Chapter 1

1.1 Irrationality of root-2

A famous mathematician G.H. Hardy suggests that real, “authentic” mathematics is best justified as an art.

He uses two classical Greek proofs to show the beauty of math: the proof of infinite primes, and, more relevant to our study, the proof that 2 is irrational.

Theorem 1.1.1: There is no rational number whose square is 2.

We use a proof by contradiction, where we follow the opposite of our statement until we reach an impossible conclusion, showing that our statement cannot be wrong.

Proof (by contradiction):
Any rational number can be written as p/q, where p and q are integers with no common factors.

We should be able to find: $(p/q)^2 = 2$.

p^2 must be even to be twice as large as q^2, so p must be even.

However, if p is even, then p^2 has a factor of 4.

Half of 4 is a factor of two, so p^2 and q^2 share a factor of 2.

We assumed they have no common factors, so have a contradiction.

p/q will never be $\sqrt{2}$, thus, $\sqrt{2}$ is not rational.

In short:
-p and q have no common factors
-p^2 is even→ p is even→ p^2 has a factor of 2*2
-q^2 is half of p^2→ q^2 has a factor of 2, is even→ q is even
-p and q are both even, but have no common factors? Impossible

----

Part of Hardy's beauty is impact, and here, the Greeks learned that rational numbers cannot describe all lengths (in this case, a right triangle's hypotenuse), something they assumed until then.

Number Systems

We want to expand the idea of the number to match this: we first start, with the "counting", or natural numbers

N= {1, 2, 3, 4, 5...}

This works for addition, but if we want subtraction(the "inverse", or opposite of addition), as well as zero (the additive "identity": has no effect when added), we need the integers:

Z = { …−3, −2, −1, 0, 1, 2, 3, ... }

This allows multiplication, but if we want division(the "inverse" of multiplication), we need fractions, or the rational numbers:

Q = { p/q for all integers p and q, where $q \neq 0$}

The ability to do addition and its inverse, multiplication and its inverse, and use the: commutative, associative, and distributive properties, means that the rational numbers are what is called a field.

Number Properties

Q also has a "natural order", meaning that for any numbers r and s, either r<s, r=s, or r>s. Notably, that means if you disprove any two of the three, the third must be true.

This order is transitive, meaning if r<s and s<t, then r<t, so we can imagine all rational numbers lined up next to each other, along the number line.

There is no empty space like in N or Z, because halfway between any two rational numbers, there is another rational number.

However, because of the above proof, we know the rational numbers do NOT include all square roots.

We can approximate those numbers very closely with rationals, but clearly they are missing from our dense set of numbers.

We create the real numbers, which include the rational numbers, and every "irrational" number that made a hole in our number line.

Now, we study the properties of these rational and irrational numbers: the study of Real Analysis.

1.2 Some preliminaries

We begin by being clear about our definitions, using some "set theory".

Definitions

Set: A group/collection of things.
Elements: The things within a set.

This symbol means that x is an element in the set A.
$x \in A$

Union of two sets: The new set where each element was in at least one of the two sets.
The union of [A or B] ⇔ $A \cup B$

Intersection of two sets: The new set where each element was in BOTH sets.
The intersection of [A and B] ⇔ $A \cap B$

Sets

We can represent a set by:
i) Listing the first few elements that follow a pattern
$N = \{1, 2, 3 \ldots\}$

ii) Describing with words
E is the collection of even natural numbers.

iii) Creating a rule
$S = \{ r \text{ in } Q : r^2 < 2 \}$
or
$S = \{ r \in Q : r^2 < 2 \}$

Let S be the set of all rational numbers whose squares are less than two.

In this notation, r is our placeholder for any number in S.
As said before, $\in$ just means "the thing before is IN the thing after".

We also define a set that contains no elements: the empty set.

For example, {1,2} AND {3,4} = EMPTY

If two sets have no elements in common like this, they are called disjoint.

We introduce **inclusion** relationships, which compare sets:

$A \subseteq B$ (or $B \supseteq A$) means that every element of A is in B. We can say A is a **subset** of B, or B **contains** A (B is a **superset** of A, but we use that language less often.)

It isn't a coincidence that $\subseteq$ and $\leq$ look similar: they have similar meanings. That is, one set is "inside" another, but doesn't necessarily have fewer elements. How?

If $A \supseteq B$ AND $A \subseteq B$, the two sets have the same elements, so we say $A = B$. They're both subsets of each other, those subsets just happen to include every element in both.

Again, we compare to inequalities: if $a \geq b$ AND $a \leq b$, then $a = b$.

But this idea is pretty broad, and includes things you might not feel like should be "properly" called a subset. I mean, $A \supseteq A$ seems almost redundant.

So, we'll create a **proper subset**: If $A \subset B$, then A is a proper subset of B. To make it proper, we'll just say that $A \neq B$: B must have some elements that A does not.

That definition feels more like A is "inside" of B. But more importantly, this is a useful distinction, in the same way that $>$ and $\geq$ deserve separate symbols.

Often, we will want to find unions or intersections to an infinite collection of sets. These use their own notation.

Introduce infinite union and infinite intersection notation

For our example, let's use the sets:

$A_1 = N = \{1,2,3\ldots\}$ $A_2 = \{2,3,4\ldots\}$ $A_3=\{3,4,5\ldots\}$

In general, $A_n = \{n, n+1, n+2...\}$

For some examples:

Thus, we can say

$A_1 \supseteq A_2 \supseteq A_3 \supseteq A_4 \ldots$

If we want the union of all these sets, we want all elements that are contained in at least one set.

Because A_1 contains all other involved sets, the union is just A_1: no other set has any new elements.

Now, we wish to take the intersection of all sets. We try to find an element m in this intersection: however, no matter what number we pick, m is not in A_{m+1}. Thus, this intersection is the empty set.

If A is a subset of R, the real numbers, we can talk about every element of R that is NOT in A. This is called the complement of A, or A^c.

$A^c = \{x \text{ in } R : x \text{ not in } A\}$
$A^c = \{ x \in R : x \notin A \}$

Let Ac be the set of all real numbers which are not in A.

De Morgan's Laws

We introduce the De Morgan's Laws, which can be proven but we will simply state here, for their usefulness:

$(A \cap B)^c = A^c \cup B^c$
$(A \cup B)^c = A^c \cap B^c$

We note here that our definition of sets is partly intuitive, but have to start from some level: with every new level of detail, one layer below appears. We will leave those later details to other studies.

Functions

Function: Given sets A and B, takes every element of A and pairs it with one element of B.

$f: A \rightarrow B$

Notation - if x is an element in A, f(x) is the matching element in B.

A is the domain of f; the part/subset of B that is included by some f(x) is the range. Note that the range may not include all of B.

For a formal version of the range, take:

$S = \{ y \in B : y = f(x) \text{ for some } x \text{ in } A \}$

Let S be the set of all elements in B which equal some f(x), where x is an element of A.

This idea of a function, inspired by the recurring mathematician Dirichlet, is much more broad than some algebra "formula".

For example, consider:

g(x)= 1 if x in Q, else 0

No regular formula or graph can easily depict this function, but it's a function anyway. Its domain is R, and the range is the set {0, 1}. We'll come back to this function later.

Triangle Inequality

We introduce the absolute value function |x| because of its sheer importance as a measure of distance:

$|x| = \{ x \text{ if } x \geq 0, \text{ else } -x$

We note |ab| = |a||b|, as well as the important triangle inequality:

$|a+b| <= |a| + |b|$

This breaks into two cases: the equal case and the greater than case.

If a or b are negative, this states that adding two positive numbers will always be greater than subtracting one of them and taking the distance. If a and b have the same sign (both positive, both negative), then the left and right are just equal.

To see why this is called the triangle inequality, we add a third point/number, c, between a and b, to create two line segments:

$|a-b| = |(a-c) + (c-b)|$

We can use this in the triangle inequality:

$|(a-c) + (c-b)| \leq |a-c| + |c-b|$

$|a-b| \leq |a-c| + |c-b|$

Meaning that the distance between a and b will always be less than or equal to: the distance between a and c, plus the distance between c and b.

As before, we see that separating our distance into two parts can only maintain or increase the total. There is no shorter path than the direct one between a and b.

If we put the points on a 2D plane, in a triangle, and keep our idea of distance, it becomes more clear why this is the “triangle inequality”: no path between two points is shorter than a straight line between them.

This inequality is very useful. We will be making use of it a lot for distance-related problems.

Math proofs are a practiced skill: you create a set of instructions to prove the truth of your statement to the reader, where each step logically builds on either an accepted fact or a previous step.

In the introduction, we used an indirect proof by contradiction, where you negate (take the opposite of) what you want to prove, and show that the negation is logically impossible.

A direct proof, on the other hand, takes an agreed truth, and expands on it in logical steps, until you have demonstrated the statement you would like to prove.

We demonstrate the structure of a proof with the following:

Theorem 1.2.6: Two real numbers a and b are equal if and only if for every real number $e > 0$ it follows that $|a - b| < e$.

We take a small aside to note how this is worded: "if and only if" is meant to show that a statement must be true both ways: if A, then B ($A \rightarrow B$) AND if B, then A ($B \rightarrow A$). This shorthand can be shortened further: "iff".

This could be rephrased as "a and b are equal iff the distance between them is smaller than any positive real number."

We first prove the forward statement directly:

If $a=b$, then $|a - b| < e$ (for every $e > 0$).

If $a=b$, we can say that $a-b = 0$.

$|0|=0$, and thus, $|a - b| < e$ is definitely true because $0 < e$ if e is a positive real number, no matter which number you choose.

Now we prove the backward statement:

If $|a - b| < e$, then $a=b$ (for every $e > 0$).

For the moment, we use a proof by contradiction. So, we assume that a does not equal b.

We want this to be true for "every" positive real number, so we want our statement to be true, no matter what e we choose.

If a does not equal b, then we can get the distance

$e_0 = |a-b|$

However, if $|a - b| < e$ for every e, then it should also be true for $e = e0$.

$e_0 = |a-b|$ and $e_0 < |a-b|$ cannot be true at the same time. Thus, we have demonstrated that "a does not equal b" cannot be true.

Thus, $a=b$, and we have proven our backward statement.

With both statements validated, our proof is complete.

In short:

-If they are equal, their difference is zero, and thus always smaller than any possible e

-If they were not equal, then they would have a nonzero distance, and this distance cannot both equal a number, while also being smaller than it (if e were set at that value)

Using the key words “for all/every” and “there exists” is very useful for analysis proofs, because of how it can manage an entire set of elements more easily.

Induction

Finally, we introduce induction, an argument in proof-writing that uses the natural “counting” numbers to extend a sequence of objects. This can be useful for defining these sequences, or proving things about them.

The main idea is that you can create the natural numbers if you:

1. Include the number 1
2. Add n+1 to the set if you have the number n

This ends up being a formal way to say “if you have the number 1 and can count up, you can define all the natural numbers.”

When building a sequence, this can be restated as, “if you have the first object in your sequence, and a rule for building the next object based on the last one, you have defined that entire sequence.”

For example, take the following:

$x_1 = 1$

$x_{n+1} = (1/2)*x_n + 1$

With these two statements, you have defined x_n for all n in N: the rule gives you all you need to continue. The entire sequence is included.

When proving a sequence, you can restate induction as, “if you can prove that this statement is true for the first object, and that it is true for the next object if it’s true for the current object, then you have proven it for the entire sequence.”

Let’s say we’re curious whether this sequence is always increasing, since the first few elements increase.

We will prove the following:

For all n in N: $x_n \leq x_{n+1}$

In this case, because we are comparing two elements, the "first object" is the relationship between x_1 and x_2.

$x_1 = 1$ $\quad\quad$ $x_2 = (1/2)*x_1 + 1 = 3/2$

So in the case of the first comparison, 3/2 > 1, and $x_2 > x_1$.

Now, we need to prove the n vs n+1 comparison, where each object is the comparison between two elements of the sequence: if $x_n \leq x_{n+1}$, then $x_{n+1} \leq x_{n+2}$.

-

This can either be solved by substituting out the starting expression:

$x_{n+1} = (1/2)\, x_n + 1 \quad \rightarrow \quad x_n = 2 * (x_{n+1} - 1)$ Solve for x_n...

$x_n \leq x_{n+1} \quad \rightarrow \quad 2(x_{n+1} - 1) \leq 2(x_{n+2} - 1)$ Substitute in x_n & n+1

$2(x_{n+1} - 1) \leq 2(x_{n+2} - 1) \quad \rightarrow \quad x_{n+1} \leq x_{n+2}$ Simplify algebra for your result!

Or by scaling up both sides of the inequality, so you can substitute in the higher equations:

$x_n \leq x_{n+1} \rightarrow x_n + 1 \leq x_{n+1} + 1 \rightarrow (1/2)*(x_n + 1) \leq (1/2)*(x_{n+1} + 1)$ Both sides modified

$x_{n+1} \leq x_{n+2}$ Substitute using rule

Using both proofs, we have successfully shown that if $x_n \leq x_{n+1}$, then $x_{n+1} \leq x_{n+2}$.

Combining the proof for the first object, and the proof for all later objects, we can safely say the claim

$x_n \leq x_{n+1}$ for all n

Has been proven by induction.

In short:
-We observe by calculation that the first pair of elements is increasing.
-We can show with algebra and substitution that if one pair of elements is increasing or the same, the next pair is.
-By induction, we can extend this to say that every pair of elements is increasing.

1.3 The Axiom of Completeness

What is R?

Up to now, we've mostly identified R as extending Q to fill in the "gaps" we found before.

Before we improve that definition, we'll admit that we're using an "intuitive" idea of real numbers. However, we'll not worry about that for now, so we can handle more reasonable problems.

R contains Q, and acts as the same kind of field: addition and additive inverses (subtraction/negatives), multiplication and multiplicative inverses (division/reciprocals), exist for every real number.

Zero does not have a multiplicative inverse, because division by zero is… unpleasant.

On top of that, addition and multiplication are: commutative, associative, and distributive.

All real numbers have an order resembling that of Q, as well.

Because they meet this huge laundry list, the R gets the honor of being named an ordered field.

Axiom Of Completeness

Finally, the property that separates R from Q: it has to somehow "fill in the gaps" that exist within Q. We resolve this problem using the Axiom Of Completeness.

Axiom of Completeness: Every nonempty set of real numbers that is bounded above has a least upper bound.

Bounds

First, we need to learn what this means, starting with some definitions.

Nonempty just means that this set is not the empty set. Important to state, but simple.

A set $S \subseteq R$ is bounded above if there is a number b in R that is greater than or equal to every element in S. Meaning, you can find a number "above" the entire set. This number b is called an upper bound for S.

There is a matching idea of bounded below, and a lower bound, for a number "below" the entire set (less than or equal to the entire set).

Now, what's the least upper bound?

As the name suggests, it is a type of upper bound, but it is the "smallest" one: the least. It's like talking about the "shortest tall person". This is formalized by saying:

For any upper bound b, the least upper bound s is less than or equal to it. It is at the "bottom" of the set of upper bounds, and there is only one: if two existed, they couldn't both be below each other, so they'd have to be equal.

To prove you are in possession of the least upper bound, you must meet both these requirements: "below" every other upper bound, and "above" every non-identical element of the set.

The least upper bound is often called the supremum, and the greatest lower bound (again, matching concept) is called the infinimum.

We will start with an example:

$A = \{1/n : n \in N\} = \{1, ½, ⅓, ...\}$

A is bounded above: 2 is larger than any number within this set, because the sequence decreases from 1. Additionally, it is bounded below: -1 is smaller than any number in this set, because they're all positive.

We want to find the least upper bound. We claim sup(A) = 1.

We need to show that 1 is an upper bound, and that no upper bounds are lesser.

To show that 1 is an upper bound, we note that 1 is the greatest value in the set, because every element after the first is smaller. So, no elements are greater than 1.

To prove it is the least upper bound, we note that 1 is an element of the set: if we had another upper bound, it would need to be greater or equal to every element in the set, including 1. If we focus on 1, this is exactly what defines the supremum.

We cannot yet prove it rigorously, but intuition does not betray us when it suggests that inf(A) = 0.

In short:
-1 is an upper bound because it is larger than every later element 1/n
-1 is the least upper bound because it is in the set: another bound would need to be greater or equal.
-We have met the criteria for the definition of sup(A).

An important takeaway: sup(A) and inf(A) may or may not be part of set A: in this example, inf(A) = 0, and we know 0 is not in our set: there is no natural number n such that 1/n=0; this is part of what defines a field (no multiplying inverse for zero)

We now take a moment to separate the maximum from the supremum.

A number a_0 is a maximum of the set A if a_0 is in A, and a_0 is greater than or equal to every element of A.

This can also be said as "a_0 is in A, and is an upper bound of A." Using similar logic to the above proof, it is also always the least upper bound, if it exists. There is not necessarily always a maximum for a set.

The difference is made clear by an open vs closed interval:

$(0,2) = \{ x \in R: 0 < x < 2\}$
$[0,2] = \{ x \in R: 0 \leq x \leq 2\}$

Both intervals have the same least upper bound: 2. However, the top set does NOT have a maximum: there is no "greatest" value inside the set (0,2).

Q versus R

We now see that we cannot be sure that a set will have a maximum, but the Axiom of Completeness guarantees that there will be a least upper bound. Because axioms are claims that we assume before working, it will not be proven: it is part of what defines the real numbers, not a derived fact.

What makes the Axiom of Completeness so important is that it does NOT apply to the rational numbers: a least upper bound is not guaranteed for a set with an upper bound. Let's demonstrate.

$S = \{r \in Q, r^2 < 2\}$

We can see that the set is bounded above: 2 is a fine upper bound, for example. However, if we search for a least upper bound in the rational numbers, we get closer and closer, with more and more precise fractions, but we will never reach a supremum.

We can prove it's not possible by comparing Q to R: the supremum is unique, and Q is a subset of R, so if the supremum exists in Q, we can also find it in R.

The supremum for this set in R is $\sqrt{2}$. There is no $\sqrt{2}$ in the rational numbers (as shown by our first proof), so there is no supremum for this set in Q.

This kind of example formalizes the "gaps" in Q, and the way R fills them: R extends Q to include any supremum that was missing from the rational numbers. All these suprema gaps are filled in.

The math necessary to fully process this example will be discussed in section 1.4.

This lemma provides an alternative way to state the requirement for a least upper bound, this time assuming you have an upper bound.

Lemma 1.3.7: Alternate statement of the least upper bound

Lemma 1.3.7. Assume s (in R) is an upper bound for the set $A \subseteq R$. Then, $s = \sup(A)$ if and only if, for every choice of $e > 0$, there exists an element a (in A) satisfying : $s-e < a$.

No matter how small a positive number you subtract from your supremum s, it will be smaller than some value in A.

In the simplest form: any number smaller than the supremum is no longer an upper bound.

We first prove the forward statement:

If $s=\sup(A)$, then for some a in A, $s-e < a$ (for every $e > 0$).

If s is the supremum, then by its definition, any smaller value is not an upper bound.

$s-e$ represents any smaller value: for it not be an upper bound, then it must be less than some element a in A.

This is the result we are looking for: $s-e < a$, because otherwise, $s-e$ would be an upper bound.

Then, we prove the backward statement:

We have an upper bound s.
If, for every $e > 0$, and for some a in A, $s-e<a$, then $s=\sup(A)$.

$s-e$ for any $e > 0$, can be stated as "any number less than s".

Thus, if $s-e < a$, that means "any number less than s is less than a, and thus not an upper bound."

If any number less than s is not an upper bound, then any upper bound b must be greater than or equal to s. Given that s is an upper bound, and $b \geq s$, we have both requirements for a least upper bound.

In short:
-If $s=\sup(A)$, then it is the smallest upper bound: every smaller number $s-e$ is not an upper bound, and thus less than some a.
-If $s-e<a$, and s is an upper bound, then no numbers smaller than s can be an upper bound, and so all upper bounds are greater than or equal to s. This defines s as supremum.

Similar logic applies for greatest lower bounds in all cases. The Axiom Of Completeness is equivalently stated with these lower bounds.

1.4 Consequences of Completeness

The Axiom of Completeness allows us to find a more "natural" way to say that the real numbers have no gaps.

We show this with the following theorem, where we "zoom in" on the real numbers.

Theorem 1.4.1 (Nested Interval Property): For each n in N, assume we are given a closed interval $I_n = [a_n, b_n] = \{x \in R: a_n \leq x \leq b_n\}$. Assume also that each I_n contains I_{n+1}. Then, the resulting nested sequence of closed intervals

$I_1 \supseteq I_2 \supseteq I_3 \supseteq \ldots$

Has a nonempty intersection between every set.

To prove this theorem, we only need to provide one real number x that exists in I_n for every n in N. If there is at least one element, the intersection is nonempty.

We want to use the AoC to prove this, so we will need a set with a supremum or infimum; this bound will be able to follow the nesting intervals. We can't use any particular I_n, because it won't reflect n+1, n+2, and so on.

Instead, we are looking for a sequence that is bounded on one side, and captures the intersection. In other words, we need a sequence that either does not increase or does not decrease endlessly (without bound), and includes every I_n.

Let's "include every I_n" first, so we don't get distracted. How do we do that? If we were to describe each interval, we would describe them... using their endpoints.

One might notice that the later intervals must necessarily shrink, and so both of the endpoints move slower and closer together.

Let's look closer at these endpoints.

We want a bounded set. Because each interval is a subset of the previous one, the current left endpoint a_n is bounded below by the last interval's left endpoint a_{n-1}, and bounded above by the current right endpoint b_n.

What's more, all future left endpoints have both of those bounds. The lower bound will definitely be outside of the intersection, however, since a will keep moving up. So, we instead search for the least upper bound for the left endpoints.

We define the set of left endpoints as $A = \{ a_n : n \text{ in } N\}$.

We want to use the supremum of this set. AoC guarantees that if a set is bounded above, it has a least upper bound.

Our nested intervals mean that any given b_n is greater than every element in A: a_m will never be larger than b_n, because otherwise I_{m+1} wouldn't be a subset of I_m; it would cross over. Thus, any b_n can serve as an upper bound.

This proves that sup(A) exists. sup(A) is an upper bound, so sup(A) ≥ a_n for every n. We've proven it's always above any other a.

However, because every b_n is an upper bound, and sup(A) is the least of those upper bounds, sup(A) ≤ b_n for every n. We've proven it's always below b.

With this, we can say sup(A) = x. Why? Because a_n ≤ sup(A) ≤ b_n for every n, it is contained within every single interval (a_n , b_n).

We have proven that sup(A) is in every interval, and thus that the intersection of every interval is nonempty.

In short:
-We focus on the set A of left endpoints for each interval
-This set is bounded above because any b_n is greater than any element of A.
-According to AoC, this means that sup(A) exists.
-sup(A) ≤ every a_n because it's an upper bound
-sup(A) ≥ every b_n because it's the smallest upper bound
-This means sup(A) is inside of every interval [a_n , b_n], and the intersection is nonempty!

In shorter:
-AoC proves that the set of left endpoints has a supremum because b_n is above A
-sup(A) is above all a_n(upper bound), below all b_n(least upper bound)
-This means it's in every interval, so the intersection is nonempty.

Layman:

AoC says any group of numbers that has an upper limit, has a number perfectly on that edge. The left endpoints have a number like that on their upper edge, because the right endpoints create a wall the left can't cross. This edge will always be on/between the left endpoint and the right endpoint, so that interval will never be empty.

This shows what makes R different from Q: no matter where you "zoom in", you never reach an empty spot.

Now, we look at the relative density of points in N, Q, and R. Let's start with a useful property of N.

Theorem 1.4.2 (Archimedean Property)
(i) Given any number x in R, there exists n in N satisfying n>x.
(ii) Given any number y>0 in R, there exists an n in N satisfying 1/n < y.

This essentially states "real numbers do not go higher than natural numbers, and they do not get closer to zero than the reciprocal of the natural numbers."

(i) This statement can be reframed as "there is always a larger n", or more properly, "N is not bounded above".

> We'll use a proof by contradiction, since we intuitively "know" that N is not bounded above, and plus, we want to play with AoC more. A direct proof would focus on the construction of N (spoiler, it's messy).
>
> One might find this question almost silly, and simply say "if there were a largest n, then all you have to do is add n+1 to get a bigger number."
>
> There is some merit to this logic, but in general, we want to be more careful than that: we haven't proven that N would even have a maximum ("largest n") if it were bounded above. Future situations may not be so forgiving.
>
> So instead, we use what we CAN prove with AoC: if N is bounded above, then it does have a supremum s. This s is as close as we'll get to a maximum.
>
> We don't know if s is in N, but we can find the closest other n: we know any number smaller than s is not an upper bound, and thus has an n above it.
>
> Because we're working with natural numbers, we'll use the smallest increment: a−1 is not an upper bound, and thus, there exists an n > a−1. Further, a ≥ n > a−1, so we know the distance between a and n is less than 1. Interesting.
>
> Now that we have a "close" n to a (|n−a|<1), we can cross that distance, and apply the n+1 logic we were so eager for: if n > a−1, then n+1 > a.
>
> Now, we have a natural number n+1 larger than our supposed upper bound, because N is closed. This shouldn't be possible, so we have proven that N is unbounded.

(ii) $1/n < y$. Informally, this says that there is always a smaller natural number n: there is no minimum.

> Immediately, we know that $1/n$ is no fun, because it's definitely not a natural number. We can solve this problem with a reciprocal: $1/(1/n) = n$. Now we're back to natural numbers.
>
> So, we have $n > 1/y$. We've shifted the problem over to the real side, but the real numbers are much more inclusive: $1/y$ is still a real number x.
>
> We now have $n > x$ for some n in N. We've shown that (i) and (ii) are equivalent.
>
> (Optional)
> Alternatively, you can imagine that if you measured "size" of a number by $|s|$, you might measure the "smallness" of a number as $|1/s|$. With this interpretation, you get an intuitive feel that "the 'smallness' of $1/n$ is unbounded". The same logic follows.

In short:
-If N were bounded, there would be an n right underneath sup(N), to which you could add 1 and be above the supremum. This is nonsense, so N is unbounded and there exists n such that $n>x$.
-If you take the reciprocal and realize that $1/y$ is another real number x, (ii) is just a restatement of (i).

Theorem 1.4.3 (Density of Q in R): For every two real numbers a and b with $a<b$, there exists a rational number q such that $a < q < b$.

To simplify, this says that there's a rational number between any two real numbers.

We will start with the simplified case where $0 \le a$: a is non-negative. The case where $a < 0$ follows simply from this result. If not, we could solve it after anyway.

We want to find a q between a and b. Because we're focused on the structure of Q and R directly, we will use a direct proof.

First, our target. q, like all rational numbers, can be expressed as the ratio m/n (m and n in Z. Because the result will be positive, we can assume m and n are in N: m>0 and n>0).

We want to approach and enter the interval (a,b) in a way we can control, so we know what's happening. Thus, let's examine the structure of the number we're trying to fit into this interval: what is a rational number? How do we control it?

The spirit of rational numbers is in division. m/n is often interpreted as "how many pieces of n size can 'fit' into m". At first, this seems unhelpful, but breaking m into "pieces" might be promising, because we can control those pieces.

This might suggest we think of rational numbers in pieces: how do we break m/n into pieces we can manage? How would you break up $\frac{2}{3}$ into parts? Well, you could be creative about it, or you could just say "wait, isn't that $\frac{1}{3}$ + $\frac{1}{3}$?" You're breaking m=2 into 1+1. That might work.

Thus, we have a suggestion: what if we break up m/n into a sum of 1/n, m times? Essentially, we've traded our last question into "how many pieces of 1/n size can 'fit' into m/n". The answer is clearly m, but that means we can play with both m and 1/n separately. We'll take 1/n-size steps.

This train of logic for breaking up m/n may feel obvious once you've done it before, but it's important to remember and use.

We want to make sure the last step ends up inside the interval. According to the archimedean property, we can make m as large as we want, so there's no worry about not being able to "reach" a high enough value (m > na, m/n > a).

However, m and n are in N. Numbers in N have a fixed space between them, so some distances are impossible. It's possible that, with the wrong choice of n, our steps will be too large, and we'll go too high, over the interval (m/n<a, too low, (m+1)/n>b, too high). There would be no possible m that would put m/n in the right range.

So, we need each step to be smaller than the distance between a and b: if the last step is smaller than |b−a|, then that step can't reach b, even if it starts at a.

Thus, b−a > 1/n. We can treat b−a as a single real number because R is closed, so we have a statement of the form x > 1/n. The second part of the archimedean property states that there is definitely an n that satisfies this, so it's possible. Thus, we have our value of n.

We begin looking for m. If 1/n is sufficiently small, it's possible for there to be multiple possible choices of m, if say, 2 or 3 steps past a weren't large enough to cross the interval to reach b. For simplicity, and so we don't accidentally overshoot, we will choose the smallest possible m: meaning that m−1 is too small, such that:

$m-1 \leq na < m$

As we said above, there's definitely a large enough m, so our upper bound is secure. But how do we show that this particular m is below b?

What we know about b: $b - a > 1/n$. Because we're controlling our m and $m-1$ based both a and $1/n$, we'll use these two to describe b: $b > a + 1/n$

We want m to be lower than b, so we'll use the $m-1 <= na$ inequality, putting m on the low end. $m < na + 1$ will let us talk about m directly.

We want to combine these to show our upper bound: $m/n<b$, or $m<bn$. So, let's modify the b equation to match(either equation works):

$b > a + 1/n \rightarrow bn > na + 1$

Conveniently, we don't need to connect these with a clever third inequality (though, we might someday):

$bn > na + 1 > m \rightarrow bn > m$.

We've now proven our lower bound: with both bounds, we have proven that there is a possible m and n, where m/n is above a and below b.

This theorem can be said as "Q is dense in R". There's no place in R where you won't find a rational number. Dense indeed.

You can also prove that irrational numbers are dense in R, but we will save this for a personal exercise.

The main motivation of R was to include the numbers that Q couldn't, like square roots. We used the following set to show this comparison:

$T = \{t \in R: t^2<2\}$

We claimed that $\sqrt{2}$ was the supremum of this set in R, and said that since $\sqrt{2}$ wasn't in Q, and the supremum is unique, then T had no supremum in Q.

While we did show $\sqrt{2}$ wasn't in Q, we didn't actually prove that it was in R. Let's do that now.

Theorem 1.4.5 (Square root of 2 in R): There exists a real number a in R, satisfying $a^2=2$.

Our relevant tool, the AoC, relies on the supremum, so we'll need a bounded set. We know the supremum exists in R, so if we want to prove $\sqrt{2}$ exists, we just need to prove that $\sqrt{2}$ is a supremum to place it within R.

The above set, by no coincidence, meets these requirements.

$T = \{t \in R: t^2<2\}$

So, we need to prove that our supremum $a = \sqrt{2}$.

As always, because we have a least upper bound, we define it by the range above(least), and below it(upper bound).

This problem gives us a chance to talk about a very related concept to the supremum: proving that two things are equal.

We want to prove that a equals $\sqrt{2}$. In this case, we'll do it by trying to cut out the ranges above and below: we'll show that $a^2>2$ is untrue, and $a^2<2$ is untrue, leaving the only possibility that $a^2=2$.

Of course, this logic applies to any ordered set, not just the world of real numbers. In some ways, the AoC is motivated by this reasoning.

While these two ideas, the (AoC and proving equivalence) are very closely linked in this problem, we want to split hairs and remember that they are not the same. Even when you can't use the AoC, you can still prove equality this way, as long as order exists.

Leaving that aside, let's prod the idea that $a^2<2$:

To clear out the range below, as normal, we'll focus on the "upper bound" part of the supremum. If $a^2<2$, can we have an upper bound?

If a were an upper bound, any larger number $a+e$ should also be an upper bound. So, let's try to find a number $a+e$ that isn't an upper bound. After all, there's some room between a and 2 we might be able to squeeze a+e in.

Let's check our tools. We could try to use denseness to find a q between a and 2, but we don't know if q is the square of a real number. Dead end, but it does support the idea that there's some space to search in.

So instead, let's start with basics: we want to control e, and it needs to be able to be very small, to squeeze into any space. Needs to be small… The archimedean principle (AP) guarantees that 1/n can always be small enough for any purpose!

So let's say e = 1/n, where $n \in N$. n is pretty manageable, and the fact that 1/n is a rational number is a nice bonus. And as said before, we know it fits our size needs.

Now let's try to find an a+1/n that isn't an upper bound: one that is inside the set. So, we'll square it, to match t^2 terms.

$(a+1/n)^2 = a^2 + (2a/n) + 1/n^2$

Those side terms are ugly, so for the moment we'll say $b = (2a/n) + 1/n^2$, so don't get lost in algebra. We'll unpack it when we need to.

$a^2 + b$

We know that $a^2<2$, we have a bit of room for b to fit in, so let's try that:

$a^2+b < 2 \rightarrow b<2-a^2$

Now, we need to somehow prove that b is less than some real number… looking through our tools, we want something good for small numbers… we could use the AP if we have a rational number we can turn into 1/n.

$b = 2a(1/n) + 1/n^2$

If we can isolate 1/n, we could move the rest over to the $2-a^2$ side, and use the AP. Let's try that.

$b = (1/n) * (2a + 1/n)$

No dice. We can't seem to wrangle these two apart… we got only partway. That $1/n^2$ is a pain, maybe there's a way to get rid of it…

If we can't do algebra on the left, and the right is already separated from the rationals, all we have left… is the inequality itself.

$b < 2 - a^2$

How do we typically use inequalities? At the start of this problem, we did e = 1/n, because 1/n can always be less than or equal to another real number e: $1/n \leq e < 2-a$, in this case. Chaining together inequalities can let you replace something tricky, with something more manageable, that can still prove your case.

Let's give that a try: We'll create a number k. If k is below both of the other terms, then that doesn't help our case: b could be above or below $2-a^2$. However, if k is between the two terms, then you can use transitive logic: if $b<k$ and $k<2-a^2$, then $b<2-a^2$. So, we want $b<k<2-a^2$.

$b = (1/n) * (2a + 1/n)$

We want to use this trick to isolate 1/n, while making k larger than b.

We mentioned that $1/n^2$ is a pain to deal with. We could replace it with something bigger, that keeps the 1/n factor… that's it, we'll just use 1/n.

$k = 2a(1/n) + 1/n \quad \rightarrow \quad k = (1/n) * (2a + 1)$

We know that $n^2 \geq n$, so $1/n^2 \leq 1/n$. Thus, we know $b \leq k$, and we've isolated 1/n.

Note that this is a common trick, especially in physics problems: if $1/n^m$ is a problem, you can turn it into something bigger (1/n) or smaller (0), depending on the situation.

Now, want to find n so that

$(1/n) * (2a+1) < 2-a^2 \quad \rightarrow \quad 1/n < (2-a^2)/(2a+1)$

Ignoring all the fancy algebra on the right, this just says "1/n is less than some real number". This is exactly what we're looking for: AP supports our claim! Reversing our steps, now that we've proven this is true:

$k = (2a+1)/n < (2-a^2) \rightarrow \quad k < (2-a^2) \rightarrow \quad b < k < (2-a^2) \rightarrow \quad b < 2-a^2 \rightarrow a^2 + b < 2 \rightarrow$
$(a+1/n)^2 < 2$

Thus, we have shown that there's a large enough n so that a+1/n is not an upper bound. Thus, if a + 1/n (>a) is not an upper bound, then a is not an upper bound. a cannot be the supremum if $a^2<2$.

We must also prove $a^2>2$ is illogical.

This can be done by following the same process with $\sqrt{2} < a - 1/n$, and approximating $1/n^2$ as 0. In this case, we show that any $a > \sqrt{2}$ cannot possibly be the supremum, because a−1/n can be a smaller upper bound.

We have proven this statement both ways. Thus, $a = \sqrt{2}$, and $\sqrt{2}$ exists in R.

By substituting 2 for any $x \geq 0$, we can show any radical $\sqrt{x}$ is in R. Furthermore, by using $(a+1/n)^m$, we can show $\sqrt[m]{x}$ exists for any $m \in N$. Nice.

Countable and Uncountable Sets

Up to this point, we've used the AoC to confirm things we expected from R, and to show that the AoC gives us a system where these things are provably true.

However, the following concept is another thing entirely, being an entirely new notion, and a strange one at that.

Right now, we have our mental picture of R: densely packed with rational and irrational numbers. For both Q and I (irrationals), you can find them in any real interval: this is what makes them dense in R.

However, one might wonder: at first, Q and I seem to be mixed in equal parts, but we've not proven this in any way. Is it really true, that they're the same size? How would we prove such a strange thing?

When we describe the size of a set, we are talking about its **cardinality**. The cardinality, for finite(not-infinite) sets, is just the total number of elements.

The set with the fewer elements is smaller: the set of Snow White's dwarves (7) is less than the set of months in a year (12) because 7<12. However, the set of dwarves (7) is the same size as the set of colors commonly listed in the rainbow, (7) because 7=7.

However, how do we talk about cardinality for infinite sets? They don't have a size you can assign an ordinary number to. So, instead, we'll compare sets to each other.

If you are given two bags (sets?) of marbles, how do you know which bag has more? You could count all of them, but that's a hassle and easy to mess up. Instead, you could pull one out of each bag at a time to pair up, until one bag is empty. The bag that is empty has fewer.

This same logic can apply to any two sets: if you can match every element in both sets, they're equal. If not, the one that has leftover elements is larger. It not only had enough elements to equal the other, but it had some to spare.

With that in mind though, how do we pair up two infinite sets? We can't exactly do that by hand. Instead, we need a system to pair up each element of each set for us, according to some rule.

If we go back through all of our tools, we know that functions are used to pair elements from two sets. Perfect. So, we'll try creating a function that does what we want: pairs every element of the two sets, and leaves none left.

This function will be notated as $f: A \rightarrow B$

To review what this means: f takes every element a of A and pairs it with an element b of B. Functions are also called **mappings,** and the word "to **map**" is the same as this act of pairing.

What are the crucial parts of our goal? First: we want every element of A and B to have a match.

Since f already uses every element of A, it needs to use every element of B. Basically, every b needs to have a partner in A. This means that f is **onto**, because f maps some a **onto** every element of B.

Onto can be formally defined as: given any $b \in B$, there is some a such that $f(a) = b$.

In short, "for every element in B, there is some a it is paired with by the function". Now that all of A and all of B have partners, we have met that requirement.

Second: if elements get "paired off", that means they can't be re-used for another pair. So, no a_1 and a_2 can pair with the same b. This is called **one-to-one (1-1)** because each element only has one partner.

1-1 can be formally defined as: if $a_1 \neq a_2$, then $f(a_1) \neq f(a_2)$

This basically says, "if these two elements are not the same, then they cannot get the same output from the function". Since these outputs are from B, then it makes sense: we don't want them to get the same partner in B.

This combination of 1-1 and onto gives us exactly what we're looking for, to replicate our marble analogy.

Below, let's formally write the following: if two sets can have a direct pairing of elements in a way that is "1-1 and onto", they're the same size.

Definition 1.4.7: Two sets A and B have the same cardinality if there exists an f: A → B that is 1-1 and onto. This can be notated as A ~ B.

Cardinality of E

Let's test out an example: let's compare E (natural even numbers) to N.

We will compare them by trying to create a function f : N → E that is both onto and 1-1. If we can't do it, then one of the sets is smaller.

Can we produce every even number (onto) using some unique (1-1) natural number? Let's see. Even numbers include every other natural number starting from 2, meaning they're all multiples of 2. In fact, the set of all even numbers includes every multiple of 2. So… to be a bit redundant, couldn't we just multiply n by 2?

1*2 = 2, 2*2 = 4… And so on. If n is a natural number, f(n) = 2n. Let's check that this meets our requirements.

Are they unique inputs? Well, if $f(n_1) = f(n_2)$, then $2n_1=2n_2$, and $n_1=n_2$. Contradiction. If they are the same number when multiplied by two, dividing shows they were the same to start with. There's no way to violate the 1-1 property.

Is all of E covered? For any number e ∈ E, we know it's a natural number and a multiple of two, which means e/2 is some other natural number, which we can find in N. Every e can be gotten from some n times two. We can be confident we have an onto function.

N ~ E.

As we used above, E is a (proper) subset of N (E ⊂ N). This may make it seem like it should be "smaller", but according to cardinality, they're the same size.

After all, how can N be bigger if there's enough even numbers to match every natural one? What a strange result.

Cardinality of Z

Another example, to explore our familiar number systems further. This time, how does Z (integers) compare to N?

$N \subset Z$, but we've since learned that this doesn't necessarily mean that Z is "bigger."
Another function f: $N \rightarrow Z$ in this case. We need it to be 1-1 and onto.

We could get all the positive numbers directly with f(n) = n, but that would miss all the negatives, and zero too. We'll need negative signs. There's no arbitrary point at which we can switch to negative numbers, either, because we never run out of positives. So, we should do them both at the same time.

How do we organize them, then? They need some kind of order to fit into n. Well, to keep it tidy, every positive number has a negative counterpart. We could do them together: {1, -1, 2, -2…}

But then, we need to put zero somewhere. Since we're going with increasing magnitude, we'll put it at the beginning. We've got { 0, 1, -1, 2, -2…}, and this should cover every integer.

Let's try to create a rule using algebra for this, to make it easier to work with. The positive and negative numbers are broken up by even and odds, so we might be able to re-use our work for $N \rightarrow E$.

Because 0 is neither positive nor negative, we'll ignore it for now. The even numbers match with the positive numbers. This is like our last function, but this time, we're going $E \rightarrow N$. So, instead of multiplying by two, we divide by two: if n is even, f(n) = n/2. Done.

The odd numbers seem harder at first, but we can actually just turn them into even numbers by doing n-1 or n+1. Which do we pick? Well, both work, but if we want to include f(1) = 0, we should do n-1, so that 1-1=0.

Now that we have even numbers n-1, we'll do the same thing as before: divide by two. This time we're going $E \rightarrow -N$ (negative of the natural numbers), so we need a sign flip as well. Thus: if n is odd, then f(n) = - (n-1) / 2.

Together, we have:

f(n) = -(n-1)/2 if n is odd.
f(n) = n/2 if n is even.

Above, we made sure that it covered all of the integers (N, -N, and zero), so we know that f is onto. Further, because the magnitude is always increasing, we know that it is 1-1: each positive and negative number of the same magnitude appear together, so on either side, magnitude is lower or higher.

Thus, we've shown that N ˜ Z. Again, one set is a proper subset of another, and yet they are the same size.

We may start to wonder if all infinite sets are the same size: is any infinite set S going to have N~S be true?

Well, let's keep exploring.

Cardinality of Q

How does N compare to Q?

We want to try to map all of N to Q according to some rule that is 1-1 and onto.

However, unlike our previous examples, Q has no space between numbers: that means we can't just go by increasing magnitude or something similar, because there is no "minimum" distance to count using.

This means that applying a simple formula using N (which has space) may be messy, if at all possible. Without a minimum distance, the gaps in N are likely to leave gaps across Q. Are we stuck?

Well, as we noted very early on, functions do not need to be algebra formulas. We can use other kinds of rules to take every n ∈ N and map it to some r ∈ Q. There's still hope. We just need to figure out what that rule looks like, and describe it.

But right now, we need some hints on how to get started. Let's look at the structure of Q for an idea as to how we can make a rule to fill it out. The rule that builds Q may give us an idea as to how we can break it into parts we can control more easily.

All of these numbers in Q will be of the form p/q, where p and q are both integers and q is not zero. In short:

Q = { p/q, such that p, q ∈ Z and q ≠ 0 }

"Such that" just means "here's what we need for this statement". In this case, we'll shorten it even further to just a colon (:).

Q = { p/q, with the requirement that p, q ∈ Z and q ≠ 0 }

Q = { p/q : p, q ∈ Z and q ≠ 0 }

p and q are part of Z, which is easier to work with for our natural numbers than Q, given the space between numbers: this gives us a starting place and avoids missing gaps. Our work showing that we can match N with Z might help us match N with these integer pairs (p,q).

But we have two problems:

##

One, if we blindly create every combination of m and n with some rule, we'll end up with duplicates: 1/1 and 2/2, for example. This means any function involved would not be 1-1: multiple inputs($n \in N$) would have the same output ($r \in Q$).

This problem can be avoided if we only use the fractions that are completely simplified. I.e., 2/4 is not allowed because 2 and 4 have a common factor, and could be stated as 1/2. There's only one way to have most completely simplified fractions.

Oh, but there's also negative numbers: how do we handle -5/8 vs 5/-8?

Well, we only need p OR q to be able to go negative in order to cover negative fractions. q already can't be 0, so we'll restrict it only to the natural numbers. Now 5/-8 isn't possible, only -5/8 is left.

$Q_{1\text{-}1} = \{ p/q : \quad p \in Z, q \in N \quad p/q \text{ is in simplest form} \}$

Simplest form is defined as we said above: p and q have no common factors greater than 1. This should avoid any 1-1 problems. (Assume 0 can have every factor, to remove 0/2, 0/3…)

###

Two, we need to figure out how we progress through Q. How do we start our function off, for example? And how do we take steps from n to n+1?

Luckily, like we said, Z and N are easier to work with than Q, because of the gaps. We can't get a minimum step size for Q, but we can get a minimum step size for p and q.

We start off with 0/1. Minimum p, minimum q. Next up is 1/1 and -1/1, since $N \sim Z$ taught us that pairing positives and negatives is helpful for ordering. And so, we've got our starting place.

Now we've got a guess for how we'll proceed through p and q: we want to increase the magnitudes of both. How do we "count" through p and q at the same time?

Well, if we increased them simultaneously (p=n, q=n), obviously we'd only get n/n=1. So it can't be exactly the same time. There are a lot of different options you might consider:

>Set a limit for p and q at the same time, get all combinations ($|p| \leq k$, $q \leq k$), increase limit
>Get all $|p/q| \leq 1$ while increasing q over and over, take the reciprocal of all to get $|p/q| \geq 1$
>Increase p, get fractions for various q. Increase q, get fractions for various p…
All of these approaches might work, as long as we can get an order that makes f 1-1 and onto.
However, they may not all be equally easy to go through.

Let's try to simplify the rational numbers: we will ignore negative fractions, because we can just take $\pm p/q$. 0 is just zero, too. Our task now: get every positive rational number, by getting every combination of p and q (that is fully simplified, but we can filter that after.)

Like we did for the idea of ordinality, let's think in finite terms first, and see which ones stick once we go infinite. Take a small starting example: all the rational numbers, where p and q are 1 to 3.

$\{ 1/1, 2/1, 1/2, 2/2, 1/3, 3/1, 2/3, 3/2, 3/3 \} = \{ p/q: \quad p,q \in \mathbb{N}, \quad p, q \leq 3 \}$

How can we organize/order these items? Well, we can organize them by p, organize by q, or… some combination (p+q, p-q, pq, magnitude of p/q).

Let's try that first one:

p = 1: { 1/1, 1/2, 1/3 }
p = 2: {2/1, 2/2, 2/3}
p = 3: {3/1, 3/2, 3/3}

Unfortunately, breaking it up this way doesn't work once we go infinite. Each p will be infinitely long, and so we'll never reach the end of p=1 and switch to p=2.

We could keep trying other ways, or trying to set limits so this method works (and, just so you know, it would). However, that would be messy work we don't need. Instead, we might notice this looks like a grid. Let's make it into one:

p/q	1	2	3	n
1	1/1	1/2	1/3	1/n
2	2/1	2/2	2/3	2/n
3	3/1	3/2	3/3	3/n
m	m/1	m/2	m/3	4/n

We want to cover this entire infinite grid. This way of looking at it is more visual, and might show us more directly how to do so. You might come up with several ways. Below, we show one, for fun:

p/q	1	2	3	n
1	1 → 1st ring	2 ↓ Second ring	9→Third ring	10 ↓ nth ring
2	4 ↓	3←	8 ↑	11 ↓
3	5→	6→	7 ↑	12 ↓
n	16 ↓	15←	14←	13←

Not only does this pattern provide a visual way to prove that every number is covered, it even shows how that order matches $n \rightarrow Q^+$ (positive real numbers).

Is it 1-1? Skipping non-simplified numbers prevents one kind of duplication. However, on top of that, no two cells have the same row and column, so the same p/q will never appear twice. Yes, it is 1-1!

We also know it's onto because we can find every single positive real number. For a number p/q, the larger of p or q tells you which ring it's on. For example, 22/7 is on the 22nd ring, and has an appropriate natural number assigned.

Now, we just need to switch from Q^+ to Q. We can add 0 before 1/1 (think of it as the 0^{th} ring, maybe?). On top of that, we can just pair up positive real numbers with their negative counterparts, like for example: {1/1, -1/1}.

If we wanted to write this in set notation, we could just say, for the k^{th} ring, it contains the set of rational numbers A_k:

$A_k = \{ \pm p/q: \quad p \in N_0, q \in N, \quad p/q \text{ is in simplest form}, \ p = k \text{ OR } q = k\}$

Where N_0 is $\{0\} \cup N$ (the counting numbers with 0 at the front)

And the order is given by the diagram above. We could write it formally, but that would be ugly, so we'll pass. Using N_0 for p lets us say $A_0 = \{0\}$ without too much hassle.

Thus, Q is ordered by:

$Q_{new} = A_0 \cup A_1 \cup A_2 \cup A_3 \cup \ldots$

Now, we have a function f: N → Q. Is it 1-1 and onto? We already showed it is! For onto: using the larger of p/q will give the proper A_k, now including negatives. 1-1 is a combination of simplified forms, and the unique row/column for each cell means p and q will never repeat. So, we've shown N ~ Q. Great!

A brief aside: the grid above makes it evident that there are many ways to break Q up into finite pieces/paths that do what we want. For example: a common choice for this proof is

$A_k = \{ \pm p/q: \ p \in N_0, q \in N, \ p/q \text{ in simplest form}, \ p+q = n \ \}$

If you draw these groups on the grid, you'll see that it makes diagonals in the / shape. This could also be used for the proof.

As we said before, there are other, non-visual ways to come to these solutions (several were tested before the above was chosen), so not coming across the intuition-friendly grid doesn't prevent you from thinking through how to proceed.

If it provides as a help to future problem-solving: every approach relied on some measure of gradually increasing the magnitude of p and q, and then putting restrictions on the order so that it is 1-1 and onto.

This was so that we could start with the "simplest" p and q, and move in an easy-to-understand way.

Cardinality of R

Finally, we attempt to show that N ~ R.

However, this is tricky: R ⊃ Q, but basically adds the idea that there's no "holes". Those holes are filled by the AoC, and you can confirm they're filled using the nested interval property (no infinitely shrinking/nested interval is empty).

These holes are represented by I, the irrationals. Meaning, R = Q ∪ I. This is pretty vague, though: "the real numbers includes the rational numbers and uh, the not-rational numbers". Not very helpful.

At this point, all that structures R are those theorems we've got: AoC and the NIP. The AoC requires a bounded set, so R, Q, and I are all out. The NIP requires nested intervals, so we need boundaries for that too. We'll need to create some kind of bounding either way.

We know all of the rational numbers, so let's start looking for irrationals. Thus, whatever set we create, needs to focus on what the rational are not. We could try to use the AoC, but we don't have any inspiration for systematically getting subsets.

##

Let's try the NIP instead: We start with some interval, defined by two rational numbers. We'll be zooming in with an infinite number of intervals, so Q's infinite size might help rather than hurt.

How will we shrink our window with Q? Well… we want to find the not-rationals, so we can start filtering those out.

If we have the rationals Q={ r_1, r_2, r_3…}, we could take our interval A_0 and create a nested interval A_1 where r_1 is outside that interval. This is always possible, because if we break A_0 into two disjoint (separate) intervals, only one of them can have r_1. We've filtered out one rational.

We can repeatedly do this for every rational number, and the NIP guarantees that the overlap of all these sets is non-empty: this infinite intersection definitely has some real number y_1 in it, and it's certainly not rational. We've found our first irrational number!

What now? Well, we want to find a different irrational number. Meaning, the set of numbers we want to avoid, now includes r_1. So, let's filter that out too! We'll choose A_1 so that it does not contain y_1, and then we'll do what we did before for every rational number.

Now, in the infinite intersection, we have an irrational number, and it can't be our first irrational. So, we must have a unique irrational y_2. Now, we avoid the elements in $Y_2 = \{y_1, y_2\}$ as well as Q. We repeat this over and over, and the result is an infinite set $Y \subseteq I$.

###

We might wonder at how powerful the NIP is: because this set is never empty, it can always automatically find a new irrational for us.

Wait. Always?

Doesn't that mean that, no matter what set we have, we'll be able to find a new irrational that was missing?

Suddenly, we realize: there's no way to create an ordered set that contains every real number this way. Even worse, we've shown that it isn't possible anyway: let's say we had the set of numbers $R = \{x_1, x_2, x_3, \ldots\}$, then shouldn't we be able to use the NIP to find a new real number?

If there are always numbers outside of our set, then we can't create a function from N to R. N and R are NOT the same cardinality!

That's even more bizarre an idea, after everything so far. N, Z, E, and Q are the same size, but R is somehow a "bigger infinity". What does that even mean?

N is the set of "counting numbers". For this reason, its called a **countable set**. If you have a larger set, then you can no longer count your elements: you have an **uncountable set**.

Now that we know that there is such a thing as "uncountable", we might want to look more closely at the properties of the countable set:

The subset of a countable set must either be empty, finite or countable. If you remove some elements, it can't possibly become larger. And a "smaller" infinite set would still be long enough to map onto every natural number. This means countable sets are the smallest infinite set.

Two countable sets combined will produce another countable set: you can show this by doing what we did to get cardinality of Z:

We just need to alternate between the two sets, and every element is still ordered and counted, just with "double" length: $\{1, 2, 3\ldots\} \cup \{-1, -2, -3\} \rightarrow \{1, -1, 2, -2, 3, -3\ldots\}$ Of course, doubling the length of an infinite set doesn't change its size, as long as you can still map it to N.

This fact teaches us strange something about I: if $R = Q \cup I$, and Q is countable, then I must be uncountable. Otherwise, two countable sets would be combining into an uncountable. This means that I, despite being less commonly used, is uncountably bigger than Q.

To keep these notes tidy:

Theorem 1.4.12. If $A \subseteq B$ and B is countable, then A is either countable, finite, or empty.

Theorem 1.4.13. (i) If $A_1, A_2, \ldots A_m$ are each countable sets, then the union $A_1 \cup A_2 \cup \cdots \cup A_m$ is countable.
(ii) If An is a countable set for each $n \in N$, then the union of every set A_n is countable.

1.5 Cantor's Theorem

We present Cantor's own, original proof that first proved that real numbers were uncountable. To make it more manageable, we use a bounded set to start:

Theorem 1.5.1. The open interval (0, 1) = {x ∈ R : 0 < x < 1} is uncountable.

We'll show that this theorem is equivalent to the one we used before, by proving that: (0,1) is uncountable iff R is uncountable.

First though, we need to prove the theorem itself. We'll use contradiction again: we'll search for f: N → (0,1), and show that no such f can be onto (contain every real number).

This time though, we won't use the structure provided by the AoC or the NIP. Instead, we want a more familiar structure. How do we introduce real numbers to students, theorems aside? What do you imagine when you think about real numbers? You probably use decimal form.

$\sqrt{(½)}$ = 0. 7   0   7   1   0   6   7 ...

We'll need a way to generalize this idea. We'll represent each digit with a variable.

$f(m) = 0.\, a_{m1}\, a_{m2}\, a_{m3}\, a_{m4}\, a_{m5}\, a_{m6}\, a_{m7} \dots$

To clarify, for $m, n \in N$, a_{mn} is the n^{th} digit of f(m).

###

Now that we have a way to represent one real number, we'll try to extend to our complete set. However, as we learned from Q, a visual representation can be really helpful: we'll try another kind of grid, since that worked well before.

If we have a 1-1 pairing between N and R, we should be able to list real numbers by counting in N. Let's try that, and make a decimal grid from that listing:

N
1 ⇔ $f(1) = .\, a_{11}\, a_{12}\, a_{13}\, a_{14} \dots$
2 ⇔ $f(2) = .\, a_{21}\, a_{22}\, a_{23}\, a_{24} \dots$
3 ⇔ $f(3) = .\, a_{31}\, a_{32}\, a_{33}\, a_{34} \dots$
4 ⇔ $f(4) = .\, a_{41}\, a_{42}\, a_{43}\, a_{44} \dots$
...   .

If f is onto and 1-1, then every real number is on this list once. For our proof, we'll need to contradict that by providing a number that isn't on this list, like we did in our last example.

###

A real number that isn't on this list would need to be somehow distinct from every number on this list: in this representation, at least one decimal place must be different from every number.

Let's try comparing to what we did before: we showed that we can use the NIP to "filter" out real numbers, one by one, until we had filtered out the entire list, and still had a real number.

So, let's try that again. How do we filter out f(1)? We can make sure one decimal place of our number, x, is different. We'll pick the first digit, because why not.

We can do this however we want. We could just add to the digit, and loop 9 around to 0 ($9 + 1 \rightarrow 0$ instead of 10). If a_{11} is 2, then the first digit of x is 3. If a_{11} is 7, x uses 8. Thus, we know $x \neq f(1)$, and we know the first digit of x.

What about f(2)? Well, we don't want to mess with our first digit, in case we accidentally change it to match f(1). So we'll edit the second digit of x to not match the second digit of f(2), the same way. Add 1 to a_{22}, unless it's 9, then it turns into 0. Now we know that $x \neq f(2)$.

We can repeat this process for every f(n), because we'll never run out of digits to modify.

###

Written formally, if we say that $x = .\, b_1 b_2 b_3 b_4 b_5 \ldots$

$$b_n = \begin{cases} a_{nn} + 1 & \text{if } a_{nn} \neq 9 \\ 0 & \text{if } a_{nn} = 9 \end{cases}$$

Visually, we can see we're moving down the diagonal. Because of this, this type of logic is often called **diagonalization**:

$$\begin{array}{lllllll} N & & & & & & \\ 1 & \Leftrightarrow f(1) = & . & a_{11} & a_{12} & a_{13} & a_{14} \ldots \\ 2 & \Leftrightarrow f(2) = & . & a_{21} & a_{22} & a_{23} & a_{24} \ldots \\ 3 & \Leftrightarrow f(3) = & . & a_{31} & a_{32} & a_{33} & a_{34} \ldots \\ 4 & \Leftrightarrow f(4) = & . & a_{41} & a_{42} & a_{43} & a_{44} \ldots \\ \ldots & \ldots & & \ldots & \ldots & \ldots & \ldots \\ x = & & . & b_1 & b_2 & b_3 & b_4 \ldots \end{array}$$

This x can't match f(1) because its first digit is different. It can't match f(2) because of the second digit. This logic continues for every digit, and at the end, we have a valid real number that cannot possibly be gotten from f(n). Thus, we've shown that f is not onto: it misses some.

Thus, (0,1) is uncountable: N is smaller than (0,1).

Now, to connect this proof to the N → R version: we can just replace the real numbers (0,1) with the real numbers R, and do exactly the same thing as before.

All that has changed is that some numbers have digits to the left of the decimal point. However, we can just ignore that section and start replacing the first digit to the right of zero, then the second, and so on, just like before.

Why? Because this still eliminates every single $x \in R$: as long as some digit is different, the rest don't matter. More digits to the left make no difference. Thus, the same proof we used for (0,1) still works.

Sets larger than R

Now that we are double sure that R is strictly larger than N, we wonder: can there be sets even larger than R? This is a tricky question, and one we'll approach carefully.

Once again, we return to the friendly land of finite sets to practice. If you have a set with n elements, how can you make a larger set? Whatever method we choose has to apply to R as well.

We could try to combine elements using an operation, like addition or multiplication. However, R is closed under both of those. "Closed under multiplication" means a set can't create new elements using multiplication. Anything we create is already in R. No good.

How else can we usually get a big, complicated thing from a simple collection of things? A bag of marbles, a collection of numbers, a deck of cards?

Well, you might have heard that there are a massive number of ways to shuffle a deck of cards: 52 factorial (52! = 52 x 51 x 50 x … x 2 x 1). That's over 8×10^{67}. That's a huge set of decks.

This is certainly bigger, but we need to see if it works in R. We want to re-order (shuffle) our set. However, R is uncountable, meaning we can't put it in a listed order in the first place.

If we cannot make a complete infinite list of every real number, then we can't rearrange that list. We'll need to try something else.

##

But, we'll keep the card idea: it does show us something notable, which is that our bigger set does not need to be filled with the same kind of things as our original set. In fact, we can have a set made out of other sets: in this case we had a set of shuffled card decks, not a set of cards.

As we've shown by shuffling, sets of sets seem to be a great way to make simple stuff complicated. So, if we can't use order, how else to make sets of sets? Can we try again to edit our original? Well, we can't create elements, and we can't order them, so why not delete some?

How do we apply this to our idea of cards? Well, imagine you have all 52 cards, and throw away 20. That's a new set of 32 cards. It would be the same to think about choosing those 32 to keep. Removing some cards can be thought of the same as grabbing only the rest.

For example, how many different hands of 4 cards can there possibly be? 52*51*50*49 = 270725. That's not nearly so big, but we can also find every hand of 3 cards, or every hand of 5 cards. These seem like they could add up pretty quickly.

Does this work for R? Well, "grabbing n elements from a set" is practically the definition of a subset. You take some group of elements from the larger set, and that becomes a new set. You can definitely take subsets of R: You could get Q, or even just the numbers {1, $\sqrt{2}$, 69}.

This seems promising! So, we'll create the set of all subsets: this is indeed a set of sets, as we expected. Just how much bigger is this set?

##

Well, let's take our deck of cards. How many subsets exist, including the empty set and the original set? Well, as we said before, for each card, we're choosing whether to keep or remove it. That's two choices.

If we had one card, we would have two sets: empty (remove), or original (keep). If we had two cards, we could choose whether to keep or remove the first card, and then whether to keep or remove the second. That means there's 2 x 2 possible scenarios.

This repeats indefinitely: each card doubles the number of subsets, because for each previous subset, there's now a version with and without that last card. Thus, for n cards, there are 2^n possible subsets. For 52 cards, that's 4.5×10^{15}. That's huge, even if it's less than the ordered.

Because the total is 2 to the power of n, we call this the **powerset** of our deck of cards. For a set A (set of cards in a deck), the powerset is written as P(A) (set of subsets of the card deck A).

–––

Power sets

Now that we've gotten a promising example that still applies to R, we need to see whether or not it's larger than R. For it to be a different cardinality, we need to show that we cannot possibly produce an f : R → P(R) that is both 1-1 and onto.

For starters, we could try to see if the same approach that worked for f: N → R will work here. In which case, we need to produce a set that isn't captured by f. That will show it is not onto. We will try to do this by eliminating each f(x) from our new set, one-by-one.

The approach used by Cantor involved making one digit different from each real number, thus eliminating each f(x) from possibly matching our new number. Maybe we could make one element different from each subset?

However, this time, we don't have an ordering to follow: not only are the subsets not in a countable sequence, but there is no n^{th} element to each set.

But this doesn't mean we have to give up. Just like how we created approaches for finite sets, and then tried to make them infinite, we'll use a countable example, and see if we can make it uncountable. This is easier to work with, and might inspire us.

First, let's make our input countable: we'll use a countable subset of R, called R_C. That way, we can work out the details for a set of sets in general.

$R_C \subset R, \quad R_C = \{ x_1, x_2, x_3, x_4 \ldots \}$

R_C
$x_1 \Leftrightarrow f(x_1) = \{ x_n : x \in R_C , ????\}$
$x_2 \Leftrightarrow f(x_2) = \{ x_n : x \in R_C , ????\}$
$x_3 \Leftrightarrow f(x_3) = \{ x_n : x \in R_C , ????\}$
$x_4 \Leftrightarrow f(x_4) = \{ x_n : x \in R_C , ????\}$
… … … … … …

Now we've got our subsets in some kind of order. But how do we display each set, so that we can match the decimal pattern? Well, our elements are countable too, so we could sort them into counting order.

R_C
$x_1 \Leftrightarrow f(x_1) = \{ x_1 \; x_3 \; x_4 \; \ldots \}$
$x_2 \Leftrightarrow f(x_2) = \{ x_1 \; x_3 \quad \ldots \}$
$x_3 \Leftrightarrow f(x_3) = \{ x_1 \; x_4 \quad \ldots \}$
$x_4 \Leftrightarrow f(x_4) = \{ x_2 \; x_3 \quad \ldots \}$
… … … … … …

##

There’s one problem we’ll run into if we try to complete the proof as we are, though: elements aren’t quite the same as digits. It may not be obvious now, but it becomes more clear as one works forward (we’ll skip that, of course. Who wants to do extra work only to get stuck)

The problem is this: a number can have the same digit in many decimal places (.555), but a set can have only one of an element ({5,5} is just {5}). It seems that the n^{th} digit of a number and the n^{th} element of an ordered subset aren’t equivalent.

This can cause several concerns:

What if we create a rule, and we end up with duplicate elements we have to omit? Is this an issue? Or we remove an element for $f(x_n)$, only to add it again later on accident (because x_n can appear in multiple columns)? What about finite sets, since we can’t do (.5 → .5000…)?

Maybe we could work around all these problems. But first, we’ll try to come up with a better way of organizing our elements, that better matches decimals.

We used decimal places for columns before, so what is the equivalent for subsets? Well, let’s compare sets to decimals in general. To make a real number, you pick a value for each decimal place. To make a subset, you pick whether to keep or remove each element of the original set.

So, each decimal place better matches each possible element: it's the choice of digit 0-9, or the choice of including or omitting the element.

So, each column is a possible x_n: if a set doesn’t have that element, we’ll leave a blank space. These spaces aren’t actually elements, they just represent the choice not to pick an element x_n.

R_C

$$
\begin{array}{llllllll}
x_1 & \Leftrightarrow f(x_1) = \{ & x_1 & x_2 & & x_4 & \ldots \} \\
x_2 & \Leftrightarrow f(x_2) = \{ & x_1 & & x_3 & & \ldots \} \\
x_3 & \Leftrightarrow f(x_3) = \{ & x_1 & & & x_4 & \ldots \} \\
x_4 & \Leftrightarrow f(x_4) = \{ & & x_2 & x_3 & & \ldots \} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots &
\end{array}
$$

This avoids the problems with elements: each column has a designed element that won’t repeat later, or appear twice, and we can add blank space at the end of a finite set. It’s temporary for solving this problem, of course. The blank spaces aren’t part of the set, they’re just a structure.

Now, we have each column n matching row n’s input: originally, row n contained f(n), so column n contained the n^{th} digit. Now, row n uses $f(x_n)$, so column n contains the x_n element.

##

There we go. Now, have all of our subsets ordered, and the elements are displayed in an ordered grid, with the column saying which element is or is not there.

Now we need to eliminate each set from possibly equalling ours. Before, we changed the n^{th} digit from $f(n)$. However, this time, we either have an element or we don't: the columns can't have multiple different elements.

To continue our parallel, we'll flip the x_n slot for $f(x_n)$.

Is x_n in this set, or not? That's all that a cell really describes. So, if we want to create a distinct set, we just need to flip it: include x_n if it's not in $f(x_n)$, and vice versa.

R_C

$x_1 \Leftrightarrow f(x_1) = \{\ x_1 \quad\quad x_3\ \ x_4\ \dots \}$

$x_2 \Leftrightarrow f(x_2) = \{\ x_1\ \ x_2 \quad\quad\quad \dots \}$

$x_3 \Leftrightarrow f(x_3) = \{\ x_1 \quad \underline{\quad}\ \ x_4\ \dots \}$

$x_4 \Leftrightarrow f(x_4) = \{ \quad x_2\ \ x_3\ \underline{\quad}\ \dots \}$

$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$

$B_C = \quad\quad \{\ \underline{\quad}\ \underline{\quad}\ \ x_3\ \ x_4 \dots \}$

Diagonalization logic returns!

B_C is unique for all the same reasons as before: it can't match $f(x_1)$ because it doesn't have x_1, it doesn't match $f(x_3)$ because it has x_3. This is true for every single subset, so B is a new subset. f is supposed to be able to match every subset, but it looks like we can always get a new subset.

We'll describe how we built B_C in a compact way:

$B_C = \{\ x \in R_C \ :\ \ x \notin f(x)\ \}$

Meaning, it will include x only if f(x) does not.

Suddenly, with our formalized version, we can easily modify it to work for the uncountable variation: we just need to replace R_C with R. This definition doesn't depend on order, because it's just a set, so uncountability isn't a problem. It can avoid every possible f(x).

$B = \{\ x \in R \ :\ \ x \notin f(x)\ \}$

The same logic applies: B can't match any f(x), because it will always either lack x or have x when f(x) does the opposite.

A short proof by contradiction: if $B = f(x)$ for some x, then both $x \in B$ and $x \notin B$ are illogical: if $x \in f(x)$, then by B's construction rules $x \notin B$, and $B \neq f(x)$. If you flip everything, you've proven it's illogical for $x \notin f(x)$.

In the same way we casually replaced R_C with R, we can actually generalize it further: this logic works for any set A. This means that not only is R smaller than P(R), A is always smaller than P(A).

Theorem 1.5.2 (Cantor's Theorem). Given any set A, there does not exist
a function $f : A \to P(A)$ that is onto.

1.6 Epilogue

We'll wrap up by looking at some definitions, and the bigger picture of what all of this means for math moving forward.

When two sets have the same cardinality ($N \sim Q$), we call that an **equivalence relation**.

All of the sets with this relation are in the same group, a group we call the **equivalence class**, containing every set of the same size.

For a bigger picture, we can imagine that every set that ever could exist lives in one of these classes: each class is disjoint, because a set can only have one of these sizes.

N, Z, and Q live in the class of countable sets, while R lives in the same class as (0,1). P(R), according to Cantor's theorem, is in another class, containing some sets bigger than R. This is true for any P(A), of course: it will be in a different class from A.

What about even bigger sets? Well, the power set worked once already, why not use it again? P(P(R)), despite seeming ridiculous, is perfectly valid, and is even bigger than P(R). So we can always create a larger class: there must be an infinite number of them, then.

This whole time, we've talked about collections of "bigger" or "smaller" sets, but it's a bit hard to describe each of them without a label. Let's try to give them labels moving forward. In fact, we'll order them from smaller to larger.

Well, for finite sets, we can just use the number of elements: it's a number for cardinality, so we'll call it a... **cardinal number**. Brilliant. We'll write the cardinal number of X as card X.

We'll extend this cardinal number to infinite sets, but this gets a bit tricky. We won't define this object very formally, because it involves a lot of set theory. And we don't really have time for that.

Well, up until now, we've described size for infinite sets in terms of what other sets are the same size. Q is N-sized, and so on. And as it turns out, one way to define card X is to pick some special set that is the same size as X. Meaning, it's the same equivalence class.

But in general, we'll treat it like a number that matches X to its class.

Like other numbers, two "cards" can be equal. So if card X = card Y, card X and card Y are the same special set in the same class. Of course, if their cards are the same, then X and Y are in the same class too. They're the same size: card X = card Y means the same thing as $X \sim Y$.

So, whatever defines $X \sim Y$ is true about card X = card Y: we need $f: X \to Y$ to be 1-1 and onto: the same requirement we had before. What about the other comparisons: $<$ and $\leq$? (There's also $>$ and $\geq$ but they're just flipped)

If Y is a bigger set than X, the special set in its class is bigger too: card X < card Y.

If this is true, $f: X \to Y$ cannot be onto: we'll always be able to find some $y \in Y$ that doesn't match any $f(x)$, because there are just too many y and not enough $x \in X$.

However, $f: X \to Y$ can definitely be 1-1: if Y is bigger than X, then there should be enough elements $y \in Y$ that every $x \in X$ can have a unique partner, with no overlap. There'll even be some y left over, since it isn't onto.

The last number comparison to explore is card X $\leq$ card Y: this would imply Y is either a larger set, or the same size as X.

Since both = and < conditions require for $f: X \to Y$ to be 1-1, we know that the 1-1 requirement is true. However, now we're unsure whether or not it is onto.

This notation gives us an easier way to state things like Cantor's Theorem: we can just say that card A < card P(A). P(A) is larger, so it has a greater cardinal number.

The fact that we can do these comparisons, and the endless sequence of card A < card P(A) < card P(P(A)) < ... does seem to imply a sort of order to the classes.

However, we need one thing to prove that we have order: only one of $<$, $>$, or $=$ can be true between any two items.

If none were true, then you can't compare those two items. Order requires comparisons, so that's out. If multiple were true, that would make it impossible to arrange them: how could we put things in order if one number is both less than and more than another number?

How can we confirm this statement, then? Well, let's think about how we usually use this fact: we often use it to prove two things are equal. If only one of the three is true, then eliminating two will force us to pick the last one: if $a < b$ is false and $a > b$ is false, then $a = b$ is true.

This works, but $<$ is annoying. It has both the 1-1 and onto requirements we might be able to negate. $\leq$ is much easier to work with: we could just show it is or isn't 1-1. Can we convert it into that form?

Indeed we can! If $a \geq b$ (not $a<b$) and $a \leq b$ (not $a>b$), then $a=b$. To translate this to cardinality, if card $A \leq$ card B and card $A \geq$ card B, then card $A =$ card B, or $A \sim B$.

What does this look like in terms of sets? If we can do $f: A \to B$ that is 1-1, and we can do $f: B \to A$ that is 1-1, then $f: A \to B$ is onto and 1-1. This proof of this is called the Schrodinger-Bernstein Theorem. We will skip it for now, but we could prove it ourselves with what we know.

As we referenced earlier, Cantor's Theorem shows us that there is no biggest set: you can always take the power set to get a bigger one. That means we can't possibly have U, the set of all possible things. It can't contain P(U).

This forces us to restrict what a set can be: we can't just say "it contains everything" and call it a day. In fact, set theory had to be carefully put together so the axioms prevented you from creating such confusing objects.

Another curiosity: as we said before, countable sets are the smallest infinite sets. If you made a smaller set S, it would have to be finite, because if it weren't, there's no reason you couldn't just fill every $n \in N$ by counting through the set forever. In which case, $S \sim N$ and you're stuck.

Because of this, the cardinal number of the countable sets is given a special symbol: $\aleph_0$ (said as "aleph null"/"aleph naught"). R is also an important set, so its cardinal number gets its own symbol too: c. We already know that $\aleph_0 < c$, of course: N is smaller than R.

But, since we're creating an order, we might ask whether there's anything in between those two cardinal numbers. Can we find a set A, where $\aleph_0 < \text{card } A < c$?

This seems like a reasonable enough thing to ask about, in the same way we noticed there's a rational number between any two real numbers. Is there something in this gap?

It turns out this question, called the **continuum hypothesis**, was actually a horrible mess to figure out. It was one of the biggest mathematical problems of the 20th century.

The twist? In 1940, Kurt Godel showed you couldn't prove the continuum hypothesis. In 1963, Paul Cohen showed that you couldn't disprove it either. You could take it as true or false, and you'll run into no contradictions: it's unprovable. Absolutely absurd.

Cantor's diagonalization proof had a significant impact on the writing of proofs in the future, for many other problems.

For example, Kurt Godel's Incompleteness Theorem proved that any axiom system built for arithmetic (i.e. elementary school math) is always "incomplete": some true statements could never be proven by your axioms. The proof used a similar type of method as diagonalization.

Another example: the "halting problem" asks whether there can be an algorithm (program/set of instructions) that can look at every program in existence, and figure out whether it eventually stops. There is no such algorithm, and this too, is proven with diagonalization.

A third example will appear later, in Chapter 6.
