c) = (a + b) + c$
(4) $(ab)c = a(bc)$
(5) $a(b + c) = ab + ac$
(6) $1a = a$
(7) $0a = 0$
(8) $0^a = 0$ (if $a > 0$), or 1 if $(a = 0)$
(9) $a \leq a$
(10) If $a \leq b$ and $b \leq c$ then $a \leq c$.
(11) If $a \leq b$ then $a + c \leq b + c$.
(12) If $a \leq b$ then $ac \leq bc$.
(13) If $a \leq b$ then $a^c \leq b^c$.

Proof: These follow from the fact that corresponding sets are equal, or we can easily set up a bijection between them. ✋☺

Theorem 24:

(1) $a^{b+c} = a^b.a^c$.
(2) $(a^b)^c = a^{bc}$.

Proof: Let A, B, C be disjoint sets such that #A = a, #B = b and #C = c.
(1) Define $\theta: A^{B+C} \rightarrow A^B \times A^C$ by $\theta(f) = (f|_B, f|_C)$.
(2) Define $\theta: (A^B)^C \rightarrow A^{B\times C}$ by $\theta(f)(x, y) = f(y)(x)$. ✋☺

Theorem 25: If $1 < a \leq b$ then $a + b \leq ab$.
Proof: Let A, B be disjoint sets such that

$$\#A = a \text{ and } \#B = b.$$

Let $a_0, a_1 \in A$ and $b_0 \in B$. Define $f: A + B \to A \times B$ by

$$f(x) = \begin{cases} (x, b_0) \text{ if } x \in A \\ (a_0, x) \text{ if } x \in B \end{cases}.$$ ✋☺

Theorem 26: If $0 < a \leq b$ then $c^a \leq c^b$.
Proof: Let A, B, C be sets such that $\#A = a$, $\#B = b$ and $\#C = c$.
If $c = 0$ the result is obvious, so suppose that $c > 0$.
Let $f: A \to B$ be a 1-1 function and let $u \in C$.
If f is a function from A to C define $\theta(f): B \to C$ by

$$\theta(f)(x) = \begin{cases} f(u^{-1}(x)) \text{ if } x \in \text{im } f \\ u \text{ if } x \notin \text{im } f \end{cases}.$$ ✋☺

Theorem 27: If $a \geq 2$ then $b < a^b$.
Proof: We have done the case $a = 2$.
So $b < 2^b \leq a^b$ by Theorem 28. ✋☺

Theorem 28: $\aleph_0 + a = \aleph_0$ for all $a \in \omega$.
Proof: Define $f: \mathbb{N} \to \mathbb{N}$ by $f(x) = x + a$.
$\#\text{im } f = \aleph_0$ and $\mathbb{N} = \text{im } f + a$. ✋☺

Theorem 29: $2\aleph_0 = \aleph_0$.
Proof: Define $f: \mathbb{N} \times 2 \to \mathbb{N}$ by $f(x, y) = \begin{cases} 2x \text{ if } y = 0 \\ 2x + 1 \text{ if } y = 1 \end{cases}$.
✋☺

Theorem 30: $\aleph_0^2 = \aleph_0$.
Proof: Define f: $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by:

$$f(x, y) = \tfrac{1}{2}(x + y)(x + y + 1) + y.$$ ✋☺

Theorem 31: $\aleph_m + n = n\aleph_m = (\aleph_m)^n = \aleph_m$
for all $0 \neq n \in \mathbb{N}$.
Proof: Use Theorems 30, 31 and 32, as well as induction. ✋☺

§9.7. Examples of Cardinal Numbers

Example 1: $\#\mathbb{Z} = 2\aleph_0 + 1 = \aleph_0$.

Example 2: $\aleph_0 \leq \#\mathbb{Q} \leq \aleph_0^2 = \aleph_0$, so $\#\mathbb{Q} = \aleph_0$.

Example 3:
$\aleph_1 = \#\{\text{binary sequences}\}$
$\leq \#[0,1)$ (represent these reals in decimals)
$\leq \#\{\text{binary sequences}\}$ (represent them in binary)
$\leq \aleph_1$.
Hence, by Schröder-Bernstein:
$\#[0,1) = \#\text{binary sequences} = \aleph_1$.

Every real number corresponds to a pair (n, x) where the integer part $n \in \mathbb{Z}$ and the fractional part x belongs to the interval $[0, 1)$.
Hence $\#\mathbb{R} = \aleph_1.\aleph_0 = \aleph_1$.
Thus $\#\mathbb{C} = \aleph_1^2 = \aleph_1$.

§9.8. Applications of Cardinal Numbers

Essentially the study of infinite set theory is an end in itself and for the most part mathematicians happily go on with their work without paying much attention to it. But there are some places where a theorem in some other area can be proved by applying infinite set theory.

An **algebraic number** is a complex number α such that $f(\alpha) = 0$ for some non-zero integer polynomial. The rest, if any, are called **transcendental**. While it's a little difficult to prove that specific numbers, such as e and π, are transcendental, the existence of transcendental numbers can be proved by observing that the number of algebraic numbers is $\aleph_0$ while the number of complex numbers altogether is $\aleph_1$. The fact that $\aleph_0 < \aleph_1$ shows that transcendental numbers must exist, and plenty of them!

Computer programs can be written that compute a specific real number by printing out its decimal expansion. Although it can only ever print out a finite number of decimal places in finite time, they can be written so that every decimal place will eventually appear if the program runs for long enough. A program to print out the decimal equivalent of a rational number such as 22/7 would be very easy. One that printed out the decimal expansion of π would be somewhat harder, but it can be done. Can a suitable program be written for *every* real number. The answer is most definitely "no". There are $\aleph_1$ possible real numbers. But since a computer program is a finite string of symbols there are only $\aleph_0$ programs that

are possible, even if we had infinite time to write them. Therefore some (in fact the vast majority) of real numbers cannot be computed.

To give a specific example of such a number would be difficult; because once a number is described precisely it's not too difficult to write a suitable computer program. The problem is that there can be only $\aleph_0$ possible descriptions so only $\aleph_0$ real numbers that can be defined.

Most of the applications of infinite cardinal numbers depend on recognising the difference between $\aleph_0$ and $\aleph_1$. Occasionally $\aleph_2$ is used outside of infinite set theory. But the bigger cardinal numbers are never used. Their interest lies solely is satisfying one's curiosity as to what is out there.

§9.9. Even Bigger Cardinal Numbers

The list of cardinal numbers that we've described (though not yet defined as sets) goes well beyond what a mathematician would find useful. Yet these very large infinite numbers deserve to have their existence acknowledged. The situation is analogous to astronomers discovering remote galaxies that are so far from earth that. their practical significance is zero. Yet we find them fascinating.

The numbers we've discovered so far come in two infinite lists: 0, 1, 2, 3, ……

$\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ldots\ldots\ldots\ldots$

That's exciting enough, but as they say in TV ads for steak knives, "there's more". That is, there's an infinite cardinal number that is bigger than *any* of the $\aleph_n$'s!

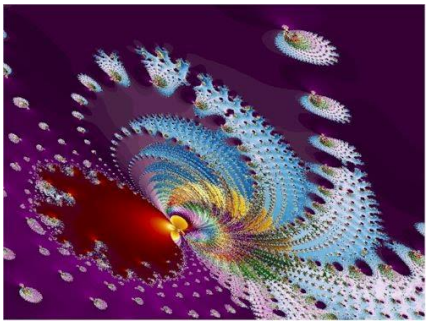
For each $n \in \mathbb{N}$ take a set S_n such that $\#S_n = \aleph_n$. For example we could take S_n to be $\wp^n\mathbb{N}$. Now take their union $\cup\{S_n \mid n \in \mathbb{N}\}$. We define $\aleph_\omega$ to be the cardinal number of this set.

Theorem 32: For all $n \in \mathbb{N}$, $\aleph_\omega > \aleph_n$.
Proof: Let $n \in \mathbb{N}$. Then $S_{n+1} \subseteq \cup\{S_n \mid n \in \mathbb{N}\}$
and so $\aleph_{n+1} \leq \aleph_\omega$.
But $\aleph_n < \aleph_{n+1}$ and so $\aleph_n < \aleph_\omega$. ✋☺

And, of course, it doesn't stop there. We can define $\aleph_{\omega+1} = 2^{\aleph_\omega}$ and so begin a third row in our table. Having described the infinitely long 3rd row of cardinal numbers we can carry out a similar construction to the above get a 4th row, the first entry of which is denoted by the symbol $\aleph_{\omega 2}$.

Why not $\aleph_{2\omega}$? Ah, well, the symbol ω represents something called an *ordinal number* and with the multiplication we'll define then, $2\omega = \omega$ while $\omega 2$ is bigger. But you'll have to wait till we've discussed ordinal numbers.

So we can produce infinitely many rows, each with infinitely many cardinal numbers, and we wouldn't have them all. We can produce a second page corresponding to the union of all the sets that correspond to the numbers on



the first page etc. Before long we'd have infinitely many libraries, each with infinitely many floors, each with infinitely many rooms, each with infinitely many shelves, each with infinitely many books, each with infinitely many pages, each with infinitely many rows, each infinitely long, and still there'd be even bigger numbers.

In fact it is easy to see that for every set of cardinal numbers there's a cardinal number bigger than them all. If the set has a largest just raise 2 to that power. If it doesn't, the same process that got us to $\aleph_\omega$ could be used.

So what about the set of *all* cardinal numbers? Can there really be one bigger than them all, which of course would mean that it would be bigger than itself? Certainly not. Yet we have outlined a proof, haven't we?

Not quite. This paradox simply means that the class of all cardinal numbers is not a set. It must therefore be a proper class.

We now know quite a bit about cardinal numbers but we haven't actually defined them. They're defined to be ordinal numbers with a certain property so we need to define 'ordinal number' first. But there is something more important to discuss in the next chapter – the Axiom of Choice. This is probably the most important part of Axiomatic Set Theory for other areas of mathematics.

§9.10. Further Arithmetic of Cardinal Numbers

Theorem 33: If a, b are cardinal numbers and a is finite and b is infinite then $a + b = b$.

Proof: Choose disjoint sets A, B such that

$$\#A = a \text{ and } \#B = b.$$

Let $C \subseteq B$ such that $\#C = \aleph_0$ and let $D = B - C$.

Let $\#D = d$.

Then $b = \aleph_0 + d$ so $a + b = a + \aleph_0 + d$

$$= \aleph_0 + d$$

$$= b. \text{ ✋☺}$$

Theorem 34: If a is an infinite cardinal number then:

$$a + a = a.$$

Proof: Choose A so that $\#A = a$.

Let $F = \{\text{bijections } f : X \times 2 \to X \mid X \subseteq A\}$.

$F \neq 0$ (take X with $\#X = \aleph_0$).

F is partially ordered by extension.

By Zorn's Lemma there exists a maximal function

$$f: X \times 2 \to X \text{ for some } X \subseteq A.$$

If A – X is infinite this contradicts the maximality of f so
A – X is finite.
#X + #X = #X and #A = #X + #(A – X) so
#A + #A = #X + #X + 2#(A – X)
= #A + #(A – X)
= #A. ✋☺

Theorem 35: If $a \leq b$ are cardinal numbers and b is infinite then $a + b = b$.
Proof: Choose A, B with #A = a, #B = b.
Since $a \leq b$, $a + b \leq b + b = b$.
But $b \leq a + b$ so $a + b = b$. ✋☺

Theorem 36: If a is an infinite cardinal number then:
$a.a = a$.
Proof: Choose A so that #A = a.
Let F = {bijections f : X × X → X | X ⊆ A}.
F ≠ 0 (take X with #X = $\aleph_0$).
F is partially ordered by extension.
By Zorn's Lemma there exist maximal f: X × X → X
for some X ⊆ A.
Let #X = x. Then $x.x = x$.
Suppose $x < a$.
Then #(A – X) = a and so A – X has a subset, Y
with #Y = x.
Then #[(X × Y) + (Y × X) + (Y × Y)] = $3x.x$
= x so
there exists a bijection from

$$(X \times Y) + (Y \times X) + (Y \times Y) \text{ to } Y.$$

We can thus extend f to a bijection

g: $(X + Y) \times (X + Y) \to X + Y$, a contradiction.

Hence $x = a$ and so $a.a = a$. 👌☺