Continuity, Metric Spaces and Topologies

1 Foundations on $\mathbb{R}$

This lecture is based largely on drawing analogies of important properties about real numbers to a more general setting. To this end, we will begin with a few observations that will motivate the forthcoming theory. One thinks of $\mathbb{R}$ intuitively as a "number line". If we were to plot the integers on this number line then we would be subdividing this number line into line segments, each of which of the form $[n, n+1)$, where $n \in \mathbb{Z}$. It is not difficult to see that the points in $[n, n+1)$ are in bijection with the points of $[0,1)$ through the map $x \mapsto x - n$, so $\mathbb{R}$ is in fact a union of a number of copies of $[0,1)$, where by a *copy*, we mean the image of a bijective map. This is significant because if we are interested in local information about $\mathbb{R}$, i.e., at the level of individual points and the points surrounding them, we can restrict to a finite interval.
The closed interval $[0,1]$ is an "uninterrupted" line segment of length 1 between two points 0 and 1, fixed in space (we view 0 and 1 as having some concrete existence, and define a system of distance by defining something as having length in multiples of the distance between 0 and 1). We will treat $[0,1]$ as an entity and make more rigorous this property of being uninterrupted presently. Certain elements of $[0,1]$ are distinguished because they have concrete meaning. Any rational number $\frac{p}{q}$, for instance, corresponds to the ratio of two positive integers p and q (for instance, to describe the relative sizes of a set of p objects and a set of q objects). We can be more precise.

Definition 1.1. Let S, T be sets. The *(Cartesian) product set* $S \times T$ is the set of all ordered pairs (s,t), where $s \in S$ and $t \in T$. When $S = T$, we write $S^2 := S \times S$.
An *equivalence relation* on a set S is a subset $R \subseteq S \times S$ that satisfies the following properties: i) $(a,a) \in R$ for every $a \in S$ (reflexivity); ii) if $(a,b) \in R$ then $(b,a) \in R$ (symmetry); iii) if $(a,b),(b,c) \in R$ then $(a,c) \in R$ (transitivity).
We write $a \sim b$ to imply that $(a,b) \in R$, and also refer (by abuse of notation) to the relation $\sim$ between elements as an *equivalence relation.* We say that two elements $a, b \in S$ are *equivalent* with respect $\sim$ if $a \sim b$, and refer to the *equivalence class* of an element $a \in S$ as the set of all $b \in S$ such that $a \sim b$.

In general, we will not refer explicitly to the set R implicit in the above definition; instead, we will only allude to its existence through statements like $a \sim b$, where $\sim$ is understood to refer to R.

Examples 1.2. a) Let S be a set. The simplest example of an equivalence relation is equality of elements S. Clearly, $a = a$; if $a = b$ then $b = a$; and if $a = b$ and $b = c$ then we must have $a = c$. In this case, we can state that $a \sim b$ if and only if $a = b$, and the corresponding subset of the product $S \times S$ is the *diagonal subset*, i.e., $\{(s,s) : s \in S\}$. The equivalence classes of any $a \in S$ with respect to this relation is the singleton $\{a\}$. b) An important example for our purposes is provided by letting $S = \mathbb{N} \times \mathbb{N}$ and setting $R := \{((m_1,n_1),(m_2,n_2)) \in S \times S : m_1n_2 - m_2n_1 = 0\}$. Writing $\mathbf{a}_j$ to denote the pair (m_j, n_j), where $j = 1$ or 2, it is clear that $\mathbf{a}_j \sim \mathbf{a}_j$, since $m_jn_j - m_jn_j = 0$, and since $0 = -0$, $\mathbf{a}_1 \sim \mathbf{a}_2$ implies that $\mathbf{a}_2 \sim \mathbf{a}_1$. Finally, if $\mathbf{a}_3 := (m_3, n_3)$ then observe that

$$m_2(m_3n_1 - m_1n_3) = m_3m_2n_1 - (m_3m_1n_2 - m_3m_1n_2) - m_2m_1n_3 = m_3(m_2n_1 - n_2m_1) - m_1(m_2n_3 - m_3n_2) = 0.$$

Since $m_2 \neq 0$, $m_3n_1 - n_3m_1 = 0$. Thus, $\mathbf{a}_1 \sim \mathbf{a}_3$, and $\sim$, as defined in this example, is an equivalence relation.
Observe that $\frac{a}{b} = \frac{c}{d}$, where a, b, c, d are positive integers, if and only if $ad - bc = 0$, in other words, that $(a,b) \sim (c,d)$. Setting $t := \gcd(a,b)$, we observe that if $a = a't$ and $b = b't$, where $\gcd(a',b') = 1$ then $t(a'd - b'c) = 0$. Again, $t \neq 0$, so this implies that $(a',b') \sim (c,d) \sim (a,b)$; thus, by transitivity, (a,b) and (a',b') are equivalent. This shows that any two equal fractions are

equivalent, and moreover, are equivalent to a corresponding fraction in lowest terms. This might be obvious, but it provides our more rigorous formulation of rationals.

Definition 1.3. A *positive rational number* is an equivalence class of pairs of integers $(a, b) \in \mathbb{N}^2$ with respect to the relation from the previous example. A *rational number* is either 0, a positive rational number or the product of a positive rational number and -1. We denote by $\mathbb{Q}$ its set.

Remark 1.4. We will refer to a rational number as the unique element of its equivalence that is in lowest terms, i.e., (a, b) such that $\gcd(a, b) = 1$, as we considered in b). Thus, according to this convention, the *numerator* and *denominator* refer to the coprime elements a and b, respectively.

Let $N \in \mathbb{N}_0$. We denote by $\mathcal{F}(N)$ the set of all distinct $\frac{p}{q} \in \mathbb{Q}$ such that $1 \leq p \leq N$. We call this set the *Farey sequence of order* N. We leave the following exercises to the reader.

Exercise 1.1. Fix $N \in \mathbb{N}$ and let $\frac{a}{b}, \frac{c}{d} \in \mathcal{F}(N)$, $\frac{a}{b} < \frac{c}{d}$.
a) Show that $\mathbb{Q} \cap [0, 1] = \bigcup_{N \geq 1} \mathcal{F}(N)$.
b) Prove the mediant property: $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$. Deduce that if $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive elements in $\mathcal{F}(N)$ then $b + d > N$.
c) Prove that $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive in $\mathcal{F}(N)$ if, and only if, $bc - ad = 1$. Deduce from this and the previous exercise that

$$\left|\frac{a}{b} - \frac{c}{d}\right| \leq \min\left(\frac{1}{b(N+1-b)}, \frac{1}{d(N+1-d)}\right) < \frac{2}{N+1}.$$

d) Though, strictly speaking, unrelated to the task at hand, we provide a challenging, number theoretical interlude for the reader about Farey sequences, namely giving a bound on the number of elements in $\mathcal{F}(N)$. The reader may want to return to this problem after having studied the next section after this one. i) Prove that $1 < \sum_{n \geq 1} \frac{1}{n^2} < 2$.
ii) Let $\phi(n) := \{1 \leq a \leq n : \gcd(a, n) = 1\}$. Show that $|\mathcal{F}(N)| = \sum_{n \leq N} \phi(n)$.
iii) Show that $\phi(p^k) = p^k - p^{k-1}$ whenever p is a prime and $k \geq 1$. Also, show that $\phi(p^k q^l) = \phi(p^k)\phi(q^l)$. Deduce that $\phi(mn) = \phi(m)\phi(n)$ whenever $\gcd(m, n) = 1$, and that $\phi(n) = n \prod_{p|n}\left(1 - \frac{1}{p}\right)$.
iv) Let $\mu(n)$ be the Möbius function, defined by $\mu(n) = 0$ if n is divisible by the square of any prime, and otherwise, $\mu(n) = (-1)^k$, where k is the number of distinct prime factors of n. Show (relating divisors of integers without square factors to subsets of the set of primes dividing n and using the principle of inclusion/exclusion) that $\sum_{d|n} \mu(d) = 0$ if $n > 1$, and $\mu(1) = 1$. Using iii), prove that $\phi(n) = \sum_{d|n} \frac{\mu(d)}{d}$, where the sum runs through all of the divisors of n.
v) Observe that $\left|\sum_{n \geq 1} \frac{\mu(n)}{n^2}\right| \leq \sum_{n \geq 1} \frac{1}{n^2}$. Using iv), show that $\left(\sum_{n \geq 1} \frac{\mu(n)}{n^2}\right)\left(\sum_{n \geq 1} \frac{1}{n^2}\right) = 1$. Deduce from this and the rest of the exercise that $\frac{1}{4}N^2 < |\mathcal{F}(N)| < \frac{1}{2}N^2$. Note that the non-triviality in this exercise is that $|\mathcal{F}(N)| > \frac{1}{2}N^2$. In fact, $|\mathcal{F}(N)|$ is about $\frac{3}{\pi^2}N^2$, the improved constant coming from an exact value for the sum in i).

Exercise a) shows that we recover all rational numbers in $[0, 1]$ by considering elements of Farey sequences. Exercises b) and c) together show that by taking mediants of consecutive elements in a Farey sequence of a given order, we produce new rational numbers that lie in gaps of the initial Farey sequence, and moreover, that these gaps decrease in size at least as quickly as $\frac{2}{N}$. Hence, by iteratively taking mediants, we get shorter and shorter gaps between rational numbers. Unfortunately, there is something going on between these gaps.

Exercise 1.2. Prove that if n is not a square, $\sqrt{n}, \frac{1}{\sqrt{n}} \notin \mathbb{Q}$.

As an illustration of important concepts in Mathematical analysis that we will treat systematically later, we can approach what is occurring in these gaps using the following concept.

Definition 1.5. Let S be a set. A *sequence* of elements in S is a function $f : \mathbb{N}_0 \to S$. We will also refer to a set $\{a_n\}_n$ defined by $a_n := f(n)$, as a sequence.

Examples 1.6. a) A trivial example of an integer-valued sequence is $a_n = n$ for $n \geq 0$. Less trivial, but well-known, is the Fibonacci sequence which assigns to n a number a_n satisfying $a_n = a_{n-1}+a_{n-2}$ for $n \geq 2$ (which, of course, requires an initial condition on a_0 and a_1, namely $a_0 = a_1 = 1$).
b) Let $N \in \mathbb{N}$ and let $r := \frac{p}{q} \in \mathcal{F}(N)\backslash\{1\}$. Let $r' := \frac{p'}{q'}$ be the smallest element in $\mathcal{F}(N)$ larger than $\frac{p}{q}$. Define $\text{med}(r, r') := \frac{p+p'}{q+q'}$ to be the mediant of r and r'. We may define a sequence as follows. Take $a_0 := r$, $a_1 := r'$ and for $n \geq 2$, $a_n := \text{med}(a_{n-1}, a_{n-2})$. Thus, $\{a_n\}_n \subset \mathbb{Q}$ whose elements, by our previous remarks, approach one another.
c) Let $\alpha \in \mathbb{R}$. We define $a_0 := \lfloor \alpha \rfloor$ and $\alpha_1 := \alpha - a_0$. For $n \geq 1$, define $a_n := \left\lfloor \frac{1}{\alpha_n} \right\rfloor$ and let $\alpha_{n+1} := \alpha_n - a_n$. The sequence $\{a_n\}_n$ is called the *continued fraction expansion* of α, and the a_n are called *partial quotients* of α. Observe that $\alpha_{n+1} := \{\frac{1}{\alpha_n}\}$ for each $n \in \mathbb{N}$. This is called the *Gauss map.*

Recall the following properties of the absolute value.

Exercise 1.3. Let $x, y \in \mathbb{R}$. Check that: i) $|x| = 0$ if, and only if, $x = 0$; ii) $|xy| = |x||y|$; iii) $|x+y| \leq |x|+|y|$. Clearly, iii) implies that $|x+y| \leq 2\max |x|, |y|$. Property iii) is called the triangle inequality. A function that satisfies these iii) properties is called a *generalized absolute value.*
Also, prove the following equivalent form of the triangle inequality: $||x| - |y|| \leq |x - y|$.
As a side note, suppose $\|\|$ is a function on $\mathbb{R}$ that satisfies i) and ii) above, and $\|x + y\| \leq \max\{\|x\|, \|y\|\}$ for any $x, y \in \mathbb{R}$. Prove by induction that $\|n\| \leq \|1\|$ for any integer n, evaluate, in two ways, $\|(x + 1)^n\|$, where $x \in \mathbb{Q}$. (We will see this absolute value come up later on. In contrast to the regular absolute value, which is called *Archimedean*, this type of absolute value is called *non-Archimedean.*)

Example b), above, is an important motivator for the following definition.

Definition 1.7. A sequence $\{a_n\} \subset \mathbb{R}$ is called a *Cauchy sequence* if for any $\epsilon > 0$ there is an $M = M(\epsilon) \in \mathbb{N}$ such that for any $n, m \geq M$, $|a_n - a_m| < \epsilon$.

Let us observe a few consequences of this definition.

Proposition 1.8. *a) Let $N \in \mathbb{N}_0$ and $b_n := a_{N+n}$ for each $n \geq 0$. Then $\{b_n\}_n$ is Cauchy if, and only if, $\{a_n\}_n$ is Cauchy.*
b) If $\{a_n\}_n$ is Cauchy then there is some $B > 0$ such that for all $n \in \mathbb{N}$, $|a_n| \leq B$, i.e., Cauchy sequences are bounded.
c) Let $\{a_n\}_n$ and $\{b_n\}_n$ be Cauchy. Then $\{c_n\}_n$ and $\{d_n\}$ are Cauchy, where $c_n := a_n + b_n$ and $d_n := a_n b_n$.

Part a) implies that we only care about the *tail behaviour* of Cauchy sequences, i.e., what occurs for elements with sufficiently large index. We'll refer to the sequence $\{b_n\}_n$ in a) as a *truncation* of $\{a_n\}_n$, since we lopped off the beginning of it.

Proof. a) If $\{a_n\}_n$ is Cauchy then by choosing $M'(\epsilon) := M(\epsilon) - N$ in the definition, we have $|b_m - b_n| = |a_{N+m} - a_{N+n}| < \epsilon$ whenever $m, n \geq M'$ (so that $N + n, N + m \geq M'$). Conversely, we take $M'(\epsilon) = M(\epsilon) + N$, and the same argument holds.
b) By the Cauchy property, we can choose $M \in \mathbb{N}$ such that for $n, m \geq M$, $|a_n - a_m| < 1$. Therefore, if we let $B' := \max\{|a_0|, \ldots, |a_M|\}$ and $B := B' + 1$, then clearly, $|a_n| \leq B' < B$ when $0 \leq n \leq M$, and when $n > M$, by the triangle inequality, $|a_n| \leq |a_n - a_M| + |a_M| \leq B' + 1 = B$.
c) By the triangle inequality, $|c_n - c_m| \leq |a_n - a_m| + |b_n - b_m|$. Thus, if M_1 and M_2 are such that,

respectively, $n, m \geq M_1$ implies that $|a_n - a_m| < \epsilon/2$ and $n, m \geq M_2$ implies that $|b_n - b_m| < \epsilon/2$ then $|c_n - c_m| < \epsilon$ whenever $n, m \geq \max(M_1, M_2)$.
For the second statement, we have the identity

$$|d_n - d_m| = |(a_n b_n - a_n b_m) + (a_n b_m - a_m b_m)| \leq |a_n||b_n - b_m| + |b_m||a_n - a_m|.$$

By b), we can find $A, B > 0$ such that $|a_n| \leq A$ and $|b_n| \leq B$ for all $n \in \mathbb{N}$. Hence, if we choose $M \in \mathbb{N}$ large enough so that for $n, m \geq M$ both $|a_n - a_m| < \frac{\epsilon}{2(A+1)}$ and $|b_n - b_m| < \frac{\epsilon}{2(B+1)}$ (i.e., choosing M to be the larger of M_1 and M_2 satisfying the bound for $\{a_n\}_n$ and $\{b_n\}_n$, respectively, then we have $|d_n - d_m| < A\frac{\epsilon}{2(A+1)} + B\frac{\epsilon}{2(B+1)} \leq \epsilon$, as required. □

Intuitively, one might suggest that the "uninterrupted" property of the real numbers is manifest in the fact that when elements are sufficiently close together, there is *some* real number between them to which they mutually get close. In effect, we can consider real numbers as representing Cauchy sequences in a similar sense as how rational numbers represent equivalence classes of pairs of integers.
Let Q be the set of all Cauchy sequences of rational numbers. We define an equivalence relation on Q by saying that $\{a_n\} \sim \{b_n\}$ if for each $\epsilon > 0$ there is some $N \in \mathbb{N}$ such that for $n \geq N$, $|a_n - b_n| < \epsilon$. Thus, equivalent sequences are eventually close to one another. According to this definition, a Cauchy sequence $\{a_n\}$ and its truncations are equivalent. Notice, as well, that any eventually constant sequence, i.e., such that $a_n = a$ for all $n \geq N$ and N some positive integer, is also Cauchy, and is equivalent to the constant sequence $(a, a, \ldots)$. Thus, we can identify rational numbers with Cauchy sequences by representing them by a constant sequence. At last, we can give a precise definition of $\mathbb{R}$.

Definition 1.9. The *set of real numbers* is the set of all equivalence classes in Q with respect to the equivalence relation given above.

This is not a satisfactory definition because we relate to elements of $\mathbb{R}$ as if they were individual numbers, not equivalence classes. Let us identify the equivalence class of $\{a_n\}_n \in Q$ with a number a such that for every $\epsilon > 0$ there is some $M \in \mathbb{N}$ such that for any $n \geq M$, $|a_n - a| < \epsilon$. Note that this is well-defined because if $\{a_n\}_n$ and $\{b_n\}$ are equivalent, they will both be uniformly close to a for sufficiently large N in the sense described earlier. This construction closes the gaps that were made evident by example b) above. Note also that both statements in c) of Proposition 1.8 allude to the fact that $\mathbb{R}$ is closed under multiplication and addition. With the following exercise, this will show that $\mathbb{R}$ is also closed under division by non-zero real numbers.

Exercise 1.4. Let $\{b_n\}_n$ be a Cauchy sequence that is not equivalent to the zero sequence. Show that $\{\frac{1}{b_n}\}_n$ is also Cauchy. Deduce that if $\{a_n\}_n$ is also Cauchy then so is $\{\frac{a_n}{b_n}\}_n$. (Hint: the argument invokes b) and c) of Proposition 1.8.)

By virtue of our identification of $\mathbb{Q}$ with the constant Cauchy sequences in $\mathbb{R}$, it follows, consistent with out intuition, that $\mathbb{Q} \subseteq \mathbb{R}$. It might also seem, because real numbers, by construction, are arbitrarily close to rational numbers, the rational numbers should represent a large proportion, in some sense (however vague), of the reals. We can illustrate this intuition using finite sets of integers, for instance. The smallest gap between integers is 1, and therefore if $F \subseteq \{1, \ldots, N\}$ is such that any two consecutive elements have a gap smaller than or equal to 2 between them then $|F| \geq \frac{1}{2}N$, or equivalently, that the ratio of the size of F to its superset $\{1, \ldots, N\}$ is $\frac{1}{2}$. Unfortunately, this sort of intuition is meaningless when we deal with infinite sets. This motivates the following sets of definitions, some of which may be familiar to the reader.

Definition 1.10. Let A, B be sets and let $f : A \to B$ be a map. We say that f is *injective* (or 1-1) if, whenever $f(a_1) = f(a_2)$ with $a_1, a_2 \in A$, $a_1 = a_2$. We say that f is *surjective* (or *onto*) if,

for each $b \in B$ there is (at least) an $a \in A$ such that $f(a) = b$. We say that f is *bijective* if f is injective and surjective.
If A, B are sets such that there exists an injective map $f : A \to B$ then we write $|A| \leq |B|$. If a surjective map exists between them then we write $|A| \geq |B|$. If a bijective map exists between them then we write $|A| = |B|$, and in this case, we say that the sets A and B are *equicardinal.*

Among finite sets, we can easily determine which sets are equicardinal. We caution the reader that we will be working in an abstract setting in which no explicit characterization of elements of our sets will be given. If he/she so chooses, he/she may as well think of these as real numbers until further notice.

Exercise 1.5. Given a collection $\mathcal{C}$ of sets, check that $\leq$ is a reflexive and transitive relation, i.e., that $|A| \leq |A|$ and if $|A| \leq |B|$ and $|B| \leq |C|$ then $|A| \leq |C|$. Note that it is evidently not an equivalence relation, as the following proposition already shows for finite sets.

Proposition 1.11. *Let A, B be finite sets containing m and n distinct elements, respectively. Then A and B are equicardinal if, and only if $m = n$.*

Proof. Write $A := \{a_1, \dots, a_m\}$ and $B := \{b_1, \dots, b_n\}$. If $m = n$ then the map $f : A \to B$ given by $f(a_j) = b_j$ is clearly a surjection, as each b_j is produced this way. Now suppose f were not injective. Let r_i denote the number of elements in A that map to b_i. By definition, f is injective if, and only if, each $r_i = 1$. It is easy to see that $\sum_{i=1}^n r_i = m$, since f acts on every element of A. On the other hand, since each r_i is a non-negative integer and f is surjective, each r_i is at least one, so we have $m \geq \sum_{i=1}^n 1 = n$, with equality if, and only if, $r_i = 1$ for each i. Since $m = n$, it must be the case that $r_i = 1$ for each i and hence f is also injective. Thus, f is a bijection and A and B are equicardinal.
Conversely, suppose, without loss of generality, that $m < n$ but A and B are equicardinal (otherwise, by symmetry, we can change the roles of A and B). By definition, there is a map $f : A \to B$ that is surjective. Thus, each $b \in B$ belongs to $f(A)$. However, $f(A) = \{f(a_1), \dots, f(a_m)\}$ has only at most $m < n$ elements. If f is surjective, on the other hand, each b_i must be represented in $f(A)$, so $f(A)$ must have at least n elements, a contradiction. Thus $m = n$. □

Infinite sets are bizarre when we use the finitary intuition implicit in the proposition above. For example, we should expect that $\mathbb{Z}$ should have *more* (again, in some sense) elements than $\mathbb{N}$ because it contains essentially a positive and a negative copy of $\mathbb{N}$. All the more so, $\mathbb{Q}$ should have more elements than $\mathbb{Z}$ and thus than $\mathbb{N}$, as $\mathbb{Q}$ corresponds to pairs of integers, rather than single one, in which either member can change (up to the coprimality condition, of course). However, the next result suggests otherwise.

Proposition 1.12. *i) Suppose $A \subseteq B$. Then $|A| \leq |B|$.*
ii) Suppose there is a surjection $B \to A$. Then there is an injection $A \to B$. iii) There exists an injective map $f : \mathbb{N} \to \mathbb{Q}$ and an injective map $g : \mathbb{Q} \to \mathbb{N}$.
iv) If $A \subset \mathbb{N}$ is infinite then there is an injective map $\mathbb{N} \to A$.

Part iii) is significant because of the following important theorem, which is beyond the scope of these notes (I give a proof in the appendix).
(For those that are counting or that are aware of these things, our proof of ii) requires the Axiom of Choice. As far as I know, there is no way to prove it otherwise.)

Theorem 1.13 (Cantor-Schröder-Bernstein)**.** *Let A and B be sets such that $|A| \leq |B|$ and $|B| \leq |A|$. Then $|A| = |B|$. In other words, if there is an injection in one direction and a second injection in the opposite direction between two sets (these maps being not necessarily inverses of one another) then A and B are equicardinal.*

The finite case of this theorem is essentially the content of the earlier proposition. But as the proof of Proposition 1.12 might suggest, the infinite case is far from intuitive.

Proof of Proposition 1.12. i) The first statement is trivial: the identity map sends an element $a \in A$ to the same element $a \in B$. If $f(a) = f(a')$ then $a = a'$, as claimed.
ii) Let $\phi : B \to A$ be the surjection. Define a map $\psi : A \to B$ by assigning to each a a single element in $\phi^{-1}(a) := \{b \in B : \phi(b) = a\}$. If $\psi(a) = \psi(a') = b$ then $\phi(b) = a$ and $\phi(b) = a'$ by construction. Thus, $a = a'$ and ϕ is indeed injective.
iii) The first injective map is trivial given i).
The second injective map is also not hard. Indeed, for $\frac{a}{b} \in \mathbb{Q}$ let $u = 1$ if $\frac{a}{b} < 0$ and $u = 0$ otherwise. Let $g(\frac{a}{b}) = 2^u 3^a 5^b$. By unique factorization, if $g(\frac{a}{b}) = g(\frac{a'}{b'})$ then $2^u 3^a 5^b = 2^{u'} 3^{a'} 5^{b'}$ if, and only if, $u = u'$, $a = a'$ and $b = b'$, whence $\frac{a}{b} = \frac{a'}{b'}$.
iv) Order the elements of A according to size and for each $a \in A$ let $F(a) := \{a' \in A : a' \leq a\}$. Hence, $F(a)$ consists of all of the elements of A that are less than or equal to a. Define $\phi : \mathbb{N} \to A$ such that $\phi(n) := a$, where $F(a)$ is equicardinal to $\{1, \ldots, n\}$. Note that this is well-defined because A is infinite, so there are at least n elements in A for each $n \in \mathbb{N}$. Suppose $a, a' \in A$. If $\phi(a) = \phi(a')$ then $|F(a)| = |\{1, \ldots, n\}| = |F(a')|$. By the Proposition 1.11, this means that $F(a')$ and $F(a)$ contain the same number of elements. If $a < a'$ and $b \in F(a)$ then $b \leq a < a'$, hence $b \in F(a')$, so that $F(a) \subset F(a')$. On the other hand, $a' \notin F(a)$, so $|F(a') \backslash F(a)| \geq 1$. This contradicts the fact that $F(a')$ and $F(a)$ have the same number of elements. Symmetrically, $a > a'$ gives the same contradiction. Thus, $a = a'$ and ϕ is an injection. Therefore, $|\mathbb{N}| \leq |A|$. □

By virtue of $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$ and $|\mathbb{N}| = |\mathbb{Q}|$, it follows that there is an injective map $\mathbb{Q} \to \mathbb{Z}$ by composing the injection from Proposition 1.12iii) with the trivial injection from i). Thus, $|\mathbb{Q}| = |\mathbb{Z}|$, as well. Since we have already demonstrated that there is some common class of sets, namely $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and any infinite subset of these, thus far, that are equicardinal, we may as well give a name to this class.

Definition 1.14. A set S is *countable* if $|S| = |\mathbb{N}|$, and is called *at most countable* if it is either countable or finite. If it is not at most countable, it is called *uncountable.*
If S is finite, we say that it has *cardinality* n, and we write $|S| = n$ if there is a bijective map between S and $\{1, \ldots, n\}$. If S is countable, we say that is has *countable cardinality,* and we express this as $|S| = \aleph_0$.

By convention, we say that the empty set is finite and has cardinality zero.
Thus, $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$ are all countable. Besides those that we've considered thus far, there are "many" other countable sets, as we have operations that produce countable sets from other countable sets.

Proposition 1.15. *i) If A and B are countable then $A \times B$ is countable. In particular, any finite product of countable sets is countable.*
ii) If $\{A_i\}_{i \geq 1}$ is a sequence of countable sets then $A := \bigcup_{i \geq 1} A_i$ is countable.

Note that from Proposition 1.12i) it is already clear that the intersection of any number of countable sets is at most countable (it might be empty, as the example $A_i := \{i, i+1, \ldots\}$, for $i \geq 1$, shows.

Proof. i) Let $\phi : A \to \mathbb{N}$ and $\psi_B : B \to \mathbb{N}$ be injections implied by the countability property and let $g : A \times B \to \mathbb{N}$ be defined by $g(a, b) := 2^{\phi(a)} 3^{\psi(b)}$. This is injective by unique factorization, since if $g(a, b) = g(a', b')$ then $\phi(a) = \phi(a')$ and $\psi(b) = \psi(b')$, and since ϕ and ψ are both injective, $a = a'$ and $b = b'$. Thus, $|A \times B| \leq |\mathbb{N}|$. On the other hand, since $|\mathbb{N}| = |A| \leq |A \times B|$ by fixing $b_0 \in B$ and letting $h(a) := (a, b_0) \in A \times B$ (which is easily seen to be injective), by the Cantor-Schroder-Bernstein theorem, $|A \times B| = |\mathbb{N}|$. We leave the claim about finite products of countable sets to the reader (noting that $(A_1 \times \cdots A_{n-1}) \times A_n = A_1 \times \cdots \times A_n$).

ii) Since there are infinitely many primes, the set of primes is countable, and we can set $\Pi : \mathbb{N} \to \mathcal{P}$ to be a bijection. Without loss of generality, we may assume that the sets A_i are disjoint (otherwise, we can set $B_1 := A_1$ and for $j \geq 1$, set $B_{j+1} := A_{j+1} \backslash \left(\bigcup_{i=1}^{j} B_i \right)$, and the union of the B_j is also A). Let $\phi_i : A_i \to \mathbb{N}$ be the injection implied by countability for each of the A_i and let $\phi : A \to \mathbb{N}$ be defined by $\phi(a) = \Pi(i)^{\phi_i(a)}$, where $a \in A_i$. Again, by unique factorization, $\phi(a) = \phi(a')$ if, and only if, $a, a' \in A_i$ for the same index i, and moreover, $\phi_i(a) = \phi_i(a')$, which, by injectivity, implies that $a = a'$. Thus, ϕ is an injection, and $|A| \leq |\mathbb{N}|$. Since $A_1 \subseteq A$ and thus $|\mathbb{N}| = |A_1| \leq |A|$, by the Cantor-Schroder-Bernstein theorem, $|A| = |\mathbb{N}|$. □

Remark 1.16. Note that in all of the above proofs, we only used the existence of an injection from a set A to $\mathbb{N}$, not a bijection. Therefore, in all of the statements above we may replace countable with *at most countable.*

Exercise 1.6. Recall that a set A is said to be countable if there is a bijection between A and $\mathbb{N}$, and uncountable otherwise.
i) Prove that the set of all finite subsets of a countable set is countable.
ii) Deduce from i) that the set of real numbers that are roots of an integer polynomial is countable. This will be of interest in a later set of lectures.
iii) Prove that if $|A| = |B|$ then if $\mathcal{P}(A)$ denotes the power set of A (the set of all subsets of A) then $|\mathcal{P}(A)| = |\mathcal{P}(B)|$, and hence deduce that given any countable set A, $\mathcal{P}(A)$ is uncountable.
iv) Let $x, y \in (0, 1)$, and write their decimal expansions $x = \sum_{j\geq1} a_j 10^{-j}$ and $y = \sum_{j\geq1} b_j 10^{-j}$. Let $f(x, y) := 0.a_1b_1a_2b_2\ldots$. Prove that f is an injective map. Also, prove that if $(0, 1)$ is uncountable then (a, b) is uncountable. Finally, observe that $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$ is a bijection, so deduce from this, or otherwise, that $\mathbb{R}^2$ and hence $\mathbb{R}^n$ are both equicardinal to $\mathbb{R}$.

So we have lots of countable sets. Is $\mathbb{R}$ countable? As a simple model, observe that every number can be written in a binary expansion. A binary expansion is almost, but not quite unique, because we can always replace a finite binary expansion, which is a rational number, with an infinite one in which a 1 at the end is replaced with an infinite sequence of 1's following a zero. For instance, in binary, the number 1 and the number $0.111\cdots$ are equivalent because they are arbitrary close to one another (a fact we make more explicit in the next section, in which the partial sums are partial binary expansions). In any case, up to this ambiguity, we can describe real numbers uniquely in terms of the binary sequence corresponding to the digits in its binary expansion. In other words, there is (at least) a surjection from $\mathbb{R}$ to the infinite product set $\{0, 1\}^{\mathbb{N}}$.
Note, however, that Proposition 1.15 only deals with *finite* products of at most countable sets (see the remark), while $\{0, 1\}^{\mathbb{N}}$ is an infinite product. As Cantor showed, this makes a infinity of difference (pun intended).

Theorem 1.17 (Cantor)**.** *The set $\{0, 1\}^{\mathbb{N}}$ is uncountable. In particular, $\mathbb{R}$ and the set of irrational numbers are both uncountable.*

Proof. We proceed according to Cantor's method. Let $X := \{0, 1\}^{\mathbb{N}}$ and suppose $|X| = |\mathbb{N}|$, so we can enumerate the real numbers $\{r_1, r_2, \ldots\}$ (this is the same thing as saying that we have a bijection $n \mapsto r_n$). We label the nth element of the sequence r_m as r_{mn}. We construct a sequence $a := \{a_n\}_n \in X$ such that if $a_n := 1 - r_{nn}$. Since $r_{nn} \in \{0, 1\}$, so is a_n, and $\{a_n\}_n$ is well-defined. By assumption, $a \in X$ so it is equal to some r_j. On the other hand, for any j, $a_j \neq r_{jj}$, which means that $a \neq r_j$. This contradiction implies that X and $\mathbb{N}$ are not equicardinal. At the very least, we have an injection $\mathbb{N} \to X$ by writing the integer n in terms of its binary expansion (this time to the left of the decimal point) and appending an infinite sequence of zeros after it. This is easily seen to be injective, as the reader may verify.
Now suppose $\mathbb{R}$ were countable. We already established that $\mathbb{R} \to X$ is a surjection. If $|\mathbb{N}| = |\mathbb{R}|$ then

we should have a surjection $\mathbb{N} \to \mathbb{R} \to X$, and hence $|X| \leq |\mathbb{N}|$ (note that we are using Proposition 1.11 here). But by what we just showed, $|\mathbb{N}| \leq |X|$, so by Cantor-Schroder-Bernstein, $|X| = |\mathbb{N}|$, which contradicts the fact that X and $\mathbb{N}$ are not equicardinal. Thus, $\mathbb{R}$ is also uncountable.
Lastly, since $\mathbb{R}$ is the disjoint union of $\mathbb{Q}$, which is countable, and the set of irrational numbers, and $\mathbb{R}$ itself is uncountable, the set of irrational numbers can not be countable, for by using Proposition 1.15ii) with $A_1 := \mathbb{Q}$ and $A_j := \mathbb{R}\backslash\mathbb{Q}$ for each $j \geq 2$, $\mathbb{R} = \bigcup_{j\geq 1} A_j$ would then be countable, a contradiction. □

So in fact, the gaps "between" rational numbers contain more information than the rationals themselves. If we now reassess our construction of real numbers, this seems plausible: the rationals, as Cauchy sequences, were equivalent to *constant sequences*. The set of non-constant sequences obviously has the potential to be much larger. All this to say that the math you may have learned earlier did not tell the whole story.

Remark 1.18. The reader is cautioned that much of the theory developed in this section belongs to the realm of Set theory, which is, in essence, at the very foundation of Mathematics and its point of contact with Philosophy. Mathematics, particularly as it pertains to infinitary objects, was developed based on certain self-evident truths about finitary sets. For example, the principle of Mathematical Induction implies that if we have a set and a rule for adjoining an element to it then we can adjoin elements arbitrarily many times. From this we abstract the idea of an infinite set. Strictly speaking, this means that the set contains at least N elements, for any positive integer N, i.e., that it is the result of adjoining N elements for any N, and that the first M elements, for $M < N$, correspond to the first M elements of the set when it only contained N elements. As we just observed, there are some subtleties involved in infinite sets, and therefore there is controversy over what sorts of abstractions and axioms should be foundational. We won't go any further with this, but if the reader has interest for such things, he/she may find my opinions on my webpage (this essay coming soon).

Appendix: Proof of the Cantor-Schröder-Bernstein Theorem

Recall that the C-S-B theorem allowed us to conclude, from the existence of injections from a set A to a set B and from B to A that A and B are equicardinal. We prove this presently.
Let $f : A \to B$ and $g : B \to A$ be the injections in question. Of course, f is a bijection onto $f(A) \subset B$ and g is a bijection from B onto $g(B)$. We will extend f to a surjective map onto all of B, while maintaining injectivity. Observe first that if $a \in g(B)$ then $g^{-1}(a)$ is well defined, and gives a map on all of $g(B)$ that is surjective onto B. If $a \in A\backslash g(B)$ then $b = f(a) \in B$, and thus b. We define two sequences as follows. Let $Y_0 := A\backslash g(B)$, and for $j \geq 1$ let $Y_j := g(f(Y_{j-1}))$. Note that g^{-1} is well-defined on Y_j. If $Z := \bigcup_{j=0}^{\infty} Y_j$ then $Z \subset A$. Intuitively, Z is built out of a set that g^{-1} misses, and so it is somewhat natural to consider a map that has characteristics of f on this set and of g^{-1} outside of this set. We will prove that the map given by $h(x) := f(x)$ when $x \in Z$ and $h(x) = g^{-1}(x)$ when $x \notin Z$ is injective and surjective.
Suppose that $h(a_1) = h(a_2)$ for some $a_1, a_2 \in A$. If $a_1, a_2 \in Z$ then $f(a_1) = f(a_2)$ implies that $a_1 = a_2$. Similarly, if $a_1, a_2 \notin Z$ then the injectivity of g implies that $a_1 = a_2$. Thus, without loss of generality, assume that $a_1 \in Z$ and $a_2 \notin Z$. Then $f(a_1) = g^{-1}(a_2)$, and $a_2 = g(f(a_1))$. But then $a_2 \in Z$ by definition, a contradiction. Thus, a_1 and a_2 are both in the same set, and hence $a_1 = a_2$ and h is injective.
Now, suppose that $b \in B$. If $b \in f(Z)$ then b is obviously in the range of h since $h = f$ on this set. Thus, assume $b \in B\backslash f(Z)$. If $b = f(c)$, where $c \in Z$ then $h(c) = f(c) = b$. Thus, assume that $b \in B\backslash f(A)$. Then $g(b) \notin g(f(A))$, hence $g(b) \notin Z$. Thus, $b = g^{-1}(g(b)) = h(g(b))$, and hence b is

in the range of h. This completes the proof. (Thanks to Wikipedia for a little bit of guidance on the proof.)