CONSTRUCTION OF NUMBER SYSTEMS

N. MOHAN KUMAR

1. Peano's Axioms and Natural Numbers

We start with the axioms of Peano.

Peano's Axioms. $\mathbb{N}$ *is a set with the following properties.*

(1) $\mathbb{N}$ *has a distinguished element which we call '1'.*
(2) *There exists a distinguished set map* $\sigma : \mathbb{N} \to \mathbb{N}$.
(3) σ *is one-to-one (injective).*
(4) *There does not exist an element* $n \in \mathbb{N}$ *such that* $\sigma(n) = 1$. *(So, in particular* σ *is not surjective).*
(5) *(Principle of Induction) Let* $S \subset \mathbb{N}$ *such that a)* $1 \in S$ *and b) if* $n \in S$, *then* $\sigma(n) \in S$. *Then* $S = \mathbb{N}$.

We call such a set $\mathbb{N}$ to be the set of natural numbers and elements of this set to be natural numbers.

Lemma 1.1. *If* $n \in \mathbb{N}$ *and* $n \neq 1$, *then there exists a unique* $m \in \mathbb{N}$ *such that* $\sigma(m) = n$.

Proof. Consider the subset S of $\mathbb{N}$ defined as,

$$S = \{n \in \mathbb{N} \mid n = 1 \text{ or } n = \sigma(m), \text{ for some } m \in \mathbb{N}\}.$$

By definition, $1 \in S$. If $n \in S$, clearly $\sigma(n) \in S$, again by definition of S. Thus by the Principle of Induction, we see that $S = \mathbb{N}$. Further injectivity of σ implies uniqueness as claimed in the lemma. This proves the lemma. □

We define the operation of addition (denoted by $+$) by the following two recursive rules.

(1) For all $n \in \mathbb{N}$, $n + 1 = \sigma(n)$.
(2) For any $n, m \in \mathbb{N}$, $n + \sigma(m) = \sigma(n + m)$.

Notice that by lemma 1.1, any natural number is either 1 or of the form $\sigma(m)$ for some unique $m \in \mathbb{N}$ and thus the defintion of addition above does define it for any two natural numbers n, m.

Similarly we define multiplication on $\mathbb{N}$ (denoted by $\cdot$, or sometimes by just writing letters adjacent to each other, as usual) by the following two recursive rules.

(1) For all $n \in \mathbb{N}$, $n \cdot 1 = n$.
(2) For any $n, m \in \mathbb{N}$, $n \cdot \sigma(m) = n \cdot m + n$.

Agian, lemma 1.1 assures that this defines multiplication of any two natural numbers. This procedure seems a bit informal and logically suspect. You are of course right. To be completely precise, we have to prove the Universal Property of Natural Numbers, first. This may look very abstract and so, if you wish, you may just trust the above formulas. But, we prove it for those more skeptical and those who are not afraid of abstractions.

1.1. **Universal Property of Natural Numbers.**

Theorem 1.1 (Universal Property of Natural Numbers)**.** *Let S be any set, $f : S \to S$ be any function and let $s \in S$ be a fixed element. Then there exists a unique function $\phi : \mathbb{N} \to S$ such that $\phi(1) = s$ and $\phi \circ \sigma = f \circ \phi$.*

Proof. The follwing proof is rather long, so we will discuss at least some part of thinking, which is not part of the proof itself. So, this will be in blue, while the proof itself will be in black.

The theorem asserts the existence and uniqueness of a function ϕ with some properties. We will not worry about the uniqueness, which is easy and concentrate on the existence. Since, at present, we just have some knowledge of set theory and we have assumed Peano's axioms, but little else and none of these tell us how to construct a function. By the first property we know that $\phi(1) = s$. If we call $\sigma(1) = 2, \sigma(2) = 3$ etc., which are just names we have given, the second property says $\phi(2) = \phi(\sigma(2)) = f(\phi(1)) = f(s)$, $\phi(3) = f(f(s))$ etc. But, we have no logical way of interpreting 'etc.'. So, we seem to be at an impasse. Thus, we are forced to rethink our path. Since we know a bit about sets and how to define them, can we interpret ϕ in terms of a set? We have seen that a function gives the graph, which is a set and we can retrieve the function from the graph. So, let me recall this.

Lemma 1.2. *Let A, B be sets. Then a subset $\Gamma \subset A \times B$ is the graph of a function from A to B if and only if $p : \Gamma \to A$, the first projection, is a bijection.*

Proof. If Γ is the graph of a function $\phi : A \to B$, then $\Gamma = \{(a, \phi(a)) | a \in A\}$. Then for any $a \in A$, $p^{-1}(a) = \{(a, \phi(a))\}$ and thus p is injective and surjective. So, p is a bijection.

Now, assume that $p : \Gamma \to A$ is a bijection. Then we can define ϕ as $\phi(a) = q(p^{-1}(a))$, where $q : \Gamma \to B$ is the second projection and $p^{-1} : A \to \Gamma$ is the inverse of p, which makes sense since p is a bijection. I will leave you to check that then Γ is the graph of this function ϕ. $\square$

So, in our situation, we need to find a subset $\Gamma \subset \mathbb{N} \times S$ such that $p : \Gamma \to \mathbb{N}$ is a bijection. What other properties should it have, if this is going to be the graph of ϕ asserted in the theorem? Since $\phi(1) = s$, we must have $(1, s) \in \Gamma$. Next let us interpret the second condition. Since Γ is expected to be the graph of the yet unconstructed function ϕ, for any $n \in \mathbb{N}$, we must have elements of the form $(n, \phi(n)) \in \Gamma$. So, if $(n, t) \in \Gamma$, then t is expected to be $\phi(n)$. Then, the second condition says the $\phi(\sigma(n)) = f(\phi(n))$ and hence $(\sigma(n), f(t)) \in \Gamma$. This says, if we define $\theta : \mathbb{N} \times S \to \mathbb{N} \times S$ as $\theta((n, t)) = (\sigma(n), f(t))$, then $\theta(\Gamma) \subset \Gamma$. So, these three conditions will ensure what we need. So, we start the proof.

First, we prove the existence of ϕ, uniqueness will be easy. We will construct the graph of ϕ, which in turn will define ϕ. So, we plan to construct a suitable subset Γ of $\mathbb{N} \times S$. First, we have a function $\theta : \mathbb{N} \times S \to \mathbb{N} \times S$, given by $\theta((n, t)) = (\sigma(n), f(t))$ for $n \in \mathbb{N}, t \in S$. We will have the required function if we can construct such a subset Γ satisfying the following three properties.

(1) $p : \Gamma \to \mathbb{N}$, the first projection is a bijection.
(2) $(1, s) \in \Gamma$.
(3) $\theta(\Gamma) \subset \Gamma$.

If we look for such a subset, clearly nothing immediately strikes one as a possible candidate. So, we are still stuck. Since Γ has to satisfy the three conditions above, may be we can find some set satisfying some of the conditions easily? Here we strike gold, since the set $\mathbb{N} \times S$ itself satisfy the second and third conditions. So, may be we should study all sets satisfying the last two conditions and then look for Γ among them. At least our search has narrowed down.

Let $\mathcal{C}$ be the set of all subsets of $\mathbb{N} \times S$ satisfying the last two conditions above. Then $\mathbb{N} \times S \in \mathcal{C}$ and hence this collection is non-empty.

How do we distinguish our Γ from these sets? If $X \in \mathcal{C}$, then we have $(1, s) \in X$ and $\theta(X) \subset X$. Since $\theta((1, s)) = (2, f(s))$, $\theta(2, f(s)) = (3, f(f(s)))$ etc. and these are precisely the elements expected to be in Γ, it seems that $\Gamma \subset X$. So, the Γ we are looking for seems to be the 'smallest' element in $\mathcal{C}$. So, it makes sense to look at the set Γ which is the intersection of all the sets in $\mathcal{C}$.

Let Γ be the intersection of all elelements in $\mathcal{C}$, which makes sense since this collection is non-empty. First let us check that $\Gamma \in \mathcal{C}$. This is easy, since $(1, s) \in X$ for all $X \in \mathcal{C}$ and Γ beng the intersection of such sets, $(1, s) \in \Gamma$. Similarly, for any $X \in \mathcal{C}$, $\theta(\Gamma) \subset \theta(X) \subset X$ and thus $\theta(\Gamma) \subset X$ for all $X \in \mathcal{C}$. So, by definition of Γ, $\theta(\Gamma) \subset \Gamma$. So, Γ satisfies the last two conditions and hence $\Gamma \in \mathcal{C}$.

Now, we tackle the first condition for this Γ. First, let us look at the set $G = \theta(\Gamma) \cup \{(1, s)\}$. Then clearly $G \subset \Gamma$. On the other hand, $(1, s) \in G$ and $\theta(G) \subset \theta(\theta(\Gamma)) \cup \theta(\{(1, s)\} \subset \theta(\Gamma) \subset G$. So, $G \in \mathcal{C}$ and thus by definition of Γ, we get $\Gamma \subset G$. This shows that

$$\theta(\Gamma) \cup \{(1, s)\} = G = \Gamma. \tag{1}$$

Now, we check that the first projection $p : \Gamma \to \mathbb{N}$ is a bijection. We use induction for this and so define a set,

$$T = \{n \in \mathbb{N} | p^{-1}(n) \subset \Gamma \text{ has exactly one element}\}.$$

Notice that as usual, we have defined this set so that if we can show $T = \mathbb{N}$, then p would be a bijection.

We have $p((1, s)) = 1$ and we wish to show that $p^{-1}(1) = \{(1, s)\}$. If not, say $(1, t) \in p^{-1}(1)$ with $t \neq s$. By equation 1, we see that $(1, t) \in \theta(\Gamma)$. So, there exists an element $(n, u) \in \Gamma$ such that $(1, t) = \theta((n, u)) = (\sigma(n), f(u))$. This says in particular, $1 = \sigma(n)$ contradicting Peano's axiom. This proves that $p^{-1}(1) = \{(1, s)\}$ and hence $1 \in T$.

Next, assume that $n \in T$. Then by definition, we have $p^{-1}(n) = \{(n, w)\}$. Since $(n, w) \in \Gamma$, we know that $\theta((n, w)) = (\sigma(n), f(w)) \in \Gamma$, since $\theta(\Gamma) \subset \Gamma$. Thus $p^{-1}(\sigma(n))$ contains $(\sigma(n), f(w))$. If we can show this is the only element in this set, we would have shown $\sigma(n) \in T$ and then by induction, we would be done. So, assume that $(\sigma(n), x) \in \Gamma$ and we want to show that $x = f(w)$. By Peano's axiom, $(\sigma(n), x) \neq (1, s)$ and hence, from equation 1, we see that $(\sigma(n), x) \in \theta(\Gamma)$ and thus there is an element $(m, y) \in \Gamma$ with $\theta((m, y)) = (\sigma(n), x)$. Then $\sigma(m) = \sigma(n)$ and $f(y) = x$. By injectivity of σ, we get $m = n$. Since $(m, y) = (n, y) \in \Gamma$ and $n \in T$ implies $y = w$. Then $x = f(w)$ and thus $(\sigma(n), x) = (\sigma(n), f(w))$ proving what we set out to prove. Thus $T = \mathbb{N}$ and hence p is a bijection.

Thus we have created a function $\phi : \mathbb{N} \to S$ satisfying the two required conditions.

The above argument shows uniqueness too, since Γ determines ϕ and Γ was forced to be the intersection of all elements in $\mathcal{C}$, so it had no choice. But, no harm in reproving it.

To show uniqueness, let $\phi' : \mathbb{N} \to S$ be another function satisfying the two conditions of the theorem. We wish to show $\phi(n) = \phi'(n)$ for all $n \in \mathbb{N}$ and so it makes sense to define the set $U = \{n \in \mathbb{N} | \phi(n) = \phi'(n)\}$. Then since $\phi(1) = (1, s) = \phi'(1)$ by the first condition, we see that $1 \in U$. If $n \in U$, then we have $\phi(n) = \phi'(n)$ and thus $\phi(\sigma(n)) = f(\phi(n))$ by second condition and hence, $f(\phi(n)) = f(\phi'(n)) = \phi'(\sigma(n))$

again by the second condition for ϕ'. So, $\phi(\sigma(n)) = \phi'(\sigma(n))$, proving $\sigma(n) \in U$. By induction, $U = \mathbb{N}$ and the theorem is proved. □

Just to illustrate the use of the above theorem, we prove the characterizing property of *infinite sets*, which we will define below.

Lemma 1.3. *Let S be any set. Then there exists an injective function $\phi : \mathbb{N} \to S$ if and only if there exists an injective but non-surjective function $f : S \to S$.*

Proof. Assume first that such a function ϕ exists and let $A = \phi(\mathbb{N})$ and $B = S - A$. Then $\phi : \mathbb{N} \to A$ is bijective, since ϕ is injective by assumption and surjective by choice of A. Define $f : S \to S$ by $f(b) = b$ for all $b \in B$ and $f(a) = \phi(\sigma(\phi^{-1}(a)))$ for any $a \in A$. Since $f(B) \subset B$ and $f(A) \subset A$, we prove the properties of f for A, B separately. f is injective and surjective on B by definition. Since σ, ϕ are injective, we see that f is injective on A. So, f is injective on S. We claim that f is not surjective. For this, let $\phi(1) \in A \subset S$. If there exists an $s \in S$ such that $f(s) = \phi(1)$, since $f(B) \subset B$, we must have $s \in A$. Then, $\phi(1) = f(s) = \phi(\sigma(\phi^{-1}(\phi(1)))) = \phi(\sigma(1))$. Since ϕ is injective, this means, $1 = \sigma(1)$, which contradicts Peano's axioms.

Next, we prove the converse. So, assume that we are given an injective non-surjective function $f : S \to S$. So, there exists an $a \in S$ such that there is no $s \in S$ with $f(s) = a$. There may be many such elements, but we fix any one of them and call it a. So, by universal property, we get a function $\phi : \mathbb{N} \to S$ such that $\phi(1) = a$ and $\phi(\sigma(n)) = f(\phi(n))$ for all $n \in \mathbb{N}$. We claim that ϕ is injective and then we will be done. We use induction to prove injectivity of ϕ. As always, when we need to prove something using induction, we must set up a suitable subset of $\mathbb{N}$, which if we can prove is all of $\mathbb{N}$, we should have proved what we set out to prove. This principle in mind, we define the set

$$T = \{n \in \mathbb{N} | \phi(n) = \phi(m), \text{ for some } m \in \mathbb{N}, \text{ then } m = n\}.$$

We first check that $1 \in T$. So, assume that $a = \phi(1) = \phi(m)$ for some $m \in \mathbb{N}$. If $m \neq 1$, then by lemma 1.1 we can write $m = \sigma(p)$ for some $p \in \mathbb{N}$. Then we get, $a = \phi(\sigma(p)) = f(\phi(p))$, which contradicts to our choice of a. So, $m = 1$ and thus $1 \in T$.

Next, assume that $n \in T$. We wish to show that $\sigma(n) \in T$. So, again, we start with the equation $\phi(\sigma(n)) = \phi(m)$ for some $m \in \mathbb{N}$. Then, $m \neq 1$, since $1 \in T$. So, we can write $m = \sigma(p)$ for some $p \in \mathbb{N}$, again by lemma 1.1. Thus, we get $\phi(\sigma(n)) = \phi(\sigma(p))$, which in turn implies by the property of ϕ that $f(\phi(n)) = f(\phi(p))$. Since f is injective, we get that $\phi(n) = \phi(p)$ and since $n \in T$, we get $n = p$. This implies,

$\sigma(n) = \sigma(p) = m$, which is what we wanted to prove. Thus, $\sigma(n) \in T$. Thus $T = \mathbb{N}$ by induction and then ϕ is injective. □

Definition 1. *We say that a set S is infinite if there is an injective function $\phi : \mathbb{N} \to S$.*

1.2. **Definition of addition and multiplication.** Now we are ready to rigorously define addition and multiplication of natural numbers. Fix any $m \in \mathbb{N}$. We will define an operation for any $n \in \mathbb{N}$, $m + n \in \mathbb{N}$ such that $m + 1 = \sigma(m)$ and $m + \sigma(n) = \sigma(m + n)$. For this, consider $S = \mathbb{N}$, $f = \sigma$ and $s = \sigma(m)$ in the above theorem 1.1. Then we have a function $\phi : \mathbb{N} \to \mathbb{N}$ such that $\phi(1) = \sigma(m)$ and $\phi \circ \sigma = \sigma \circ \phi$. So, if we call $\phi(n) = m + n$ (the addition symbol just represents this function), then it is trivial to check both the above properties.

To define multiplication, again we take, fixing an $m \in \mathbb{N}$, $S = \mathbb{N}$, $f : \mathbb{N} \to \mathbb{N}$ be, $f(n) = n + m$ (which is laready defined) and $s = m$. Then we get a function $\phi : \mathbb{N} \to \mathbb{N}$ by the Universal property, so that $\phi(1) = m$ and $\phi(\sigma(n)) = \phi(n) + m$. So, if we define $m \cdot n = \phi(n)$, then it staisfies both the properties of multiplication stated earlier and since m was any element of $\mathbb{N}$, we have defined multiplication for any two natural numbers. To get a feel for how we identify this set $\mathbb{N}$ as our usual number system, let me *prove* some of the properties we are familiar with. Remember, we may use only the axioms, definitions and whatever we have proved before to prove the successive statements. This principle should be rigidly adhered to follow our rules of logic.

Lemma 1.4 (Associativity)**.** *If $x, y, z \in \mathbb{N}$, then $x+(y+z) = (x+y)+z$.*

Proof. As before let us define a subset of $\mathbb{N}$ as follows.

$$S = \{z \in \mathbb{N} \mid \forall x, y \in \mathbb{N},\ x + (y + z) = (x + y) + z\}$$

To prove the lemma, we must show that $S = \mathbb{N}$ and again we plan to use the Principle of Induction. To apply the Principle, we must check two things and we will check them below.

Step 1: $1 \in S$.

For any $x, y \in \mathbb{N}$, we have,

$$\begin{aligned} x + (y + 1) &= x + \sigma(y) \quad \text{(by definition of addition)} \\ &= \sigma(x + y) \quad \text{(by definition of addition)} \\ &= (x + y) + 1 \quad \text{(by definition of addition)} \end{aligned}$$

Thus we get $1 \in S$.

Step 2: If $z \in S$, then $\sigma(z) \in S$.

For any $x, y \in \mathbb{N}$, we have

$$\begin{aligned} x + (y + \sigma(z)) &= x + \sigma(y + z) \quad \text{(by definition of addition)} \\ &= \sigma(x + (y + z)) \quad \text{(by definition of addition)} \\ &= \sigma((x + y) + z) \quad \text{(since } z \in S\text{)} \\ &= (x + y) + \sigma(z) \quad \text{(by definition of addition)} \end{aligned}$$

This proves the lemma. □

Lemma 1.5 (commutativity of addition)**.** *For any $x, y \in \mathbb{N}$, $x + y = y + x$.*

Proof. As always we start with a subset S of $\mathbb{N}$.

$$S = \{y \in \mathbb{N} \mid \forall x \in \mathbb{N}, \quad x + y = y + x\}$$

To use induction, we need to check two things. Of course, if we show $S = \mathbb{N}$, we would have proved the lemma.

Step 1: $1 \in S$.

For this we define a new subset T of $\mathbb{N}$ as follows.

$$T = \{x \in \mathbb{N} \mid x + 1 = 1 + x\}$$

We apply induction to this set T.

Step a): Clearly $1 \in T$, since $1 + 1 = 1 + 1$.

Step b): Assume $x \in T$. Then

$$\begin{aligned} 1 + \sigma(x) &= \sigma(1 + x) \quad \text{(by definition of addition)} \\ &= \sigma(x + 1) \quad \text{(since } x \in T\text{)} \\ &= \sigma(\sigma(x)) \quad \text{(by definition of addition)} \\ &= \sigma(x) + 1 \quad \text{(by definition of addition)} \end{aligned}$$

Thus we see that $T = \mathbb{N}$. Going back, we see that this implies $1 \in S$.

Step 2: If $y \in S$, then $\sigma(y) \in S$.

Let $x \in \mathbb{N}$.

$$\begin{aligned} x + \sigma(y) &= x + (y + 1) \quad \text{(by definition of addition)} \\ &= (x + y) + 1 \quad \text{(by associativity, proved before)} \\ &= (y + x) + 1 \quad \text{(since } y \in S\text{)} \\ &= 1 + (y + x) \quad \text{(since } 1 \in S\text{)} \\ &= (1 + y) + x \quad \text{(by associativity)} \\ &= (y + 1) + x \quad \text{(since } 1 \in S\text{)} \\ &= \sigma(y) + x \quad \text{(by definition of addition)} \end{aligned}$$

□

The correct order to prove some of the remaining properties of $\mathbb{N}$ after the above is the following. (There may be other possibilities, but at least this order will work).

(1) (Cancellative law) For any $x, y, z \in \mathbb{N}$, if $x + z = y + z$, then $x = y$.
(2) (Distributive law) If $x, y, z \in \mathbb{N}$ then $x \cdot (y + z) = x \cdot y + x \cdot z$ and $(y + z) \cdot x = y \cdot x + z \cdot x$.
(3) (Associative law for multiplication) For any $x, y, z \in \mathbb{N}$, $x(yz) = (xy)z$.
(4) For any $x \in \mathbb{N}$, $1 \cdot x = x$.
(5) (commutative law for multiplication) For any $x, y \in \mathbb{N}$, $xy = yx$.

Let me prove the distributive law now. We will only prove one of them.

Lemma 1.6 (Distributive law)**.** *For all $x, y, z \in \mathbb{N}$, $x(y+z) = xy+xz$.*

Proof. Again, let

$$S = \{z \in \mathbb{N} \mid x(y + z) = xy + xz, \forall x, y \in \mathbb{N}\}$$

Step 1: $1 \in S$.

$$\begin{aligned} x(y + 1) &= x\sigma(y) \quad \text{(by definition of addition)} \\ &= xy + x \quad \text{(by definition of multiplication)} \\ &= xy + x \cdot 1 \quad \text{(by definition of multiplication)} \end{aligned}$$

Step 2: If $z \in S$, then $\sigma(z) \in S$.

$$\begin{aligned} x(y + \sigma(z)) &= x\sigma(y + z) \quad \text{(by definition of addition)} \\ &= x(y + z) + x \quad \text{(by definition of multiplication)} \\ &= (xy + xz) + x \quad \text{(since } z \in S\text{)} \\ &= xy + (xz + x) \quad \text{(by associativity of addition)} \\ &= xy + x\sigma(z) \quad \text{(by definition of multiplication)} \end{aligned}$$

Thus by Induction, $S = \mathbb{N}$ and we have proved the lemma. □

Exercise 1. Prove the remaining properties stated above. Remember, you may use anything proved earlier for a proof, but no later property may be used in the proof.

1.3. **Ordering on $\mathbb{N}$.** Next we introduce the ordering on $\mathbb{N}$.

Definition 2. *If $n, m \in \mathbb{N}$, we say that n is less than m, written $n < m$, if there exists a $k \in \mathbb{N}$ such that $m = n + k$. We also write $n \leq m$, read n is less than or equal to m, to mean that either $n = m$ or $n < m$.*

Remark 1. As usual, we will write $m > n$, read m is greater than n to mean $n < m$. Similarl meaning is assigned to $m \geq n$.

Lemma 1.7. *If $x, y, z \in \mathbb{N}$ and $x < y$ and $y < z$ then $x < z$.*

Proof. The assumption $x < y$ means $y = x + k$ for some $k \in \mathbb{N}$. Similarly we get $z = y + l$ for some $l \in \mathbb{N}$. Thus, we get,

$$\begin{aligned} z &= y + l \\ &= (x + k) + l \\ &= x + (k + l) \quad \text{(by associativity)} \end{aligned}$$

Thus by definition $x < z$ since $k + l \in \mathbb{N}$. □

The same argument can be used to show the following.

Lemma 1.8. *Let $x, y, z \in \mathbb{N}$.*

(1) *If $x \leq y$ and $y < z$, then $x < z$.*
(2) *If $x < y$ and $y \leq z$, then $x < z$.*
(3) *If $x \leq y$ and $y \leq z$, then $x \leq z$.*

Exercise 2. If $x, y, z \in \mathbb{N}$ and $x < y$ show that $x + z < y + z$ and $xz < yz$.

Lemma 1.9. *For any $m \in \mathbb{N}$, $m \neq m + 1$.*

Proof. As usual define a subset $S \subset \mathbb{N}$ as follows:

$$S = \{n \in \mathbb{N} \mid n \neq n + 1\}$$

Clearly, $1 \in S$, since if not, $1 = 1+1 = \sigma(1)$, by definition of addition and $1 \neq \sigma(k)$ for any $k \in \mathbb{N}$ by Peano's axioms (number 4).

Assume $n \in S$. If $\sigma(n)$ is not in S, then $\sigma(n) = \sigma(n) + 1$. Then by definition of addition, $\sigma(n) = \sigma(\sigma(n))$. By the third axiom, this means, $n = \sigma(n)$ which in turn is $n + 1$ by definition of addition. This is impossible since $n \in S$. □

Lemma 1.10. *For any $m, k \in \mathbb{N}$, $m \neq m + k$.*

Proof. Again define a subset $S \subset \mathbb{N}$ as follows:

$$S = \{k \in \mathbb{N} \mid \forall m \in \mathbb{N}, m \neq m + k\}$$

From the previous lemma, we see that $1 \in S$. If $k \in S$, we want to show that $\sigma(k) \in S$ and then by induction we would be done. That is, we want to show that $m \neq m + \sigma(k)$ for any m. Notice that $m + \sigma(k) = \sigma(m + k)$, by definition of addition.

Let us define a subset T as follows:

$$T = \{m \in \mathbb{N} \mid m \neq \sigma(m + k)\}$$

Clearly $1 \in T$ by axiom 4. Assume $m \in T$. Want to show that $\sigma(m) \in T$. If $\sigma(m) = \sigma(\sigma(m) + k)$, by axiom 3, we see that $m =$

$\sigma(m) + k$. Thus we have,

$$\begin{aligned} m &= \sigma(m) + k \\ &= (m+1) + k \quad \text{(by definition of addition)} \\ &= m + (1+k) \quad \text{(by associativity)} \\ &= m + (k+1) \quad \text{(by commutativity)} \\ &= m + \sigma(k) \quad \text{(by definition of addition)} \\ &= \sigma(m+k) \quad \text{(by definition of addition)} \end{aligned}$$

But this contradicts our assumption that $m \in T$.

□

Lemma 1.11 (Well ordering of $\mathbb{N}$)**.** *If $n, m \in \mathbb{N}$, then exactly one of the following is true. Either $n < m$ or $n = m$ or $m < n$.*

Proof. Let us first prove only one of these can hold. If $n < m$, then by definition, $m = n + k$ for some element $k \in \mathbb{N}$. By the previous lemma, we see that $m \neq n$. If $m < n$, then there exists an $l \in \mathbb{N}$ such that $n = m + l$ which implies $m = (m + l) + k = m + (l + k)$ which again is not possible by the previous lemma. The other cases are similar and left as an exercise.

To finish the proof we consider the set S, fixing an n.

$$S = \{m \in \mathbb{N} \mid n < m, n = m \,\text{or}\, m < n\}$$

Step 1: $1 \in S$.

If $n = 1$, clearly $1 \in S$. If $n \neq 1$, then by lemma 1.1, $n = \sigma(k) = k + 1 = 1 + k$ and thus by definition of our ordering, $1 < n$.

Step 2: If $m \in S$ then $\sigma(m) \in S$.

Assume $m \in S$. This means we have three possibilities, namely $n < m$ or $n = m$ or $m < n$. First, let us look at the case $n < m$. Then $m = n + k$ for some k. Thus $\sigma(m) = \sigma(n + k) = n + \sigma(k)$ and so $n < \sigma(m)$ and hence $\sigma(m) \in S$. Next possibility is $n = m$. Then $\sigma(m) = m+1 = n+1$ and thus $n < \sigma(m)$ and again $\sigma(m) \in S$. Finally, we have the possibility of $m < n$. Thus $n = m + k$ for some element $k \in \mathbb{N}$. If $k = 1$, then $n = \sigma(m)$ and thus $\sigma(m) \in S$. If not, by lemma 1.1, $k = \sigma(l)$ and thus $n = m + \sigma(l)$ and thus,

$$\begin{aligned} n &= m + \sigma(l) \\ & m + (l+1) \quad \text{(by definition of addition)} \\ &= m + (1+l) \quad \text{(by commutativity)} \\ &= (m+1) + l \quad \text{(by associativity)} \\ &= \sigma(m) + l \quad \text{(by definition of addition).} \end{aligned}$$

Therefore $\sigma(m) < n$ by definition of our ordering and thus $\sigma(m) \in S$. Thus in any case $\sigma(m) \in S$. Therefore by induction, $S = \mathbb{N}$ and we are done.

□

Lemma 1.12. *If $a, b, c \in \mathbb{N}$ with $a < b$ then $ac < bc$.*

Proof. If $a < b$, then $b = a + k$ for some natural number k. Thus $bc = ac + kc$. Since $kc \in \mathbb{N}$, by definition, $ac < bc$. □

Corollary 1.1. *If $a < b$ for $a, b \in \mathbb{N}$, then $a^2 = a \cdot a < b^2 = b \cdot b$.*

Proof. From the previous lemma, since $a < b$, we get $a^2 < ab$. Applying the lemma again, we get $ab < b^2$. Putting them together, we get $a^2 < b^2$. □

Lemma 1.13. *Let $a, b \in \mathbb{N}$ and $1 < a$. Then there exists a natural number N such that for all $n \geq N$, $b < a^n$. (Recall, a^n is just a convenient way of writing the product of a, n times).*

Proof. Consider the set

$$S = \{b \in \mathbb{N} \mid \text{there exisits } N \text{ such that } \forall n \text{ with } n \geq N,\ b < a^n\}$$

Then $1 \in S$. For this take $N = 1$ and apply the previous lemma as follows. Let

$$T = \{n \in \mathbb{N} | 1 < a^n\}.$$

Then $1 \in T$ since $1 < a$. If $n \in T$, then we have, $1 < a^n$. Multiplying by a, from the previous lemma, we have $a < a^{n+1}$. Putting these together, we have $1 < a^{n+1}$ and thus $n + 1 \in T$ and thus by induction we see that $T = \mathbb{N}$, proving $1 \in S$.

Assume that $b \in S$. So, there exists N such that $b < a^n$ for all $n \geq N$. I claim that for $\sigma(b)$, we can take instead of N, $N + 1$. Let $n \geq N + 1$. Since $n \neq 1$, we may write $n = m + 1$ and $m \geq N$. Thus by induction hypothesis, $b < a^m$ and thus by the lemma above, $ab < a^n$. Since $1 < a$, we can write $a = k + 1$. Thus $ab = bk + b$ So, $ab = bk + b \geq b + 1$. Thus we get $b + 1 < a^n$. So $b + 1 \in S$ and we are done by induction.

□

Next I want to prove some alternate forms of induction which are frequently used. We start with a definition.

Definition 3. *Let $S \subset \mathbb{N}$. Then an element $n \in S$ is called a least element if for any $m \in S$, $n \leq m$.*

Lemma 1.14. *Let $S \subset \mathbb{N}$. If S has a least element, then it is unique. So it makes sense to use the definite article 'the' instead of 'a' if a least element exists and call it 'the least element'.*

Proof. Assume S has two least elements, say n and m. Thus we see that $n \leq m$ and $m \leq n$. From the well ordering lemma, it is clear that $n = m$.

□

Theorem 1.2 (First alternate form of Induction)**.** *If $S \subset \mathbb{N}$ and $S \neq \emptyset$ then S has a least element.*

Proof. As usual let

$$T = \{n \in \mathbb{N} \mid \text{if } n \in S \subset \mathbb{N}, \text{ then } S \text{ has a least element}\}$$

Caution: Here T is a fixed set defined as above. But, S is a variable subset of $\mathbb{N}$.

I will leave it as an exercise to check that $1 \in T$. Next assume that $n \in T$ and we want to show that $\sigma(n) \in T$. So let $S \subset \mathbb{N}$ with $\sigma(n) \in S$. If $n \in S$, then by hypothesis, we know that S has a least element. So assume that n is not in S and consider $A = S \cup \{n\}$. Then $n \in A$ and thus A has a least element by hypothesis. Let us call this least element a.

There are two possibilities. Either $a = n$ or $a \neq n$. If $a \neq n$, then $a \in S$. If $m \in S$, then clearly $m \in A$ and thus $a \leq m$. So we see that $a \in S$ is a least element.

Next assume $a = n$. Then I claim that $\sigma(n)$ is the least elemnt of S. If $m \in S$, we know that $a = n \leq m$. But $n \notin S$ and thus $n \neq m$. Thus by definition of less than or equal to, we see that $n < m$. Thus we may write $m = n + k$ for some $k \in \mathbb{N}$ by defintion. If $k = 1$, then $m = n + 1 = \sigma(n)$. If $k \neq 1$, then by the first lemma, $k = \sigma(l)$ for some $l \in \mathbb{N}$. Thus

$$m = n + (l + 1) = n + (1 + l) = (n + 1) + l = \sigma(n) + l$$

and thus $\sigma(n) < m$. Thus we see that $\sigma(n)$ is a least element of S and we are done.

□

Some of you may be more familiar with the following form of induction, though all the three are equivalent.

Theorem 1.3 (Second alternate form of Induction)**.** *Let $P(n)$ be mathematical statements for $n \in \mathbb{N}$. Assume*

1. *$P(1)$ is true.*
2. *If $P(n)$ is true, then $P(n + 1)$ is true.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Define a set $S = \{n \in \mathbb{N} | P(n) \text{ is false}\}$. We wish to show that $S = \emptyset$. If non-empty, by the previous form of induction,Theorem 1.2, we have a least element $m \in S$. By the first hypothesis of the theorem, $m \neq 1$. Then $m = p + 1$ for some $p \in \mathbb{N}$. Since $p < m$ and m being the least elelement of S, we know that $p \notin S$ and thus $P(p)$ is true by definition of the set S. Now, by our second hypothesis, $P(p+1) = P(m)$ is true and hence $m \notin S$, a contradiction, proving the result. □

Theorem 1.4 (Third alternate form of Induction)**.** *Let $P(n)$ be mathematical statements fos $n \in \mathbb{N}$. Asuume,*

(1) *$P(1)$ is true.*
(2) *If $n > 1$ and $P(k)$ is true for all $k < n$, then $P(n)$ is true.*

Then $P(n)$ is true for all n.

Proof. As before, let $S = \{n \in \mathbb{N} | P(n) \text{ is false}\}$ and we wish to show that $S = \emptyset$. So assume that it is non-empty and let $m \in S$ be the least element assured by Theroem 1.2. Again, as before, by the first hypothesis, $1 \notin S$ and thus $m > 1$. By minimality of m, if $k < m$, then $k \notin S$ and hence $P(k)$ is true. Thus by second hypothesis $P(m)$ is true and thus $m \notin S$, which is a contradiction, proving the theorem. □

We will use these forms in the next section on Number Theory to prove results familiar to you. We state some more properties of natural numbers, which can be proved using the above ordering properties of $\mathbb{N}$.

(1) (cancellative law for multiplication) For any $x, y, z \in \mathbb{N}$, if $xz = yz$ then $x = y$.
(2) (uniqueness of identity) For some $x, y \in \mathbb{N}$, if $xy = x$, then $y = 1$.

At this point, we will use our usual nomenclature for natural number. We already have called a special number 1 and then we call $2 = 1 + 1, 3 = 2 + 1$ etc. in the usual fashion.

2. Finite sets

We defined earlier in definition 1.3 that a set S is infinite if we have an injective function $\phi : \mathbb{N} \to S$. In this section, we wish to define a finite set and prove some elementary properties.

For a natural number n, let $\Sigma_n = \{m \in \mathbb{N} | m \leq n\}$. We have a natural inclusion $i_n : \Sigma_n \to \Sigma_{n+1}$.

Lemma 2.1. (1) *For all $n \in \mathbb{N}$, $\Sigma_{n+1} = i_n(\Sigma_n) \cup \{n + 1\}$. So, $\Sigma_n = \{1, 2, \ldots, n\}$.*

(2) *If $a \neq b \in \Sigma_n$, there exists a bijection $\phi_{a,b} : \Sigma_n \to \Sigma_n$ such that $\phi_{a,b}(a) = b, \phi_{a,b}(b) = a$ and $\phi_{a,b}(c) = c$ for all $c \neq a, c \neq b$.*
(3) *If $f : \Sigma_n \to \Sigma_m$ is injective, then $n \leq m$ and if it is surjective, $n \geq m$. Thus, if it is bijective, $n = m$.*

Proof. (1) Clearly both $i_n(\Sigma_n)$ and $n+1$ are contained in Σ_{n+1} and hence the right hand side is contained in the left. To prove the reverse inclusion, let $a \in \Sigma_{n+1}$. If $a = n + 1$, then it is in the right hand side. If not, since $a \leq n + 1$ and $a \neq n + 1$, we get that $a < n + 1$ and hence $a \leq n$. Thus, $a \in i_n(\Sigma_n)$.
(2) This part is clear.
(3) We will prove the result for f injective, the surjective case being similar. Proof is by induction. So, let $S = \{n \in \mathbb{N} | \text{for any } f : \Sigma_n \to \Sigma_m, \text{ injective, } n \leq m\}$. Clearly $1 \in S$, since $1 \leq m$ for any $m \in \mathbb{N}$. So, assume that $n \in S$ and let $f : \Sigma_{n+1} \to \Sigma_m$ be an injection. Let $f(n + 1) = a$. If $a \neq m$, we can use the second part to construct a bijection $\phi_{a,m} : \Sigma_m \to \Sigma_m$. Then, $g = \phi_{a,m} \circ f : \Sigma_n \to \Sigma_m$ is an injection and $g(n+1) = \phi_{a,m}(f(n+1)) = \phi_{a,m}(a) = m$. Thus, in any case we may further assume that $f(n + 1) = m$. So, if $a \neq n + 1$, then $f(a) \neq m$, by injectivity. This gives an injection $f' : \Sigma_n \to \Sigma_{m-1}$. But, $n \in S$ implies that $n \leq m - 1$ and thus $n + 1 \leq m$. This proves that $n + 1 \in S$ and by induction, we are done.

□

Definition 4. *A set S is finite if either it is empty or bijective to Σ_n for some $n \in \mathbb{N}$.*

Notice from the previous lemma, that if a set is bijective to some Σ_n, then this n is unique. We will call this n the *cardinality* of S and write $|S| = n$. Once we have zero, we will say that the empty set has cardinality zero.

Corollary 2.1 (Pigeon Hole Principle)**.** *Let $f : S \to T$ be a function between finite sets. If $|S| > |T|$, then f is not injective.*

Proof. By definition, S is bijective to Σ_n and T is bijective to Σ_m for $n = |S|, m = |T|$ and by assumption, $n > m$. Any function f as above gives using these bijections a function $g : \Sigma_n \to \Sigma_m$ and f is injective (resp. surjective) if and only if g is. Now, by part 3 of lemma 2.1, g can not be injective. □

What we need to show next, as the terminology suggests, is that a set is either finite or infinite. It is again clear from the previous lemma

that a set can not be both finite and infinite. What is not clear is that any set is one of these. Before proceeding, I ask you to mull over how one could prove such a result.

It turns out that to do this, we need yet another axiom, called the Axiom of Choice. Of course, we could take our axiom to be that any set is either infinite or finite. But, for various reasons, the Axiom of Choice which will imply this, is preferred.

Axiom of Choice. *Given a collection of non-empty disjoint sets $\{X_\alpha\}$, we can form a set which consists of exactly one element from each X_α.*

We start with an easy corollary to the Axiom of Choice.

Corollary 2.2. *Let $f : A \to B$ be a surjective map of non-empty sets. Then there exists a function $s : B \to A$ such that $f \circ s$ is the identity map of B. Such a map is called a section for f.*

Proof. For any $b \in B$, let $X_b = f^{-1}(b) \subset A$. Since f is surjective, these are non-empty. Also, $X_b \cap X_c = \emptyset$ if $b, c \in B$ and $b \neq c$. So, by axiom of choice we can construct a set Γ which contains exactly one element from each X_b. If $x \in \Gamma$, then $x \in X_b$ for some $b \in B$ and hence $x \in A$. This proves that $\Gamma \subset A$. If we look at the restriction of the function f to this set Γ and call it g, we see that g is surjective, since $\Gamma \cap X_b \neq \emptyset$ for any $b \in B$ and since this set contains exactly one element, we also get that g is injective. Thus g is bijective. Define $s : B \to \Gamma \subset A$ to be the inverse of g. The rest is clear. □

Theorem 2.1. *Any set is either finite or infinite.*

Proof. Let S be any set and assume that it is not finite. Then $S \neq \emptyset$ and there is no bijection from Σ_n to S for any $n \in \mathbb{N}$. We wish to construct an injective map from $\mathbb{N}$ to S. We proceed as follows.

Let F_n denote the set of all injective maps from Σ_n to S for $n \in \mathbb{N}$. If $f \in F_{n+1}$, then if we restrict f to the subset $\Sigma_n \subset \Sigma_{n+1}$, we get a map $f' : \Sigma_n \to S$. Since f is injective, it is clear that so is f'. Thus we get a function $\pi_n : F_{n+1} \to F_n$ by $\pi_n(f) = f'$. We claim that this map is surjective for any $n \in \mathbb{N}$.

If $f \in F_n$, f can not be surjective, since then we will have a bijective map $f : \Sigma_n \to S$, contrary to our assumption. So, let $a \in S - f(\Sigma_n)$ and define a map $g : \Sigma_{n+1} \to S$ by $g(m) = f(m)$ if $m \leq n$ and $g(n+1) = a$. Easy to check that g is injective and $\pi_n(g) = f$, proving surjectivity.

Thus by the above lemma, we can construct maps $s_n : F_n \to F_{n+1}$ such that $\pi_n \circ s_n$ is the identity map on F_n. $F_1 \neq \emptyset$, since there are certainly functions from $\Sigma_1 = \{1\}$ to S, since $S \neq \emptyset$ and functions from a set with one element are always injective. So, pick an $f_1 \in F_1$.

Define recursively, $f_{n+1} = s_n(f_n)$. (Remember, recursive definition means, using the Universal Property of $\mathbb{N}$. Can you write a rigorous argument?) Define a function $\phi : \mathbb{N} \to S$ by $\phi(n) = f_n(n)$. We want to check that this map is injective. If $f_n(n) = \phi(n) = \phi(m) = f_m(m)$ and $n \neq m$, we may assume that $m > n$.

It is easier to prove a slightly general statement first. Let $k \leq n$ so that $k \in \Sigma_n$. Then for any $m \geq n$, $f_m(k) = f_n(k)$. We will use induction. So, let $S = \{m \in \mathbb{N} | m \geq n \,\text{and}\, f_m(k) \neq f_n(k)\}$. If $S = \emptyset$, we are done. If not, by induction, S has a least element, say p. Then $p > n$, and $f_{p-1}(k) = f_n(k)$ (where we have written $p-1$ for the natural number such that $\sigma(p-1) = p$, which exists, since $p > n$ and hence $p \neq 1$). We have $f_p(k) = (\pi_{p-1} \circ f_p)(k)$ since $k < p$. We also have $f_p = s_{p-1}(f_{p-1})$, by definition of f_n. Thus, $f_p(k) = (\pi_{p-1} \circ s_{p-1})(f_{p-1})(k)$. Since $\pi_{p-1} \circ s_{p-1}$ is the identity of F_{p-1}, we get that $f_p(k) = f_{p-1}(k) = f_n(k)$ and thus $p \notin S$. This contradiction proves our claim.

Going back to the proof, we see that $f_n(n) = f_m(m)$, but $f_n(n) = f_m(n)$ from the previous paragraph since $m > n$. Then $f_m(n) = f_m(m)$, contradicting the injectivity of f_m.

□

3. Integers

We will briefly desribe the construction of integers and rational numbers below and state various properties in the correct order and prove just a few to give a flavour.

Consider the set $S = \mathbb{N} \times \mathbb{N}$ and put a relation on it as follows: $(a, b) \sim (c, d)$ if and only if $a + d = b + c$. (As usual, we denote a typical element in S by an ordered pair of natural numbers)

Check that this is an equivalence relation on S. Let $\mathbb{Z}$ be the set of equivalence classes under this relation. Define an operation (addition) on $\mathbb{Z}$ as follows: If $A, B \in \mathbb{Z}$, then recall that A, B are non-empty subsets of S and thus we may pick elements $(a, b) \in A$ and $(c, d) \in B$. With our notation for equivalence classes, this means $A = [(a, b)]$ for example. Define an operation tentatively denoted by $\oplus$, to avoid confusion, as follows:

$$A \oplus B = [(a + c, b + d)]$$

There is *a priori* a problem with this definition. To make sure that the operation is *well-defined* (a term we will see several times in the sequel), we must make sure that the right hand side above has only value. This operation is supposed to be a function from $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, given two integers, we must get a well-defined integer as its sum. But,

let us look at our definition. Here we picked some $(a, b) \in A$ and $(c, d) \in B$ and declared that $A \oplus B = [(a + c, b + d)]$. We could have easily picked another $(a', b') \in A$ and $(c', d') \in B$. Then we would have got $A \oplus B$ to be $[(a' + c', b' + d')]$, which may very well be different from $[(a + c, b + d)]$ and then we really do not have function from $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ and no real definition. So, in such situations, the first order of business is always making sure that it is well-defined. In other words, with the above notation, we must check that $[(a + c, b + d)] = [(a' + c', b' + d')]$. For once let us check this.

Lemma 3.1. *The addition defined above is well-defined.*

Proof. As discussed, we must check that if $A = [(a, b)] = [(a'b')]$ and $B = [(c, d)] = [(c', d')]$, then we must check that $[(a + c, b + d)] = [(a' + c', b' + d')]$. Unwinding the defintions, this means that if $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$, then we must check $(a + c, b + d) \sim (a' + c', b' + d')$. Again looking at our relation, we get that $a + b' = a' + b$ and $c + d' = c' + d$. Adding them we get $a + b' + c + d' = a' + b + c' + d$, which implies that $(a + c, b + d) \sim (a' + c', b' + d')$. □

(Intuitively, the element $[(a, b)]$ should be thought of as $a - b$ in our familiar settings though we are yet to define *subtraction*.) Define multiplication tentatively denoted by $\otimes$ as follows:

$$A \otimes B = [(ac + bd, ad + bc)]$$

and as before make sure that this is well-defined. Now proofs of all the familiar properties of addition and multiplication of integers can be carried out, by using the definitions and corresponding properties of natural numbers.

(1) Associativity of addition.
(2) Commutativity of addition.
(3) Cancellative property of addition.
(4) For any two natural number $a, b \in \mathbb{N}$, $(a, a) \sim (b, b)$ and thus $[(a, a)] = [(b, b)]$ whic we denote by the symbol 0. Then for any $A \in \mathbb{Z}$, $A \oplus 0 = A = 0 \oplus A$.
(5) Additive inverse: If $A = [(a, b)]$, then we denote by $-A = [(b, a)]$, the *additive inverse* of A. Then $A \oplus (-A) = 0$.
(6) If $A \oplus B = 0$ then $B = -A$ and in particular the additive inverse is unique.
(7) Distributivity.
(8) Associative law for multiplication.
(9) Commutative law for multiplication.

(10) For any $a, b \in \mathbb{N}$, $(\sigma(a), a) \sim (\sigma(b), b)$ and thus we denote the equivalence class $[(\sigma(a), a)]$ for any $a \in \mathbb{N}$, by the symbol 1. Then for any $A \in \mathbb{Z}$, $A \otimes 1 = A = 1 \otimes A$.
(11) Cancellative law for multiplication: If $A, B, C \in \mathbb{Z}$ and $A \otimes C = B \otimes C$ with $C \neq 0$, then $A = B$.
(12) Uniqueness of identitiy: If $A \otimes B = A$ and $A \neq 0$, then $B = 1$.
(13) $A \otimes 0 = 0$ for all $A \in \mathbb{Z}$.
(14) If $A \otimes B = 0$, then either $A = 0$ or $B = 0$.
(15) For any $a, b, k \in \mathbb{N}$, $(a+k, a) \sim (b+k, b)$ and thus $[(a+k, a)] = [(b+k, b)]$. So define a map $f : \mathbb{N} \to \mathbb{Z}$ by $f(k) = [(a+k, a)]$ for some $a \in \mathbb{N}$. Then f is one-one and $f(a+b) = f(a) \oplus f(b)$ and $f(ab) = f(a) \otimes f(b)$.

The last property ensures that $\mathbb{N} \subset \mathbb{Z}$ via f and the addition and multiplication are respected by f. Thus we may now drop $\oplus, \otimes$ and write just $+, \cdot$.

As before we define an ordering on $\mathbb{Z}$ by saying that $A < B$ if $A = [(a, b)]$ and $B = [(c, d)]$, then $a + d < b + c$. Make sure that this is well defined and we define $A \leq B$ if either $A = B$ or $A < B$. Also show that if $a, b \in \mathbb{N}$, then $a < b$ if and only if $f(a) < f(b)$ and thus the ordering is also respected by f. We say that $A \in \mathbb{Z}$ is *positive* if $0 < A$ and *non-negative* if $0 \leq A$. If $A < 0$, we say that A is *negative*. As usual, we write $A > B$ to mean $B < A$ etc.

Lemma 3.2. *Let $A = [(a, b)]$. Then A is positive if and only if $b < a$.*

Proof. If $0 < A$, then since $0 = [(a, a)]$, we see that $a + b < a + a$ by definition of the ordering. Using the definition of ordering in $\mathbb{N}$ and cancellation, the result follows. Converse is equally easy. □

Lemma 3.3. *A is positive if and only if $A = f(k)$ for some natural number k.*

Proof. Let $A = [(a, b)]$. First assume that it is positive. Then $b < a$ from the previous lemma and thus $a = b + k$ for some natural number k and thus $A = [(b + k, b)] = f(k)$. Converse is equally easy. □

Lemma 3.4. *$A \in \mathbb{Z}$ is positive if and only if $-A$ is negative.*

Proof. Left as an exercise. □

Let us denote by $\mathbb{N}' = \mathbb{N} \cup \{0\} \subset \mathbb{Z}$, where $\mathbb{N}$ is identified with the subset $f(\mathbb{N})$. We have a map $\mathbb{Z} \to \mathbb{N}'$ called the *absolute value*, denoted by $|\ |$. $|a| = a$ if a is non-negative and $|a| = -a$ if a is negative.

Lemma 3.5. *For any $a \in \mathbb{Z}$, we have $-|a| \leq a \leq |a|$. Converseley, if $c \geq 0$ is an integer and $-c \leq a \leq c$, then $|a| \leq c$*

Proof. First assume that $a \geq 0$. Then $|a| = a$ by defintion. So, the inequlaity we need is $-a \leq a \leq a$. Since $a \leq a$ by definition, we only need to verify that $-a \leq a$. But this is same as $2a \geq 0$, by adding a to both sides of the inequality and this is clear. The case of $a < 0$ is equally easy.

To prove the second statement, we work exactly as before by looking at the two cases of $a \geq 0$ and $a < 0$. If $a \geq 0$, then $a = |a|$ and since $a \leq c$, we get that $|a| \leq c$. If $a < 0$, then $|a| = -a$. Since $-c \leq a$, we get that $c \geq -a = |a|$, proving the result. □

Lemma 3.6 (Triangle inequality)**.** *If $a, b \in \mathbb{Z}$ then $|a + b| \leq |a| + |b|$.*

Proof. From the previous lemma, we have $-|a| \leq a \leq |a|$ and $-|b| \leq b \leq |b|$. Adding these, we get,

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

Since $|a| + |b| \geq 0$, by the previous lemma, we are done. □

Finally now we can define a new operation *subtraction* denoted by $-$ in $\mathbb{Z}$ as follows. If $a, b \in \mathbb{Z}$, then $a - b = a + (-b)$. I will leave the standard properties of subtraction to be verified by the reader. Let me close this section by yet another form of induction.

Theorem 3.1 (Fourth alternate form of Induction)**.** *Let $S \subset \mathbb{Z}$ be a non-empty subset of the integers with the property that there is an $m \in \mathbb{Z}$ such that for all $a \in S$, $a > m$. (That is to say that the set S is bounded below). Then S has a least element.*

Proof. Consider the set $T = \{a - m | a \in S\}$. Since $S \neq \emptyset$ nor is T. Since $a > m$, $a - m > 0$ and hence $T \subset \mathbb{N}$. Thus, by Theorem 1.2, T has a minimal element, say p. Then $q = p + m \in S$ and we claim that q is the least element of S. If $a \in S$, then $a - m \in T$ and hence $a - m \geq p$ and thus $a \geq p + m = q$. Thus by definition of least element, q is the least element of S. □

Exercise 3. Show that if $a, b \in \mathbb{Z}$, then $a \cdot (-b) = -(ab)$.

4. Some Number Theory

We will not use most of what we prove in this section in the sequel. I have included it only to connect our discussions with facts familiar to you and thus giving you some bearing in this abstract jungle. We start with one of the earliest results you might have studied in school.

Theorem 4.1 (Division Algorithm)**.** *Let $a, d \in \mathbb{Z}$ with $d \neq 0$. Then there exists unique integers $q, r \in \mathbb{Z}$ such that $a = qd + r$ and $0 \leq r < |d|$.*

Proof. Since $d \neq 0$, d is either positive or negative. We first treat the case $d > 0$. Consider the set $S = \{a - qd | a - qd \geq 0, q \in \mathbb{Z}\}$. First I claim that $a + |a|d \geq 0$, showing that $S \neq \emptyset$. If $a = 0$, then $a + |a|d = 0$, if $a > 0$, then $a + |a|d = a + ad = a(1 + d) > 0$ and if $a < 0$, then $a + |a|d = a - ad = a(1 - d) \geq 0$, since $a < 0$ and $1 - d \leq 0$. Since S is bounded below by zero, we can apply Theorem 3.1 to conclude that S has a minimal element, say $0 \leq r = a - qd$ for some $q \in \mathbb{Z}$. So, $a = qd + r$ and if we show that $r < d$, we would have proved the existence part of the theorem. If $r \geq d$, then $a - (q + 1)d = r - d \geq 0$. Thus $r - d \in S$ and since $r - d < r$, this contradicts the minimality of r. So, $r < d$.

Next, we prove uniqueness. If $a = qd + r = q'd + r'$ with $0 \leq r, r' < d$, we get $(q - q')d = r' - r$. If $q = q'$, then this implies $r = r'$, proving uniqueness. If $q \neq q'$, then taking absolute values, we get, $d \leq |(q - q')d| = |r - r'|$. But since $0 \leq r, r' < d$, we see that $|r - r'| < d$, which leads to a contradiction. Thus uniqueness is proved.

Finally, if $d < 0$, let $e = -d = |d| > 0$. Thus by the previous part, we have $a = qe + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < e = |d|$. So, $a = (-q)d + r$ as desired. □

As usual if $0 \neq d \in \mathbb{Z}$, and $a \in \mathbb{Z}$, we say that d *divides* a if $a = md$ for some $m \in \mathbb{Z}$. Symbolically, this is written as $d \mid a$.

Definition 5. *Let $a, b \in \mathbb{Z}$ with at least one of them non-zero. Then the greatest common divisor of a, b, written* $\gcd(a, b)$ *is a number $d \in \mathbb{N}$ satisfying the following two properties.*

(1) *$d \mid a$ and $d \mid b$.*
(2) *If $e \in \mathbb{N}$ divides both a and b, then $e \mid d$.*

Notice that gcd is defined only for two numbers with at least one of them non-zero. Also, notice that it is a natural number. What is not clear from the definition is whether such a number exists and if it exists whether it is unique and these we proceed to prove.

Lemma 4.1. *Let $a, b \in \mathbb{Z}$ with at least one of them non-zero. If* $\gcd(a, b)$ *exists, then it is unique.*

Proof. Let $d = \gcd(a, b)$ and $e = \gcd(a, b)$. We wish to show that $d = e$. By first property in the definition applied to e we get that $e \mid a, e \mid b$. Now applying the second property, we see that $e \mid d$. Reversing the roles, we see that $d \mid e$. It is easy to see then $d = e$, though let me give an explicit proof.

$e \mid d$ means we can write $d = pe$ with $p \in \mathbb{Z}$. Since $d, e \in \mathbb{N}$, this forces $p > 0$ and hence $p \in \mathbb{N}$. Similalry we get $e = qd$ with $q \in \mathbb{N}$.

Substituting, we get, $d = pqd$. Cancelling $d > 0$ (which we have shown earlier), we get $1 = pq$. If we show that this implies $p = q = 1$, we would be done. If $p \neq 1$, we could write $p = s + 1$, $s \in \mathbb{N}$ from earlier properties of natural numbers. Thus we get $qs + q = 1$. If $q = 1$, we have $d = e$, so we may also assume that $q > 1$ and thus we can write $q = t+1$. Then $qs+t+1 = \sigma(qs+t) = 1$, but Peano's axioms say that 1 is not in the image of σ, a contradiction, proving what we need. □

Theorem 4.2 (Existence of gcd)**.** *Let $a, b \in \mathbb{Z}$ with at least one of them non-zero. Then* $\gcd(a, b)$ *exists.*

Proof. Consider the set, $S = \{ma + nb > 0 | m, n \in \mathbb{Z}\}$. As usual, we claim that this set is non-empty. Since at least one of a, b is non-zero, we may assume that $a \neq 0$, possibly after renaming them. Then let $m = a, n = 0$. So, $a \cdot a + 0 \cdot b = a^2 > 0$ and hence $a^2 \in S$. So S is non-empty. Since $S \subset \mathbb{N}$, by Theorem 1.2, S has a minimal element, say d. By definition of the set S, we have $m, n \in \mathbb{Z}$ so that $d = ma+nb$. I claim that $d = \gcd(a, b)$.

For this, we need to check the two conditions in the definition of gcd. The second one is easy. If e divides both a, b, then we can write $a = pe, b = qe$ for some $p, q \in \mathbb{Z}$ and substituting, we get, $d = ma+nb = mpe + nqe = (mp + nq)e$ and since $mp + nq \in \mathbb{Z}$, we see that $e \mid d$.

For the first condition, we proceed as follows. First, we show that $d \mid a$, the other case being similar, we shall omit it. By division algorithm, we can write $a = qd + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < d$. If $r = 0$, $d \mid a$, so let us assume that $r > 0$ and then we will arrive at a contradiction. Then $r = a - qd = a - q(ma + nb) = (1 - qm)a + (-qn)b$ and by definition of the set S, we see that $r \in S$. But, $r < d$, contradicting the minimality of d. □

The above proof gives something stronger and it is in this form it is often used, so let us state this explicitly.

Corollary 4.1. *Let $a, b \in \mathbb{Z}$ with at least one of them non-zero. Then* $\gcd(a, b)$ *exists and is unique. Further, it is the smallest natural number of the form $ma + nb$ with $m, n \in \mathbb{Z}$.*

Here is an immediate application.

Corollary 4.2. *Let $p, a, b \in \mathbb{Z}$ and assume that $p \neq 0$. If $p \mid ab$ and* $\gcd(p, a) = 1$*, then $p \mid b$.*

Proof. $\gcd(p, a) = 1$ implies there exists integers m, n such that $1 = mp + na$. Multiplying by b we get, $b = mpb + nab$. Since $p \mid ab$, we can write $ab = sp$ for some $s \in \mathbb{Z}$. Thus we get, $b = mpb+nsp = (mb+ns)p$ and since $mb + ns \in \mathbb{Z}$, we get that $p \mid b$. □

Definition 6. *A natural number p is a prime number if $p \neq 1$ and the only natural numbers dividing it are 1 or p. An integer n is called a composite number if $|n| \neq 0, 1$ and $|n|$ is not a prime.*

I will not give examples of primes, since I am sure most of you are familiar with them.

Exercise 4.
(1) Show that if p is a prime then for any $a \in \mathbb{N}$, $\gcd(p, a) = 1$ or $\gcd(p, a) = p$.
(2) Show that if p is a prime and $p \mid ab$ for $a, b \in \mathbb{Z}$, then $p \mid a$ or $p \mid b$.
(3) Show that if a prime p divides a prime q, then $p = q$.
(4) Show that if a prime p divides q^n for a prime q and $n \in \mathbb{N}$, then $p = q$.

Theorem 4.3 (Fundamental Theorem of Arithmetic, Part 1)**.** *Let $1 \neq n \in \mathbb{N}$. Then there exists primes, $p_1, p_2, \ldots, p_k$ such that $n = p_1 p_2 \cdots p_k$. That is, any $n \in \mathbb{N}$, $n > 1$ is a product of primes.*

Proof. As usual we start with the set

$$S = \{n \in \mathbb{N} | n > 1, n \text{ is not a product of primes}\}.$$

We wish to show that S is empty and if it is non-empty, we let $n \in S$ be the least element, assured by Theorem 1.2. Now, n can not be a prime, since then n is the product of one prime. Thus by definition of a prime, there exists a natural number d which divides n and $d \neq 1, n$. Then by properties of natural numbers, we get $1 < d < n$. Since $n = de$ for some natural number e, we get that $1 < e < n$. By minimality of n, we get $d, e \notin S$ and since they are not 1, they are product of primes. So, $d = p_1 \cdots p_r$, $e = q_1 \cdots q_s$ for primes p_i, q_j. Thus,

$$n = de = p_1 \cdots p_r \cdot q_1 \cdots q_s$$

and hence n is a product of primes. So, $n \notin S$, leading to the desired contradiction. □

Thus, if $n \in \mathbb{N}$ and $n > 1$, we can write $n = p_1 \cdots p_k$ for primes p_is. Collecting the primes and ordering them, we may assume that there exists primes $p_1 < \cdots < p_m$ and natural numbers $a_1, \ldots, a_m$ such that $n = p_1^{a_1} \cdots p_m^{a_m}$.

Theorem 4.4 (Fundamental Theorem of Arithmetic, Part 2)**.** *Let $n > 1$ be a natural number and let $n = p_1^{a_1} \cdots p_m^{a_m}$ as in the previous paragraph. Then this expression is unique. That is, if $n = q_1^{b_1} \cdots q_l^{b_l}$ with $q_1 < \cdots < q_l$ primes and $b_i \in \mathbb{N}$ for all i, then $m = l$, $p_i = q_i$ and $a_i = b_i$ for all i.*

Proof. Again, let S be the set of natural numbers not equal to 1 and which have two different such decompositions into primes. We wish to show that S is empty and if not pick n, the least element assured by Theorem 1.2. So, we have,

$$n = p_1^{a_1} \cdots p_k^{a_k} = q_1^{b_1} \cdots q_l^{b_l}$$

with p_i, q_j primes, $p_1 < p_2 < \cdots < p_k$, $q_1 < q_2 < \cdots < q_l$, $a_i, b_j \in \mathbb{N}$. Now, p_1 divides n and thus p_1 divides $q_1^{b_1} \cdots q_l^{b_l}$. By exercise 4, we then have $p_1 = q_i$ for some i. Similarly, $q_1 = p_j$ for some j.

We first look at the case $i > 1$. Then $p_1 \leq p_j$ (j may be 1) and since $p_j = q_1 < q_i = p_1$, we get $p_1 < p_1$, a contradiction. So, we see that $i = 1$ and $p_1 = q_1$. Next we claim that $a_1 = b_1$. If not we may assume by well ordering, that $a_1 > b_1$ (or the other way around, but we can interchange p, q. Cancelling $p_1^{b_1}$, we get,

$$p_1^{a_1 - b_1} p_2^{a_2} \cdots p_k^{a_k} = q_2^{b_2} \cdots q_l^{b_l}.$$

But then $p_1 = q_i$ for some $i > 1$ as before, but all these are greater than $q_1 = p_1$, which is impossible. This proves that $a_1 = b_1$. Then we have,

$$m = p_2^{a_2} \cdots p_k^{a_k} = q_2^{b_2} \cdots q_l^{b_l}.$$

Since $m < n$, $m \notin S$ and thus, by minimality of n, the theorem is true for m. This implies, $k = l$, $p_i = q_i$ for $i > 1$ and $a_i = b_i$ for $i > 1$. This says, since $p_1 = q_1$ and $a_1 = b_1$, that the theorem is true for n and hence $n \notin S$, which is the desired contradiction, proving the theorem. □

The following theorem is one of the most ancient and well-studied theorem and proof, noted for its elegance and simplicity. But, the classical proof below has its pitfalls.

Theorem 4.5 (Infinitude of primes)**.** *There are infinitely many primes.*

Proof. We know that there at least some primes, for example $2, 3$. Assume there are only finitley many primes, say $p_1, \ldots, p_k$. Consider the natural number $N = p_1 p_2 \cdots p_k + 1$. Then clearly $N > 1$ and hence by the fundamental theorem, there exists a prime number q which divides N. But, since the p_is are all the primes, $q = p_i$ for some i. Thus p_i divides N. Then p_i divides $N - p_1 \cdots p_i \cdots p_k = 1$ and no prime number can divide 1, leading to a contradiction. This proves the theorem. □

5. Rational numbers

The idea is the same. So, I will briefly sketch the construction. Now consider the set $S = \mathbb{Z} \times (\mathbb{Z} - \{0\})$ and put a relation as follows. $(a, b) \sim (c, d)$ if $ad = bc$. One easily checks that this is indeed an

equivalence relation on S and the set of equivalence classes is called the rational numbers and denoted by $\mathbb{Q}$. As usual, addition is defined by $[(a,b)] \oplus [(c,d)] = [(ad+bc, bd)]$ and multiplication by $[(a,b)] \otimes [(c,d)] = [(ac,bd)]$, where now we have addition and multiplication of integers inside the brackets. As usuaul, we check that this is well defined and all the standard properties. We also have a one-one map $f : \mathbb{Z} \to \mathbb{Q}$ given by $f(a) = [(a,1)]$ and then $f(a+b) = f(a) \oplus f(b)$ and $f(ab) = f(a) \otimes f(b)$. Intuitively, we are thinking of $[(a,b)]$ as a/b. Using f, we can identify $\mathbb{Z}$ as a subset of $\mathbb{Q}$ in the usual way. Thus again, we can drop $\oplus, \otimes$ and write the usual symbols for addition and multiplication.

In $\mathbb{Q}$, we have a new operation, *division*, as usual denoted by a/b for $a, b \in \mathbb{Q}$ and $b \neq 0$. This is defined as follows: If $A = [(a,b)]$ and $B = [(c,d)]$, with $a, b, c, d \in \mathbb{Z}$, $b \neq 0 \neq d$, then $B \neq 0$ implies that $c \neq 0$. Define $A/B = [(ad, bc)]$ and make sure that this is well defined. Also notice that since both b and c are not zero, $bc \neq 0$.

Lemma 5.1. *If A is any rational number, then $A = [(a,b)]$, for some $a, b \in \mathbb{Z}$ with b positive.*

Proof. By defintion, $A = [(a,b)]$ with $b \neq 0$, $a, b \in \mathbb{Z}$. If b is positive, then we are done. If not, we know that $-b$ is positive. One easily sees that $A = [(-a, -b)]$ and thus we are done. □

We introduce an order on $\mathbb{Q}$ as follows. If $A, B \in \mathbb{Q}$, write $A = [(a,b)]$ and $B = [(c,d)]$ with b, d both positive. Then we define, $A < B$ if $ad < bc$. One checks that this is well defined and has all the usual properties. As always, write $A \leq B$ to mean either $A = B$ or $A < B$. Absolute value of a rational number can be defined as before, after defining what is positive, negative etc. Again, I will assume that we can prove all the usual properties of ordering on $\mathbb{Q}$.

Exercise 5.
(1) Show that if $a < b$ are two rational numbers, then there exists $c \in \mathbb{Q}$ with $a < c < b$.
(2) Prove triangle inequality (see lemma 3.6) for $\mathbb{Q}$.

6. Real Numbers

Let me start with some observations. We started with natural numbers which had this very important property of Induction. In other words, any non-empty subset had a least element. But the system lacked operations like subtraction and thus we were forced to enlarge the system to integers, which had subtraction, and at least a suitable form of induction; namely, any non-empty subset which is *bounded below* had a minimal element. But integers still lacked division and thus

we enlarged our system to the rational numbers to rectify this lacuna. But, now, alas we have lost any semblance of the minimal element property which was so important and desirable. In other words, there are now subsets of $\mathbb{Q}$ which are bounded below with no minimal element in sight. (Not just in the set, but not even in $\mathbb{Q}$). So, the plan of constructing real numbers is to rectify this important problem with $\mathbb{Q}$.

Let me elaborate the above statement. Let

$$S = \{x \in \mathbb{Q} \mid x^2 > 2 \quad \text{and } x \text{ is positive}\}.$$

Clearly this set is non-empty and bounded below. For example, $2 \in S$ and if $x \in S$, then $x > 1$. But I claim that this set has no minimal element even in $\mathbb{Q}$. Assume $a \in \mathbb{Q}$ is such a minimal element. Then a has two properties. First, $a \leq x$ for all $x \in S$. Second, if $b \in \mathbb{Q}$, with $b \leq x$ for all $x \in S$, then $b \leq a$. I claim no such rational number exists. Since 1 works as a b, we see immediately that $a \geq 1$. In particular a must be positive. I claim that $a^2 \leq 2$. If not, $a^2 > 2$ and $a \in S$. So $a^2 = 2 + r$ for some positive rational number. We can choose a large natural number N so that $2a/N < r$. Consider $c = a - (1/N) < a$. Then

$$c^2 = a^2 - 2\frac{a}{N} + \frac{1}{N^2} = 2 + (r - 2\frac{a}{N}) + \frac{1}{N^2} > 2.$$

Thus $c \in S$ and $c < a$. This is a contradiction. Thus $a^2 \leq 2$.

Next I leave it as a (non-trivial) exercise that there is no rational number with $a^2 = 2$ and thus we must have $a^2 < 2$. Then again write $2 = a^2 + r$ with r a positive rational number. As before we can choose a large natural number N such that $rN - 2a \geq 1$ and thus $2a/N + 1/N^2 < r$. Now consider $b = a + (1/N) > a$. One easily checks that $b^2 < 2$ and thus $b \leq x$ for $x \in S$. This again is a contradiction.

Thus, though rational numbers had several of the arithmetic properties for numbers that we desire, it lacks a very important property necessary for Mathematics. In real life, we rarely have to deal with a real number which is not rational in some strict sense, considering, any non-rational number is usually approximated to a rational number, like all the numbers you may get out of a calculator or computer. For Mathematics, approximation of this kind is never sufficient, if you want to be rigorous and precise. So, this is some justification for the construction of real numbers.

6.1. **Dedekind cuts.** One way to construct real numbers starting with rational classically is to use what is known as *Dedekind cuts*, named for an important mathematician Richard Dedekind, who invented them. We will later construct reals by Cauchy sequences, which, while harder, more useful.

So we start with the definition.

Definition 7. *A proper non-empty subset $A \subset \mathbb{Q}$ is called a Dedekind cut, abbreviated DC, if it satisfies the following.*

(1) *If $a \in A$ and $b \leq a$, then $b \in A$.*
(2) *If $a \in A$, there exists $b \in A$ with $a < b$.*

Example 1. (1) Let $a \in \mathbb{Q}$. Then $A = \{b \in \mathbb{Q} | b < a\}$ is a DC.
(2) The set $A = \{a \in \mathbb{Q} | a \leq 0 \text{ or } a^2 < 2\}$ is a DC.

Lemma 6.1. *If A, B are DC, then either $A \subset B$ or $B \subset A$.*

Proof. Assume that A is not contained in B. So, there exists an $a \in A$ and $a \notin B$. If $b \in B$, therefore, $b < a$, since if $b \geq a$, then $a \in B$ by the definition of DC. But then $b \in A$ again by definition and thus $B \subset A$. □

Lemma 6.2. *If A, B are DC, then so is $A+B = \{a+b | a \in A, b \in B\}$.*

Proof. Since □

6.2. **Cauchy Sequences.** Now we plunge into this construction. As you would expect, real numbers are got by *approximating* rational numbers. So we make a couple of definitions.

Definition 8. *A* sequence *of rational numbers is a set map $f : \mathbb{N} \to \mathbb{Q}$. In other words, we are given rational numbers x_n for every natural number n. This is usually abbreviated by the notation $\{x_n\}$.*

Definition 9. *A sequence $\{x_n\}$ is a* Cauchy Sequence *(we will abbreviate it by writing CS), if given* any *$0 < \epsilon \in \mathbb{Q}$, there exists a natural number N (which is allowed to depend on ϵ) such that for all $n, m \geq N$, natural numbers, we have $|x_n - x_m| < \epsilon$.*

I suggest that you mull over this important definition. We will give a few examples below. Typically, given a sequence of rational numbers $\{x_n\}$, to show that it is *not* a CS, we will have to exhibit one positive rational number ϵ such that for any $N \in \mathbb{N}$, there exists $n, m \geq N$ with $|x_n - x_m| \geq \epsilon$. On the other hand, if we wished to show that the sequence is a CS, we must show that for *any* positive rational number ϵ, there exists an $N \in \mathbb{N}$ and for any $n, m \geq N$, $|x_n - x_m| < \epsilon$.

To warm yourself to the concept of CS, here is are some easy exercises.

Exercise 6. (1) If a is any rational number, show that the sequence defined as $x_n = a$ for all $n \in \mathbb{N}$ is a CS.

(2) Let $g : \mathbb{N} \to \mathbb{N}$ be an increasing function. That is, if $n > m$ then $g(n) > g(m)$. If $\{x_n\}$ is a Cauchy sequence, show that $\{y_n\}$ is a CS, where $y_n = x_{g(n)}$. ($\{y_n\}$ is called a *sub-sequence* of $\{x_n\}$).
(3) If $\{x_n\}$ is a CS of rational numbers, show that for any rational number a, $\{ax_n\}$ is a CS.

Example 2. (1) Let $x_n = n$ for all $n \in \mathbb{N}$. Then this is not a CS.

As I said in the previous paragraph, we need to find just one positive rational number ϵ which violates the CS condition. For this take $\epsilon = 1$. If an N existed, then we must have $|x_n - x_m| < \epsilon = 1$ for all $n, m \geq N$. But if we take $n = N$ and $m = N + 1$, clearly we get a contradiction. (How did I decide to take ϵ to be 1? Usually, one works backwards and analyzes what one needs. Sometimes this can be tricky.)

(2) Let $x_n = 2^{-n}$ for all $n \in \mathbb{N}$. Then $\{x_n\}$ is a CS.

Here, we are not allowed to pick an ϵ. We must take any positive rational number ϵ and figure out an appropriate N guaranteeing the CS condition. So, let $\epsilon > 0$ be given. (Here again, to figure out the N, one may have to work backwards and can be tricky.) Let us do the analysis for once.

We need an N so that for any $n, m \geq N$, we must have,

$$|x_n - x_m| = |2^{-n} - 2^{-m}| < \epsilon.$$

Without loss of generality, we may assume that $n \geq m \geq N$. Then $|2^{-n} - 2^{-m}| = 2^{-m}|2^{m-n} - 1|$. Since $0 < 2^{m-n} \leq 1$, because $n \geq m$, we see that $|2^{m-n} - 1| < 1$ and thus we see that $|2^{-n} - 2^{-m}| < 2^{-m}$. But, since $m \geq N$, we also have $2^{-m} \leq 2^{-N}$. So, if we had $2^{-N} < \epsilon$, then we would be done. Notice that we finally ended up with a condition on N and not on which $n, m \geq N$ we need to look at. This is usually the essence of most such arguments.

So, we see that we only need to find an N such that $2^{-N} < \epsilon$, for the given positive ϵ. Since $\epsilon \neq 0$, we have $\delta = \epsilon^{-1} \in \mathbb{Q}$ and since ϵ is positive so is δ. So, we want an N so that $2^N > \delta$. If we write $\delta = a/b$ with a, b positive integers, then it is clear that $a \geq \delta$. So, if we found an N so that $2^N > a$, we would be done. But this is essentially the content of lemma 1.13.

So if you want to write a proof, we invert the steps, so that it is easy to follow the arguments. Let me do this now for illustration.

Proof. We wish to show that the sequence $\{x_n\} = \{2^{-n}\}$ is a CS.

So, let $\epsilon > 0$ be a positive rational number. Then $\delta = \epsilon^{-1} \in \mathbb{Q}$ is a positive rational number. Writing $\delta = a/b$ for positive integers a, b, we see immediatley that $a \geq \delta$. By lemma 1.13, since $2 > 1$, there exists an N such that

$$2^N \geq a \geq \delta \tag{2}$$

Then we claim that for any $n, m \geq N$, (for the N from the equation above) $|x_n - x_m| < \epsilon$, which will prove that our sequence is a CS.

We may assume that $n \geq m \geq N$. Then

$$|x_n - x_m| = |2^{-n} - 2^{-m}| = 2^{-m}|2^{m-n} - 1|.$$

Since $n \geq m$, we have $0 < 2^{m-n} \leq 1$ and thus, $|2^{m-n} - 1| < 1$. So, we get $|x_n - x_m| < 2^{-m}$ if $n \geq m \geq N$. Since $m \geq N$, we have $2^{-m} \leq 2^{-N}$ and thus we get,

$$|x_n - x_m| < 2^{-m} \leq 2^{-N} \leq \delta^{-1} = \epsilon,$$

where the last inequlaity follows from equation 2 and this is valid for $n \geq m \geq N$. This is what we set out to prove and thus the sequence $\{x_n\}$ is a CS. □

(3) For any natural number n, we have $2n^2 < 2n^2 + 1$. Thus by exercise (5), we can choose a rational number y_n such that $2n^2 < y_n < 2n^2 + 1$. Then the sequence $\{x_n = y_n n^{-2}\}$ is a CS.

Given $\epsilon > 0$, easy to see that there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, we have $0 < n^{-2} < \epsilon$. If $n \geq m \geq N$, then we have,

$$|x_n - x_m| < 2 + m^{-2} - 2 = m^{-2} < \epsilon$$

(4) Let

$$x_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

Then $\{x_n\}$ is not a CS. This is left as a not too easy exercise. Or may be look it up in your calculus text.

6.3. **Construction of Real Numbers.** Next let S be the set of all Cauchy Sequences in $\mathbb{Q}$. We put a relation on S as follows: $\{x_n\} \sim \{y_n\}$, if given any $0 < \epsilon \in \mathbb{Q}$, there exists an $N \in \mathbb{N}$ such that for all $n, m \geq N$, we have $|x_n - y_m| < \epsilon$.

This is indeed an equivalence relation. Reflexivity is just the fact that the sequences are Cauchy. Symmetry is obvious since $|x - y| = |y - x|$. Transitivity follows from triangle inequality (see Exercise 5).

Thus we consider the set $\mathbb{R}$ as the set of equivalence classes and call it the real numbers.

Before we introduce the operations on $\mathbb{R}$, let us prove a few things about CS.

Lemma 6.3. *Let $\{x_n\}$ be a CS. Then there exists a rational number $M \in \mathbb{Q}$, such that $|x_n| < M$ for all n. That is to say, the sequence is* **bounded**.

Proof. Pick $\epsilon = 1$. So we have an $N \in \mathbb{N}$ satisfying the properties of the definition of a CS. Let

$$M = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, |x_N|\} + 1.$$

By choice, $|x_n| < M$ if $n \leq N$. If $n > N$, then by choice, we have, $|x_n - x_N| < 1$ and then using triangle inequality (see Exercise 5), we are done.

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N| \leq M$$

□

As usual we define the operations as follows. If $A = [\{x_n\}]$ and $B = [\{y_n\}]$, define $A \oplus B = [\{x_n + y_n\}]$ and $A \otimes B = [\{x_n y_n\}]$. Let us check that these make sense.

Lemma 6.4. *If $\{x_n\}$ and $\{y_n\}$ are CS, then so are $\{x_n + y_n\}$ and $\{x_n y_n\}$.*

Proof. Given ϵ, we can find $N_1, N_2 \in \mathbb{N}$ such that for all $n, m \geq N_1$ we have $|x_n - x_m| < \epsilon/2$ and for all $p.q \geq N_2$, $y_p - y_q| < \epsilon/2$. Take $N = \max\{N_1, N_2\}$ and use the triangle inequality to show that $\{x_n + y_n\}$ is Cauchy.

For the multiplication, we use lemma 6.3 and thus there exists $M_1, M_2 \in \mathbb{Q}$ such that $|x_n| < M_1$ for all n and $|y_n| < M_2$ for all n. Clearly, we may replace M_1, M_2 by $M = \max\{M_1, M_2\}$. Given $\epsilon > 0$, consider $\delta = \epsilon/2M > 0$. As before, there exists N_1, N_2 such that for all $n, m \geq N_1$, we have $|x_n - x_m| < \delta$ etc. Again we may repalce N_1, N_2 by $N = \max\{N_1, N_2\}$. Thus we get,

$$\begin{aligned} |x_n y_n - x_m y_m| &= |(x_n - x_m)y_n + x_m(y_n - y_m)| \\ &\leq |(x_n - x_m)y_n| + |x_m(y_n - y_m)| \end{aligned}$$

by triangle inequality. If $n, m \geq N$, this in turn gives

$$|x_n - x_m||y_n| + |x_m||y_n - y_m| < \delta M + M\delta = \epsilon$$

proving that our sequence is Cauchy.

□

Now, as usual, I will leave it to you to check that this definition of addition and multiplication are well defined. (Please try to check this– it is a good exercise). I will check that addition is well defined below, just for illustration.

Lemma 6.5. *Addition of real numbers as defined above is well defined.*

Proof. Our definition was if $A = [\{x_n\}]$ and $B = [\{y_n\}]$, two real numbers, then $A \oplus B = [\{x_n + y_n\}]$ and we have checked in the previous lemma that $\{x_n + y_n\}$ is a CS. If $A = [\{t_n\}]$ and $B = [\{u_n\}]$, then $A \oplus B$ is defined as $[\{t_n + u_n\}]$. So, we must check that $[\{x_n + y_n\}] = [\{t_n + u_n\}]$. That is $\{x_n + y_n\} \sim \{t_n + u_n\}$. So, we must check that given $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n, m \geq N$, $|x_n + y_n - t_m - u_m| < \epsilon$.

Since $[\{x_n\}] = [\{t_n\}]$, by our definition $\{x_n\} \sim \{t_n\}$ and hence we can find an $N_1 \in \mathbb{N}$ so that for all $n, m \geq N_1$, $|x_n - t_m| < \epsilon/2$. Similarly we can find an $N_2 \in \mathbb{N}$ so that for all $n, m \geq N_2$, $|y_n - u_m| < \epsilon/2$. Let $N = \max\{N_1, N_2\}$ and then for any $n, m \geq N$,

$$|x_n + y_n - t_m - u_m| \leq |x_n - t_m| + |y_n - u_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

the middle inequlity being the triangle inequality. This proves what we set out to prove. □

We also have a map $f : \mathbb{Q} \to \mathbb{R}$, sending $a \in \mathbb{Q}$ to the equivalence class of the CS, $x_n = a$ for all n (That this is indeed a CS was checked in Exercise 6). The function f is one-one and thus we can identify $\mathbb{Q}$ as a subset of $\mathbb{R}$, $f(a + b) = f(a) \oplus f(b)$ and $f(ab) = f(a) \otimes f(b)$. Thus now, we may drop our complicated notation and write the usual symbols for addition and multiplication. Also, we can check all the standard properties of this operation. We also define an ordering on $\mathbb{R}$ as follows. If $\{x_n\}$ and $\{y_n\}$ are CS, and $A = [\{x_n\}]$ and $B = [\{y_n\}]$, then $A < B$, if there exists an $\epsilon > 0$, rational number and an $N \in \mathbb{N}$ such that for all $n, m \geq N$, $y_m - x_n > \epsilon$. Again make sure that this is well defined. If we define $A \leq B$ as usual, by saying that $A < B$ or $A = B$, then one can easily check the following. $A = \{x_n\} \leq B = \{y_n\}$ if and only if for *any* positive rational number ϵ, there exists an N such that for all $n, m \geq N$, $y_n - x_m > -\epsilon$.

Exercise 7. Show that if $\{x_n\}$ is CS of rational numbers and $\{y_n\}$ is a subsequence of $\{x_n\}$ (which I have noted also must be a CS), then show that $\{x_n\} \sim \{y_n\}$.

6.4. **Supremum and Infimum.** Now we have all the machinery required to prove the following important properties of the real numbers. The proofs, while elementary, are fairly subtle. You may try it for fun.

If you do not succed, do not be discouraged. Again, as a warm up exercise, let us prove an easy lemma.

Lemma 6.6. *If $A < B$ are two real numbers, there is a rational number q, such that $A < q < B$. Remember that a rational number q as an element of $\mathbb{R}$ is just the equivalence class of the Cauchy sequence with all terms equal to q.*

Proof. Write $A = [\{x_n\}]$ and $B = [\{y_n\}]$. Then by definition, $A < B$ means, there exists a positive rational number ϵ and an $N_1 \in \mathbb{N}$ such that for all $n, m \geq N_1$, $y_n - x_m > \epsilon$. Since $\{x_n\}$ is a CS, there exists an N_2 so that for all $n, m \geq N_2$, we have $|x_n - x_m| < \epsilon/2$. Similarly, there exists $N_3 \in \mathbb{N}$ so that for all $n, m \geq N_3$, $|y_n - y_m| < \epsilon/2$. If we choose $N = \max\{N_1, N_2, N_3\}$, then for all $n, m \geq N$, all the above three inequlaities hold. Now, let $q = (x_N + y_N)/2$. I claim, this rational number is between A and B. We will show that $q < B$, and the inequality $A < q$ will be similar.

For $n \geq N$, we should compute $y_n - q$.

$$\begin{aligned} y_n - q &= y_n - \frac{x_N + y_N}{2} \\ &= \frac{y_n - x_N}{2} + \frac{y_n - y_N}{2} \end{aligned}$$

Since $n, N \geq N \geq N_1$, we have $y_n - x_N > \epsilon$. On the other hand, since $n, N \geq N \geq N_3$, we have $|y_n - y_N| < \epsilon/2$. This implies by lemma 3.5, $y_n - y_N > -\epsilon/2$. Substituting these in the above, we get,

$$y_n - q > \frac{\epsilon}{2} - \frac{\epsilon}{4} = \frac{\epsilon}{4}$$

This shows that $q < B$. Similarly one shows that $A < q$. □

Definition 10. *Let $S \subset \mathbb{R}$ and $a \in \mathbb{R}$. Then a is called a lower bound for S if $a \leq s$ for all $s \in S$. Similarly an element $b \in \mathbb{R}$ is called an upper bound for S if $s \leq b$ for all $s \in S$.*

Now, we prove the most important property of real numbers, from which all the other subtle properties can be deduced. The proof is long and so take your time mulling over the steps.

Theorem 6.1. *Let $S \subset \mathbb{R}$ be a non-empty subset and assume that it has a lower bound M. Then there exists a real number ξ (called the infimum of S) such that for any $s \in S$, $s \geq \xi$ and if $x \in \mathbb{R}$ is such that $s \geq x$ for all $s \in S$, then $\xi \geq x$.*

Proof. First, notice that we may assume the M that we have can be assumed to be rational, by choosing a smaller number by the previous

lemma. Our aim is to construct ξ as in the theorem and by now we should have the feeling that in general it is going to be a real number, but not a rational number. Thus to get there, we must construct an approriate Cauchy sequence, which will be a real number using our relation. Let us during the proof, call a number $a \in \mathbb{R}$ a *lower bound* for S if $a \leq s$ for all $s \in S$. So M is a lower bound. Now pick any $s \in S$. (This is where we use the fact that S is not empty). Again, easy to see that we can pick a rational number $N > s$. Let $q = N - M \in \mathbb{Q}$. Call $M_1 = M$. Consider the rational number $M + (q/2)$. Then there are two possibilities. Either this number is a lower bound for S or not. If it is, call $M_2 = M + (q/2)$. If it is not call $M_2 = M_1$. Let us see what we have.

By choice, M_2 is still a lower bound for S, $M_1 \leq M_2$, $M_2 - M_1 \leq q/2$ and there exists an $s \in S$ such that $s - M_2 < q/2$.

Now we repeat the process. That is consider $M_2 + (q/4)$. Then again there are two possibilities. Either it is a lower bound for S or not. If it is call $M_3 = M_2 + (q/4)$ or else call $M_3 = M_2$. Again notice that M_3 is a lower bound for S, $M_2 \leq M_3$, $M_3 - M_2 \leq q/4$ and there exists an $s \in S$ with $s - M_3 < q/4$. We continue this process by replacing $q/4$ with $q/8, q/16$ etc. to get a sequence $\{M_n\}$.

Next we check that this is indeed a Cauchy sequence and $\xi = [\{M_n\}]$ is an infimum for S. Let us repeat the basic properties of this sequence of rational numbers.

$$\begin{gathered} M_n \text{ is a lower bound for } S \, \forall n \in \mathbb{N} \\ M_1 \leq M_2 \leq \cdots \leq M_n \leq M_{n+1} \leq \cdots \\ M_{n+1} - M_n \leq \frac{q}{2^n} \quad \forall n \in \mathbb{N} \\ \exists s_n \in S \text{ such that } s_n - M_n < \frac{q}{2^{n-1}}, \quad \forall n \in \mathbb{N} \end{gathered} \tag{3}$$

If one wants to be very precise, this is how the above should be phrased. We would like to construct recursively M_n for $n \in \mathbb{N}$ satisfying the above properties. We are given a lower bound M which we call M_1. Since $S \neq \emptyset$, pick an element $s \in S$ and let N be a rational number such that $N > s$. Let $q = N - M$ a positive rational number. Then we construct M_2 as described above, satisfying the properties in equation 3 for $n = 1, 2$. So assume that we have constructed $M_1, M_2, \ldots, M_n$ satisfying the above properties. We wish to construct an M_{n+1} satisfying the above. So, we consider $M_n + \frac{q}{2^n}$. If this number is a lower bound for S, we call this M_{n+1}. Otherwise we let M_{n+1} to be the same as M_n. So, by choice, we still have M_{n+1} a lower bound for S and we

also have $M_n \leq M_{n+1}$. Since $M_{n+1} - M_n = 0$ or $q/2^n$, we also have $M_{n+1} - M_n \leq q/2^n$. So, we only need to verify the last requirement.

For this again we look at the two cases when $M_{n+1} = M_n + \frac{q}{2^n}$ or M_n. In the first case, we take $s_{n+1} = s_n$. Then

$$s_{n+1} - M_{n+1} = s_n - M_n - \frac{q}{2^n} < \frac{q}{2^{n-1}} - \frac{q}{2^n} = \frac{q}{2^n},$$

which is what we want. In the case, $M_{n+1} = M_n$, we know that $M_n + \frac{q}{2^n}$ is not a lower bound for S and thus there exists an $s_{n+1} \in S$ such that $s_{n+1} < M_n + \frac{q}{2^n}$. Since M_n is a lower bound, we have,

$$M_n \leq s_{n+1} < M_n + \frac{q}{2^n}.$$

Subtracting M_n, we get $s_{n+1} - M_n < \frac{q}{2^n}$, which is what we want. Thus recursively we can define the sequence M_n, satisfying the properties stated above.

If $n \geq m$, we have,

$$\begin{aligned} |M_n - M_m| &= M_n - M_m \\ &= (M_n - M_{n-1}) + (M_{n-1} - M_{n-2}) + \cdots + (M_{m+1} - M_m) \\ &\leq \frac{q}{2^{n-1}} + \frac{q}{2^{n-2}} + \cdots + \frac{q}{2^m} \qquad (4) \\ &= \frac{q}{2^m}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-m-2}} + \frac{1}{2^{n-m-1}}\right) \\ &< \frac{q}{2^m} \cdot 2 = \frac{q}{2^{m-1}} \end{aligned}$$

Given $\epsilon > 0$, choose an $N \in \mathbb{N}$ so that $q/2^{N-1} < \epsilon$. This can be done by lemma 1.13. If $n \geq m \geq N$, by equation 4, we get,

$$|M_n - M_m| < \frac{q}{2^{m-1}} \leq \frac{q}{2^{N-1}} < \epsilon.$$

This proves that $\{M_n\}$ is a Cauchy sequence and so $\xi = [\{M_n\}]$ is a real number.

Next, we check that ξ is a lower bound for S. For this, let $s \in S$. and let $\epsilon > 0$ be given. Then we may choose an $N_1 \in \mathbb{N}$ so that $q/2^{N_1-1} < \epsilon/2$ as before. Since M_{N_1} is lower bound for S, we have $M_{N_1} \leq s$. If we write $s = [\{u_n\}]$ for a Cauchy sequence $\{u_n\}$, there exists an $N_2 \in \mathbb{N}$ so that for all $n \geq N_2$, $u_n - M_{N_1} > -\epsilon/2$, by definition of inequality discussed earlier. Now, let $N = \max\{N_1, N_2\}$. If $n, m \geq N$, we get, using equation 4,

$$\begin{aligned} u_n - M_m &= u_n - M_{N_1} + M_{N_1} - M_m \\ &> -\frac{\epsilon}{2} - \frac{q}{2^{N_1-1}} > -\epsilon. \end{aligned}$$

This proves that $\xi \leq s$.

Finally we check that ξ is an infimum for S. So, let x be a lower bound for S. We must show that $x \leq \xi$. We prove this by contradiction. If this is not true, then $x > \xi$ and let us write $x = [\{x_n\}]$ as usual. So, there exists an $\epsilon > 0$ and $N_1 \in \mathbb{N}$ so that for all $n, m \geq N_1$,

$$x_n - M_m > \epsilon. \tag{5}$$

Choose as before an N_2 so that $q/2^{N_2-1} < \epsilon/3$. Then by equation 3, we have an $s \in S$ so that $s - M_{N_2} < q/2^{N_2-1}, \epsilon/3$. Writing this $s = [\{u_n\}]$, there exists an $N_3 \in \mathbb{N}$ so that for all $n \geq N_3$, $u_n - M_{N_2} - \epsilon/3 < \epsilon/3$. We rewrite this as,

$$M_{N_2} - u_n > -\frac{2\epsilon}{3} \tag{6}$$

Now, let $N = \max\{N_1, N_2, N_3\}$. Then for all $n, m \geq N$, we have, using equations 3, 5 and 6,

$$\begin{aligned} x_n - u_m &= (x_n - M_N) + (M_N - M_{N_2}) + (M_{N_2} - u_m) \\ &> \epsilon + 0 - \frac{2\epsilon}{3} = \frac{\epsilon}{3}. \end{aligned}$$

Thus by definition, we get that $x > s$. This is a contradiction, since x was a lower bound and hence $x \leq s$ for all $s \in S$.

□

An identical argument can be used to prove the following,

Theorem 6.2. *Let $S \subset \mathbb{R}$ be a non-empty subset and assume it has an upper bound. Then there exists a real number ζ (called the supremum of S) such that for any $s \in S$, $s \leq \zeta$ and if $x \in \mathbb{R}$ is such that $s \leq x$ for all $s \in S$, then $\zeta \leq x$.*

Now, I repeat the properties of infimum and supremum explicitly, so that we can use it again. These properties are easy, but the existence as proved above is not.

Lemma 6.7. *Let S be a non-empty set bounded below and let ξ be its infimum. Then ξ is a lower bound for S. Further, given any positive number ϵ, there exists an $s \in S$ such that $\xi \leq s < \xi + \epsilon$.*

Similarly, if S is non-empty set bounded above, then its supremum ζ is an upper bound for S. Further given any positive number ϵ there exists an $s \in S$ so that $\zeta - \epsilon < s \leq \zeta$.

Proof. The proof is easy. We already know that the infimum is a lower bound. Now, assume that we are given $\epsilon > 0$. Then by the property of infimum, $\xi + \epsilon$ can not be a lower bound for S. But this means

precisely that there is an $s \in S$ with $s < \xi + \epsilon$. The proof for supremum is identical. □

The phrase *greatest lower bound* (glb for short) is used interchageably with infimum and *least upper bound* (lub for short) for supremum.

Next we define Cauchy Sequences exactly as before, but with real numbers instead of rational numbers.

Definition 11. *A sequence of real numbers $\{A_n\}$ is a CS, if for any positive real number ϵ there exists an $N \in \mathbb{N}$ so that for all $n, m \geq N$, $|A_n - A_m| < \epsilon$.*

Of course, we could have defined a CS of real numbers with ϵ a rational number.

Definition 12. *A sequence of real numbers $\{A_n\}$ is a CS, if for any positive rational number ϵ there exists an $N \in \mathbb{N}$ so that for all $n, m \geq N$, $|A_n - A_m| < \epsilon$.*

Lemma 6.8. *The two defintions above are equivalent. That is, a sequence $\{A_n\}$ of real numbers is a CS in the first definition if and only if it is so in the second definition.*

The proof is an easy application of lemma 6.6.

Theorem 6.3. *Let $\{x_n\}$ be a Cauchy sequence of real numbers. Then there exits a unique real number x such that given any $\epsilon > 0$, real number, there exists an $N \in \mathbb{N}$, such that for all $n \geq N$, $|x_n - x| < \epsilon$.*

The above theorem states that if we repeated the construction of reals from rationals, then we get nothing new. This is usually termed as the *completeness of the real numbers*.

Proof. We first look at non-empty sets, S_i for $i = 0, 1, \ldots$ as follows. Let $S_0 = \{x_1, \ldots, x_n, \ldots\}$, $S_1 = \{x_2, x_3, \ldots, x_n, \ldots\}$ etc. These are sets and not sequences (though they look like sequences). So, S_n would be the set of elements of the sequence starting from x_{n+1}. These are all non-empty and as we have seen in lemma 6.3, these are bounded, since $\{x_n\}$ is a CS. Thus by theorem 6.1, we have ξ_n the infimum of S_n.

First notice that $\xi_n \leq \xi_{n+1}$ for all n. This is because, since ξ_n is a lower bound for the set S_n and $S_{n+1} \subset S_n$, it is clearly a lower bound for S_{n+1} and thus by property of infimum, we must have $\xi_n \leq \xi_{n+1}$. Now, since all the ξ's are bounded above, since $\{x_n\}$ is CS, we have a supremum for the non-empty set, $\{\xi_0, \ldots, \xi_n, \ldots\}$. Let this supremum be x. I claim that this x has the property stated in the theorem.

So, assume that we are given an $\epsilon > 0$ and let $\delta = \epsilon/3$. Then there exists an $N_1 \in \mathbb{N}$ so that for all $n, m \geq N_1$,

$$|x_n - x_m| < \delta. \tag{7}$$

Since x is the supremum of the set $\{\xi_n\}$, by lemma 6.7, there exists an element say ξ_{N_2} so that, $x - \delta < \xi_{N_2} \leq x$. Since $\xi_n \leq \xi_{n+1}$ for all n, we may replace N_1, N_2 by $N = \max\{N_1, N_2\}$. Thus, we have

$$x - \delta < \xi_N \leq x. \tag{8}$$

Finally, since ξ_N is the infimum of the set S_N, by lemma 6.7 again, there exists an element $x_l \in S_N$ so that

$$\xi_N \leq x_l < \xi_N + \delta \tag{9}$$

Notice that since $x_l \in S_N$, $l > N$. Putting these three equations together, we get, for $n \geq N$,

$$\begin{aligned} |x - x_n| &= |x - \xi_N + \xi_N - x_l + x_l - x_n| \\ &\leq |x - \xi_N| + |\xi_N - x_l| + |x_l - x_n| \\ &< \delta + \delta + \delta = \epsilon \end{aligned}$$

Now, we only need to prove that this x is unique. I leave it as an exercise.

□

Definition 13. *If $\{x_n\}$ is a Cauchy sequence of real numbers and x is the unique real number we found in the above theorem, we call x the limit of the sequence $\{x_n\}$ and write $x = \lim x_n$. For a set S, if it has an infimum, we denote it by* $\inf S$ *and similarly, if it has a supremum denote it by* $\sup S$.

Exercise 8.
(1) Let $x_1 \leq x_2 \leq x_3 \leq \cdots$ be a sequence (of real numbers) which is bounded above. Show that $\{x_n\}$ is a CS and $\lim x_n = \sup\{x_1, x_2, x_3, \ldots\}$.
(2) Let $\{x_n\}$ be a CS and assume that m is a lower bound for the set $\{x_n\}$ and M an upper bound. Then show that $m \leq \lim x_n \leq M$.

Theorem 6.4. *Let S be an infinite bounded set. That is, S is infinite and there exists an $M > 0$ such that for all $s \in S$, $|s| \leq M$. Then there exists an $x \in \mathbb{R}$ such that, for any $\epsilon > 0$, there are infinitely many elements $s \in S$ such that $|x - s| < \epsilon$.*

The proof is on similar lines as above and I will not prove it.

By now, it should be clear to you, that while the arguments are not difficult, they can be rather tedious. These arguments repeat themselves in several guises.

7. Continuous functions

In this section we study continuous functions from $\mathbb{R}$ to $\mathbb{R}$. You should do these for functions defined in an open interval, but I will not go into it in any detail since it is mostly routine. We start with the definition of a continuous function at a point $x \in \mathbb{R}$.

Definition 14. *A function $f : \mathbb{R} \to \mathbb{R}$ is called continuous at a point $x \in \mathbb{R}$ if for any CS, $\{x_n\}$, with $\lim x_n = x$, the sequence $\{f(x_n)\}$ is a CS. A function is continuous on $\mathbb{R}$ (or any open interval) if it is continuous at every point of $\mathbb{R}$ (respectively, at every point of the open interval).*

As usual, we will write $[a, b]$ for $a < b$, the closed interval, which is the set of all $x \in \mathbb{R}$ such that $a \leq x \leq b$. Similarly, open and half-closed intervals can be defined.

Before we give the usual examples of continuous functions, let me prove some easy lemmas.

Lemma 7.1. *If f, g are continuous functions on $\mathbb{R}$, so is $f+g$ and fg, where $(f+g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for all $x \in \mathbb{R}$.*

Proof is just an application of lemma 6.4.

Lemma 7.2. *If f is continuous at x and $\{x_n\}$ is any CS with $\lim x_n = x$, then $\lim f(x_n) = f(x)$.*

Proof. Consider a new sequence $\{y_n\}$ defined as follows. For $n = 2m - 1$, define $y_n = x_m$ and for $n = 2m$ define $y_n = x$. It is easy to check that then $\{y_n\}$ is a CS and $\lim y_n = x$. So, by continuity, $\{f(y_n)\}$ is a CS. Let $y = \lim f(y_n)$. So, given any $\epsilon > 0$, there exists an N such that for $n \geq N$, we have $|f(y_n) - y| < \epsilon$. But, if $n \geq N$ and n is even, since $y_n = x$, we get $|f(x) - y| < \epsilon$. But, since ϵ was arbitrary, this can happen only if $y = f(x)$. Now, taking $2m - 1 = n \geq N$, one has $y_n = x_m$ and so we get $|f(x_m) - f(x)| < \epsilon$, which implies that $\lim f(x_n) = f(x)$.

□

Example 3.
(1) The identity function $\mathrm{Id} : \mathbb{R} \to \mathbb{R}$ is continuous. That is, the function $\mathrm{Id}(x) = x$ is continuous.
(2) By the lemma above, we see that the function $f(x) = x^n$ is continuous for any $n \in \mathbb{N}$.
(3) Polynomial functions are continuous. These are functions of the form, $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$, where $n \in \mathbb{N}$ and $a_i \in \mathbb{R}$.

Next let me prove a fundamental theorem about continuous functions, which is frequently used as the definition, often called the ϵ-δ definition of continuity.

Theorem 7.1. *Let f be a function (from $\mathbb{R}$ or an open interval to $\mathbb{R}$). Then f is continuous at a point x if and only if, given any $\epsilon > 0$, there exists a $\delta > 0$ such that for any t such that $|t - x| < \delta$, we have $|f(t) - f(x)| < \epsilon$.*

Proof. First let us assume that f is continuous. The proof is by contradiction. So, assume that we are give an $\epsilon > 0$ and assume that there is no $\delta > 0$ satisfying the above condition. Then for any $n \in \mathbb{N}$, there exists an x_n such that $|x_n - x| < 1/n$ and $|f(x_n) - f(x)| \geq \epsilon$. Consider the sequence $\{x_n\}$. It is immediate that $\{x_n\}$ is a CS and $\lim x_n = x$. Since f is continuous, we must have $\{f(x_n)\}$ a CS and $\lim f(x_n) = f(x)$. But, this means that for all $n >> 0$, $|f(x_n) - f(x)| < \epsilon$, which contradicts the previous inequality.

To prove the converse, assume we are given a CS $\{x_n\}$ with $\lim x_n = x$. Given an $\epsilon > 0$, we have a $\delta > 0$ so that for all t with $|t - x| < \delta$, $|f(t) - f(x)| < \epsilon/2$. Since $\lim x_n = x$, there exists an N so that for all $n \geq N$, $|x_n - x| < \delta$. So, if $n, m \geq N$, we get,

$$\begin{aligned} |f(x_n) - f(x_m)| &= |f(x_n) - f(x) + f(x) - f(x_m)| \\ &\leq |f(x_n) - f(x)| + |f(x_m) - f(x)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

This shows that $\{f(x_n)\}$ is a CS, proving continuity. □

We next prove the last theorem for this topic, used often by all of you in your calculus classes.

Theorem 7.2 (Intermediate Value Theorem)**.** *Let f be a continuos function (on $\mathbb{R}$ or in the closed interval $[a, b]$) where $a < b$. Then given any real number u between $f(a)$ and $f(b)$, there exists a real number $a \leq c \leq b$ such that $f(c) = u$.*

Proof. If $f(a) = f(b)$, then $u = f(a)$ and there is nothing to prove. So, we may assume $f(a) \neq f(b)$. So, we will assume that $f(a) < f(b)$, the case $f(a) > f(b)$ being similar. We may also assume that $f(a) < u < f(b)$.

We consider the following set.

$$S = \{x \in [a, b] \mid \forall z \in [a, x], f(z) \leq u\}.$$

Notice that S is bounded (below by a and above by b) and it contains a and hence non-empty. Hence by theorem 6.2, it has a supremum, which we call c. Since $a \leq c \leq b$, if we show that $f(c) = u$, we would have proved the theorem.

First we show that $f(c) \leq u$. The proof is by contradiction. So, assume that $f(c) > u$. Notice that $c > a$, since otherwise, $c = a$ and then $f(c) = f(a) < u$. Let $0 < \delta = c - a$. Then for all $n >> 0$, natural numbers we have $a \leq c - 1/n < c$. Since c was the supremum of the set S, there must exist an $x_n \in S$ with $c - 1/n < x_n \leq c$. It is easy to see that $\{x_n\}$ is a CS and $\lim x_n = c$. So, by continuity we have $\lim f(x_n) = f(c)$. But, since $f(x_n) \leq u$ becuase $x_n \in S$, we see that u is an upper bound for the set $\{f(x_n)\}$ and hence by exercise 8, $f(c) \leq u$. This is a contradiction.

Next we show that the assumption $f(c) < u$ also leads to a contradiction and then we would have proved the result. So, assume that $f(c) < u$. Choose an $\epsilon > 0$ such that $f(c) + \epsilon < u$. (For example, you could choose $\epsilon = (u - f(c))/2$). Then there exists a $\delta > 0$, by continuity of f such that for all t with $|t - c| < \delta$, $|f(t) - f(c)| < \epsilon$. We will then show that, if $t \leq c + \delta/2$, then $f(t) \leq u$. If $t < c$, there exists an $x \in S$ such that $t < x \leq c$ and we know that for any such t, $f(t) \leq u$ by defintion of our set S. If $t = c$, this is our assumption. If $c < t \leq c+\delta/2$, we know that $|f(t)-f(c)| < \epsilon$ and thus $f(t)-f(c) < \epsilon$. This means $f(t) < f(c) + \epsilon < u$. Thus, we see that $c + \delta/2 \in S$ by our defintion of the set S. But, then $c < c + \delta/2$ can not be a supremum for S, since it is not an upper bound for S. This contradcition proves the theorem.

□